Volume 33

Advances in

Child Development and Behavior

EDITOR Robert V. Kail

ADVANCES
IN CHILD DEVELOPMENT
AND BEHAVIOR

Volume 33

Contributors to This Volume

Martha W. Alibali

Roberto Araya

Arthur J. Baroody

Kimberly Dadisman

Lisa Flook

Andrew J. Fuligni

LouAnn Gerken

Kelly S. Mix

Florrie Ng

Kim M. Pierce

Bruce F. Pennington

Eva M. Pomerantz

Soo Hyun Rhee

Catherine M. Sandhofer

Robert Siegler

Deborah Lowe Vandell

Qian Wang

Erik Willcutt

Amanda L. Woodward

ADVANCES IN CHILD DEVELOPMENT AND BEHAVIOR

edited by

Robert V. Kail
Department of Psychological Sciences
Purdue University
West Lafayette IN 47907 USA

Volume 33

2005

ELSEVIER
ACADEMIC
PRESS

Amsterdam – Boston – Heidelberg – London – New York – Oxford
Paris – San Diego – San Francisco – Singapore – Sydney – Tokyo

ELSEVIER B.V.
Radarweg 29
P.O. Box 211, 1000 AE
Amsterdam, The Netherlands

ELSEVIER Inc.
525 B Street, Suite 1900
San Diego, CA 92101-4495
USA

ELSEVIER Ltd
The Boulevard, Langford Lane
Kidlington, Oxford OX5 1GB
UK

ELSEVIER Ltd
84 Theobalds Road
London WC1X 8RR
UK

First edition 2005

ISBN: 0-12-009733-8
ISSN: 0065-2407 (Series)

⊗ The paper used in this publication meets the requirements of ANSI/NISO Z39.48-1992 (Permanence of Paper).
Printed in The Netherlands.

Contents

What Develops in Language Development?

LOUANN GERKEN

The Role of Children's Competence Experiences in the Socialization Process: A Dynamic Process Framework for the Academic Arena

EVA M. POMERANTZ, QIAN WANG AND FLORRIE NG

The Infant Origins of Intentional Understanding

AMANDA L. WOODWARD

Analyzing Comorbidity

BRUCE F. PENNINGTON, ERIK WILLCUTT AND SOO HYUN RHEE

This Page Intentionally Left Blank

Contributors

Numbers in parentheses indicate the pages on which the authors' contributions begin.

MARTHA W. ALIBALI

Department of Psychology, University of Wisconsin, Madison, WI 53706, USA (79)

ROBERTO ARAYA

Department of Psychology, Carnegie Mellon University, Pittsburgh, PA 15213, USA (1)

ARTHUR J. BAROODY

Department of Curriculum and Instruction, College of Education, University of Illinois, Champaign, IL 61820, USA (305)

KIMBERLY DADISMAN

Department of Educational Psychology, University of Wisconsin, Madison, WI 53706, USA (43)

LISA FLOOK

Department of Psychology, University of California, Los Angeles, CA 90095, USA (125)

ANDREW J. FULIGNI

Department of Psychology, University of California, Los Angeles, CA 90095, USA (125)

LOUANN GERKEN

Departments of Psychology and Linguistics, University of Arizona, Tucson, AZ 85721, USA (153)

KELLY S. MIX

Department of Psychology and Program in Cognitive Science, Indiana University, Bloomington, IN 47405, USA (305)

FLORRIE NG

Department of Psychology, University of Illinois, Champaign, IL 61820, USA (193)

BRUCE F. PENNINGTON

Department of Psychology, University of Denver, Denver, CO 80208, USA (263)

KIM M. PIERCE

Department of Educational Psychology, University of Wisconsin, Madison, WI 53706, USA (43)

EVA M. POMERANTZ

Department of Psychology, University of Illinois, Champaign, IL 61820, USA (193)

SOO HYUN RHEE
Department of Psychology, University of Colorado, Boulder, CO 80309, USA (263)

CATHERINE M. SANDHOFER
Department of Psychology, University of California, Los Angeles, CA 90095, USA (305)

ROBERT SIEGLER
Department of Psychology, Carnegie Mellon University, Pittsburgh, PA 15213, USA (1)

DEBORAH LOWE VANDELL
Department of Educational Psychology, University of Wisconsin, Madison, WI 53706, USA (43)

QIAN WANG
Department of Psychology, University of Illinois, Champaign, IL 61820, USA (193)

ERIK WILLCUTT
Department of Psychology, University of Colorado, Boulder, CO 80309, USA (263)

AMANDA L.WOODWARD
Department of Psychology and Committee on Human Development, University of Chicago, Chicago, IL 60637, USA (229)

Preface

Advances in Child Development and Behavior is designed to provide scholarly technical articles and to provide a place for publication of scholarly speculation. In these critical reviews, recent advances in the field are summarized and integrated, complexities are exposed, and fresh viewpoints are offered. These reviews should be useful not only to the expert in the area but also to the general reader. No attempt is made to organize each volume around a particular theme or topic. Manuscripts are solicited from investigators conducting programmatic work on problems of current and significant interest. The editor often encourages the preparation of critical syntheses dealing intensively with topics of relatively narrow scope but of considerable potential interest to the scientific community. Contributors are encouraged to criticize, integrate, and stimulate, but always within a framework of high scholarship.

Although appearance in the volumes is ordinarily by invitation, unsolicited manuscripts will be accepted for review. All papers—whether invited or submitted—receive careful editorial scrutiny. Invited papers are automatically accepted for publication in principle, but usually require revision before final acceptance. Submitted papers receive the same treatment except that they are not automatically accepted for publication even in principle, and may be rejected.

I acknowledge with gratitude the aid of my home institution, Purdue University, which generously provided time and facilities for the preparation of this volume.

Robert V. Kail

A COMPUTATIONAL MODEL OF CONSCIOUS AND UNCONSCIOUS STRATEGY DISCOVERY

Robert Siegler and Roberto Araya

DEPARTMENT OF PSYCHOLOGY, CARNEGIE MELLON UNIVERSITY,
PITTSBURGH, PA 15213, USA

I. Introduction

Attaining a deep understanding of change mechanisms is the holy grail of developmental psychology: a profound goal, not fully attainable but worth pursuing nonetheless. Among the obstacles that prevent full realization of the goal are the impossibility of directly observing change mechanisms, the multiple levels and time grains at which such mechanisms operate, and the sheer complexity of the developmental process. Nonetheless, pursuit of the goal is worth the effort because of what we learn along the way.

One major purpose of this chapter is to advocate a triangulation strategy for attaining a better understanding of change mechanisms. This triangulation strategy involves going back and forth among traditional studies of age-related change, microgenetic studies of children's learning, and computer simulations that generate the changes documented in the other two approaches. Two examples are presented to illustrate this triangulation strategy. One describes

Advances in Child Development and Behavior
R Kail (Editor)

acquisition of the min strategy for solving simple addition problems; the other describes acquisition of an insight strategy for solving more complex arithmetic problems.

The other major purpose of the chapter is to describe a new computational model of conscious and unconscious strategy discovery. In addition to being a crucial component of one of the examples of the triangulation strategy, this simulation significantly extends previous models of strategy choice and discovery. In particular, it generalizes the principles embodied by previous simulations (e.g., Siegler & Shipley, 1995; Shrager & Siegler, 1998) to a new discovery and age group, and adds several basic cognitive mechanisms, including priming, forgetting, and controlled attention, that are likely to be important for many types of discovery.

The large majority of studies of cognitive development have been devoted to describing age-related changes. Most of these studies have used cross-sectional methods; a smaller number have used longitudinal approaches. The studies of age-related change have succeeded in providing excellent descriptions of many aspects of cognitive growth. However, they have told us little about *how* children acquire the competencies. Understanding the acquisition process requires different methods, in particular, empirical methods that examine learning while it is occurring and formal modeling methods that simulate changes with age and learning.

Each of these three approaches—descriptions of age-related change, descriptions of learning, and formal modeling—provides unique information critical to a well-grounded account of developmental change. The descriptions of age-related change can establish key developmental phenomena that need to be explained. The studies of children's learning—using trial-by-trial analyses of changing performance—can provide detailed depictions of phenomena that arise during the course of acquisition. The formal models of development can present mechanisms sufficient to generate the observed changes in psychologically plausible ways.

In addition to providing unique information, each of these approaches also often raises questions that only use of the other two methods can answer. Empirical observations of age-related changes raise questions concerning how children get from here to there, that is questions concerning learning. Empirical observations of learning raise questions concerning whether the changes observed under the particular conditions of the learning experiment also characterize age-related changes. Both the studies of learning and the studies of age-related growth raise questions concerning mechanisms that might produce the observed changes, questions that can be answered at least in part by formal models of change mechanisms. The formal models, in turn, raise questions of whether children would behave in the ways that the model does under previously unexamined circumstances, questions that can be answered through studies of learning and age-related change. This symbiotic relation among the three methods is illustrated in the remainder of this chapter.

II. Example 1: Simple Addition

A. BASIC DEVELOPMENTAL PHENOMENA

Cross-sectional observations (e.g., Fuson, 1982) have established that children use a variety of strategies to solve simple addition problems and that the mix of strategies changes with age.[1] One common approach, especially among preschoolers, is the *sum strategy*, which typically involves putting up fingers on one hand to represent the first addend, putting up fingers on the other hand to represent the second addend, and then counting all of the fingers. A child using the sum strategy would solve $2 + 4$ by putting up two fingers, putting up four fingers, and then counting "1, 2, 3, 4, 5, 6." A second common approach is the *min strategy*, which involves counting up from the larger addend. A child using the min strategy would solve $2 + 4$ by counting "5, 6" or "4, 5, 6." A third common approach is *retrieval* of the answer from memory. Children also use a variety of other addition strategies, including *decomposition* (solving problems such as $9 + 4$ by thinking, "$10 + 4 = 14$, $14 - 1 = 13$"), and *guessing*.

Early models portrayed development of arithmetic skills as involving a shift from initial reliance on the sum strategy to later reliance on the min strategy to yet later reliance on retrieval (Groen & Parkman, 1972; Ashcraft, 1987). More detailed, observations, however, revealed that although there is a general trend in this direction, people of all ages from 4 to 70 years use multiple addition strategies (Siegler & Shrager, 1984; Geary & Wiley, 1991; LeFevre, Sadesky, & Bisanz, 1996). The strategies that they use vary with the problem that is presented, but even when the same single-digit addition problem is presented twice in close proximity, roughly one-third of children use a different strategy on the second presentation than the first (Siegler & Shrager, 1984). This variability is not attributable to learning; on almost half of the pairs of trials where one strategy was more advanced than the other, the second presentation of the problem elicited the less advanced strategy.

At the same time, the frequency of use of different strategies does change with age and experience. The sum strategy is the most common approach among preschoolers, but most children stop using it by second or third grade. The min strategy is rarely used by preschoolers, becomes quite common in first and second

[1]The term "strategies" is used throughout this article in the broad sense of any goal-directed, non-obligatory procedure, rather than in the more restricted sense of a conscious, rationally chosen procedure.

grade, and gradually becomes less frequent after that. Retrieval steadily increases from the preschool period to adulthood, though even college students continue to use the min strategy and decomposition on 10–30% of single-digit problems (Geary, 1994; LeFevre, Sadesky, & Bisanz, 1996). In addition to these changes in relative frequency of strategy use, execution of all of the strategies becomes faster and more accurate with age and experience.

This knowledge of diverse strategies allows children to adapt to the demands of problems and situations. Indeed, one of the most striking findings to emerge from cross-sectional studies of children's addition is that the children's choices among strategies are highly adaptive from the preschool years onward. The choices are adaptive in the sense that children use the fastest and least effortful strategy consistent with their achieving a high degree of accuracy on the particular problem. Thus, on the simplest problems, such as 2 + 1, even 4- and 5-year-olds generally retrieve the answer from memory, whereas on problems that are more difficult for preschoolers, such as 4 + 5, they most often use the sum strategy (Siegler & Shrager, 1984). Similarly, first graders use the min strategy most often on problems such as 2 + 9, on which the smaller addend is small and the distance between addends is large, presumably because use of the min strategy on such problems is fast and accurate in absolute terms (because it requires little counting) and because it is especially advantageous on such problems relative to the main alternative, the sum strategy (Siegler, 1987).

With age and experience solving arithmetic problems, children's choices of strategies become increasingly adaptive, in the sense that the strategy choices are calibrated increasingly precisely to the characteristics of the problems (Lemaire & Siegler, 1995). Moreover, execution of each strategy becomes increasingly fast and accurate (Lemaire & Siegler, 1995). Thus, the development of addition skills reflects discovery of new strategies, increasing use of the relatively advanced approaches from among the strategies that children know, improved execution of strategies, and increasingly adaptive choices among strategies.

B. A MICROGENETIC ANALYSIS OF DISCOVERY OF THE MIN STRATEGY

Studies of age-related changes in addition yielded a compelling general description of development in this domain. However, the studies were silent on the issue of how children acquire new strategies that expand their problem-solving skills. Both theoretical and methodological factors limited progress on this question. Until recently, most major theories of cognitive development, including Piagetian, neo-Nativist, and theory–theory approaches, focused on age-related change and placed little emphasis on learning. Consistent with this theoretical emphasis, the large majority of methods used to study the age-related changes spaced observations of emerging competence too far apart in time to yield detailed information about the learning process. The theoretical and methodological approaches were mutually reinforcing; there was little point to developing methods for studying issues that were not viewed as essential, and there was little point to theorizing about issues that could not be studied well empirically.

The advent of microgenetic methods changed the situation by making possible direct examination of acquisition of rules, strategies, and representations. As noted by Siegler and Crowley (1991), microgenetic methods are defined by three main properties:

(1) Observations span the period of rapidly changing competence.
(2) Within this period, the density of observations is high, relative to the rate of change.
(3) Observations are analyzed intensively, with the goal of inferring the representations and processes that gave rise to them.

The second property is especially important. Assessing changing competence on a trial-by-trial basis during the period of rapid change provides the temporal resolution needed to understand the acquisition process. If discoveries were usually made in the most straightforward way imaginable, such dense sampling of ongoing changes would be unnecessary. We could examine children's thinking before and after the discovery, identify the shortest path between the two states, and infer that children moved directly from the one to the other. Such straightforward paths of change are rare, however. Neither the types of experiences that lead up to discoveries nor children's emotional reaction to the discoveries nor the breadth of generalization of the discovery can be intuited—all must be observed while they are occurring. By revealing what problems and experiences immediately preceded discoveries, how children's reactions to discoveries vary, and how quickly and broadly they exhibit new approaches, microgenetic methods advance understanding of change beyond that possible through intuition or speculation.

Siegler and Jenkins (1989) performed this type of microgenetic study on children's discovery of the min strategy. Prior cross-sectional studies indicated that children typically start using the min strategy in first grade. Therefore, the participants chosen for the microgenetic study were kindergartners who knew how to add by using the sum strategy and retrieval but who did not yet know the min strategy (as indicated by their pretest performance on a set of problems on which the min strategy would have been advantageous).

The kindergartners were given 11 weeks of practice solving addition problems. During this period, the children were presented roughly three sessions per week, seven problems per session. Strategy use was assessed on each trial through a combination of videotapes of overt behavior and asking children immediately after each trial, "How did you solve that problem?" In Weeks 1–7, children were presented simple problems in which both addends were five or less. During this time, most of the children discovered the min strategy in the sense of using it at least once, but none of them used it very often. Therefore, in Week 8, children were presented challenge problems, items such as 2 + 23, with one small addend and one very large one. These problems presented both a carrot and a stick for use of the min strategy, in the sense that they could be solved relatively easily using that approach but were very hard (i.e., impossible) for the kindergartners to solve using the main alternative, the sum strategy. Finally, in Weeks 9–11, children were presented a mix of small addend problems, challenge problems, and in-between problems such as 9 + 3.

Eight main findings arose from the study:

(1) Most children used multiple strategies, the most common of which are described in Table I.

(2) Children chose adaptively among strategies; they used each approach most often on the types of problems on which that approach was most effective relative to available alternatives.

(3) The large majority of kindergartners were able to discover the min strategy from just being given practice solving addition problems.

(4) Discovery of the *shortcut sum strategy* appeared to be transitional to discovery of the min strategy. As shown in Table I, this procedure was like the already-known sum strategy, in that counting started with "1." However, it was like the soon-to-be-discovered min approach, in that it involved counting each number only once, rather than twice as with the sum strategy. Thus, the shortcut sum strategy made sense as a transitional approach. Consistent with the view that it was transitional, most children generated the shortcut sum shortly before they discovered the min strategy.

TABLE I

Preschoolers' Main Addition Strategies

Strategy	Typical use of strategy to solve 3 + 5
Sum	Put up three fingers, put up five fingers, count fingers by saying "1,2,3,4,5,6,7,8"
Shortcut sum	Say "1,2,3,4,5,6,7,8," perhaps simultaneously putting up one finger on each count beyond 5
Min	Say "5,6,7,8," or "6,7,8," perhaps simultaneously putting up one finger on each count beyond 5
Retrieval	Say an answer and explain it by saying "I just knew it"
Decomposition	Say "3 + 5 is like 4 + 4, so it is 8"

(5) Discoveries of new strategies often occurred without existing strategies having failed. The discoveries occurred following correct answers as well as errors, and on easy as well as hard problems.
(6) Generalization of the newly discovered min strategy proceeded slowly. The slow generalization was evident even among children who explicitly stated that the min strategy reduced the amount of counting below that required by the sum strategy or who praised themselves for using it by saying, "smart answer" (Siegler & Jenkins, 1989, p. 80).
(7) For children who had already discovered the min strategy, generalization increased greatly with presentation of the challenge problems in Week 8. The generalization continued to increase on the wider range of problems presented in Weeks 9–11. However, encountering the challenge problems did not lead to discoveries among children who had not yet discovered the min approach; they tended to give up when faced with them.
(8) Finally, and particularly surprising, discovery did not require trial and error. The kindergartners generated useful new strategies without ever trying conceptually flawed approaches, such as counting the first addend twice.

This last finding was particularly intriguing. Almost everyone's intuitive view of discovery, one consonant with formal models of discovery (e.g., Newell, 1990; Van Lehn, 1990), is that people initially try flawed strategies and eventually abandon them in favor of correct approaches. Yet, despite solving between 140 and 210 problems and generating at least two novel addition strategies, none of the children ever generated an illegal strategy. It was not that no illegal strategies were possible; for example, children could have added the first addend twice and ignored the second one or added the first addend to the second one and then added one or the other again. The absence of such incorrect approaches raised the issue of how children were able to invent appropriate new strategies without any trial and error.

To explain the finding, Siegler and Jenkins (1989) proposed that even before children discover the min strategy, they possess a goal sketch, a conceptual structure that indicates the goals that a legal strategy must meet. The hypothesized goal sketch for addition indicated that legal strategies must include procedures for quantifying each addend and combining the two addends into a single answer.

This hypothesis motivated Siegler and Crowley (1994) to perform a new study to test whether children possess such a goal sketch even before they discover the min strategy. An experimenter asked 5-year-olds to judge the smartness of three addition procedures that a puppet executed: the sum strategy, which all of the children already used in their own problem solving; the min strategy, which some of the children used and some did not; and counting the first addend twice, which none of the children used. The question was whether children who did not yet use the min strategy would view it as smarter than counting the first addend twice, which they also did not use.

It turned out that these children rated the min strategy as much smarter than counting the first addend twice. In fact, they rated the min strategy as slightly smarter than the strategy they used most often, the sum strategy. This finding led to the conclusion that children possess conceptual understanding of addition akin to the goal sketch before they discover the min strategy and that this understanding helps them avoid trial and error in the discovery process.

C. A FORMAL MODEL OF DISCOVERY OF THE MIN STRATEGY

These data from Siegler and Crowley (1994), together with the prior cross-sectional and microgenetic findings, provided crucial constraints on a computer simulation. The model that was generated, SCADS (*S*trategy *C*hoice *a*nd *D*iscovery *S*imulation), discovered the min strategy as well as generating improvements in speed, accuracy, and strategy choices (Shrager & Siegler, 1998). Like children, it discovered useful new strategies without trial and error.

SCADS was based on a previous computational model of strategy choice, ASCM (*A*daptive *S*trategy *C*hoice *M*odel), that produced improved speed, accuracy, and strategy choices, but did not discover new strategies (Siegler & Shipley, 1995). ASCM learned by associating problems with answers that were stated on the problems, and by associating the goodness of outcomes generated by each strategy with the problems on which the outcomes were generated and with other similar problems. Goodness of outcomes produced by a strategy reflected the speed and accuracy of solutions generated by it. Because speed and accuracy of strategies increase with use, any given level of speed or accuracy was evaluated more highly if it was produced by a strategy that was relatively novel. The logic underlying the extra strength for novelty was that the human cognitive system might know implicitly that new approaches have greater potential for improvement than better practiced ones that generate comparable performance.

ASCM produced faster and more accurate responses, increasingly frequent choices of the more advanced strategies, adaptive generalization to novel problems, and individual difference patterns that mirrored those of children. However, the associative learning mechanisms within the model did not allow discovery of new strategies.

SCADS (Shrager & Siegler, 1998) surmounted this problem; it showed the same type of learning as ASCM, and also discovered new strategies. Key to these expanded capabilities was inclusion of metacognitive as well as associative learning mechanisms. At the start of the simulation, SCADS, like the children in Siegler and Jenkins (1989), used the sum strategy and retrieval to solve single-digit addition problems. By solving problems, the model learned which strategy to choose on which problem. That part of the model was identical to

ASCM, and the same associative mechanisms produced improvements in speed, accuracy, and strategy choices in both models.

New to SCADS was a metacognitive system that generated new strategies. The metacognitive system was composed of three components: the attentional spotlight, strategy-change heuristics, and goal sketch filters. The attentional spotlight increased the cognitive resources devoted to execution of poorly learned strategies. Its operation was analogous to the way in which people focus attention on execution of unpracticed procedures in order to increase the likelihood of executing them correctly. Within SCADS, each strategy was made up of a sequence of operators. The decision of whether to devote cognitive resources to supervising execution of a particular operator was based on the current strength of the operator relative to a threshold that varied randomly from trial-to-trial. Because each operator was at first quite weak, cognitive resources were initially required to insure its appropriate execution. The more often an operator was used, the less often cognitive resources needed to be focused on its execution and the more often they could be devoted to discovering new strategies.

As cognitive resources were freed, SCADS allocated them to the strategy-change heuristics, the second component of the metacognitive system. These heuristics operated on the traces of the operations that were used to solve particular problems. SCADS included two strategy-change heuristics: (1) If a redundant sequence of behavior is detected, then delete one of the two sets of operators that caused the redundancy; and (2) If statistics on a strategy's speed and accuracy show greater success when the strategy is executed in a particular order, then create a version of the strategy that always uses that order (as opposed to the initial procedure of arbitrarily choosing which addend to quantify first).

These heuristics enabled the metacognitive system to propose a variety of strategies, some valid, others flawed. To help SCADS avoid executing invalid strategies, the metacognitive system included a third component, the goal sketch filters. These were the two standards hypothesized by Siegler and Jenkins (1989) to be essential for legitimate addition strategies and to provide criteria against which proposed new strategies can be evaluated. One filter required that both addends be represented; the other required that the sum include the representations corresponding to both addends.

SCADS demonstrated all eight of the major phenomena of strategy choice and strategy discovery listed above. It used multiple strategies throughout its run. It chose adaptively among strategies from the beginning of its run, and the choices became increasingly adaptive with experience. It discovered the min strategy on all 30 of its runs. It consistently discovered the shortcut sum strategy before the min strategy. Its discoveries followed correct as well as incorrect performance. It was slow to generalize newly discovered strategies. It never executed illegal strategies. It responded to presentation of a block of challenge problems by increasing its use of the min strategy when that strategy had already been discovered, and maintained the greater use on the mixed set of problems that followed.

SCADS also illustrated mechanisms through which the shortcut sum approach could be transitional to the min strategy. Through experience using the shortcut sum approach, which was invariably discovered before the min strategy, SCADS learned that counting objects once could lead to correct answers. The same experience also provided data from which the model could learn that starting counting with the larger addend was more efficient than counting from the smaller addend, because starting a new count at a larger number was faster and more accurate than starting it at a smaller number. It was a relatively small step, accomplished via redundancy elimination, to simply represent the larger addend and count on from it (i.e., to use the min strategy). These results indicated that the combination of associative and metacognitive processes embodied in SCADS was sufficient to generate adaptive choices among existing strategies, increasing reliance on the more effective strategies, and discovery of new strategies for solving simple addition problems. This completed one round of the triangulation strategy as applied to basic addition; new findings will in all likelihood motivate further rounds.

III. Example 2: The Inversion Problem

Inversion is the principle that adding and subtracting the same number leaves the original value unchanged. Knowing this principle allows solution of *inversion problems*, problems on which the same number is added and subtracted, for example $18 + 16 - 16$. Such problems can be solved through adding and subtracting the numbers or through recognizing that no addition and subtraction is needed and therefore simply stating the remaining number. Research on the inversion problem has focused on the age at which children begin to use the inversion principle, the influence of experimental conditions on its use, changes in its use with age and experience, and how children discover the principle.

A. BASIC DEVELOPMENTAL PHENOMENA

Starkey and Gelman (1982) reported that children as young as 3 years can solve simple inversion problems, in particular $1 + 1 - 1$ and $1 - 1 + 1$. As the investigators noted, however, the data (percentage of correct answers) did not allow determination of whether children were adding and subtracting all three numbers or whether they were using the shortcut.

To determine which of these strategies children were using, Bisanz and LeFevre (1990) devised a chronometric assessment technique that involved presenting two types of arithmetic problems: inversion problems ($A + B - B$) and standard problems ($A + B - C$). The size of B was varied on both types of problems. The logic was that if children used the shortcut strategy to solve inversion problems, the size of B would not affect solution times on them (because B would not be added and subtracted), but if children added and subtracted B, solution times would increase with the size of B, just as they would on standard problems.

Bisanz and LeFevre (1990) applied this logic to analyzing the performance on inversion and standard problems of first, second, and fourth graders and adults. Group-level analyses indicated that participants of all ages solved inversion problems faster than standard ones. However, analyses of individual performance revealed that between first and fourth grade, only around 40% of children relied on the shortcut approach. Other children added and subtracted all numbers (*the computation strategy*). Yet others used a third approach, which Bisanz and LeFevre labeled the *negation strategy*. Negation involved adding $A + B$, typically by counting up on one's fingers, but then putting down all of the fingers simultaneously and saying A. As would be expected from the nature of this procedure, children using it solved $A + B - B$ problems faster when B was small than when it was large, but the difference in times was less than when children used the computation strategy (because they only added B rather than adding and subtracting it). Use of the shortcut remained surprisingly limited at least through fourth grade.

Since Bisanz and LeFevre's (1990) pioneering study, several experiments have demonstrated rudimentary understanding of the inversion principle on nonverbal tasks even before children receive math instruction in school. In Klein and Bisanz (2000), an experimenter showed preschoolers a small number of chips and had children place the same number of chips on their own mat. Then the experimenter covered the chips on her mat, added some chips to her now hidden collection, and subtracted some chips from it. Children could see the number of chips that were added and subtracted but could not see the total number in the set. The children's task was to have a number of chips on their own mat that equaled the number in the experimenter's collection. Mean accuracy and solution times to solve inversion and standard problems did not differ for inversion and standard problems, which suggested a lack of understanding of the inversion problem. However, observation of the overt strategies that children used revealed some use of negation, and analyses of solution times on trials on which no overt behavior was produced revealed that on those trials, solutions were faster on inversion than on standard problems. Additionally, on about 30% of trials without overt behavior, children spontaneously reported that counting was unnecessary because B and C were identical.

Using different nonverbal tasks, Bryant, Christie, and Rendu (1999) and Rasmussen, Ho, and Bisanz (2003) and provided additional evidence that at least some preschoolers have rudimentary understanding of the inversion principle. In Rasmussen, Ho, and Bisanz (2003), preschoolers answered correctly more often, 49% vs. 38%, on inversion than on standard problems. As in the Klein and Bisanz study, it was uncertain whether the enhanced accuracy was due to use of the negation strategy, the shortcut strategy, or a combination of approaches. Moreover, the inconsistent results with solution times, and the relatively high frequency of errors in all of these studies, suggested that although preschoolers may possess a rudimentary understanding of the inversion principle, their reliance on it is sporadic.

As this last statement suggests, use of the shortcut strategy requires not only knowing the inversion principle but also applying it on the particular problem. This application seems likely to be influenced by the probability of encountering inversion problems. Presumably, the more consistently and frequently inversion problems are presented, the stronger the activation of the shortcut strategy, and the more often it will be chosen over computational approaches. Consistent with this hypothesis, 8- to 10-year-olds who were presented 100% inversion problems were more likely to be classified as using the shortcut strategy on inversion problems than were peers who were presented 50% inversion problems (Stern, 1992, 1993).

Reliance on the inversion principle also increases with age and experience. For example, it was used more often by first graders than by preschoolers in Rasmussen, Ho, and Bisanz (2003), despite the problems presented to first graders utilizing larger addends and subtrahends. At both age levels, however, it was uncertain whether children were using the shortcut, the negation strategy, some other approach, or a mix of approaches to solve the inversion problems. It also remained uncertain how children discover the shortcut strategy. To answer these questions, and also to determine whether there are unconscious as well as conscious strategy discoveries, Siegler and Stern (1998) performed a microgenetic study of strategy discovery on the inversion problem.

B. A MICROGENETIC STUDY OF DISCOVERY OF THE SHORTCUT STRATEGY

1. Historical Background

For reasons that will be discussed below, examining the inversion problem microgenetically presented an opportunity to address in an unusually direct way an issue of historical as well as contemporary interest: the role of consciousness in generating insights. Long before there was a scientific field of psychology, philosophers, mathematicians, and scientists speculated about the role of consciousness in generating insights. Archimedes' experience of stepping into the bath, seeing the water rise, and exclaiming "Eureka," is probably the prototypic insight: a sudden change from not knowing a problem's solution to being consciously aware of it. Other thinkers emphasized the contribution of unconscious processes to their discoveries. Perhaps the prototypic account of this type is Kekule's dream of intertwined snakes, which led him to "see" the structure of the benzene ring (Gruber, 1981).

Accounts of these two discoveries share the view that whatever led up to it, the insight itself was sudden. Other accounts differ, though. Wittgenstein (1969), for example, compared acquisition of new ideas to a sunrise; the amount of light increases steadily over a prolonged period, but our experience is that the new day suddenly "dawns."

Psychologists' arguments about whether discoveries are conscious or unconscious, and whether they are generated gradually or abruptly, have shown a similar division. A number of theorists, both venerable and modern, have depicted insights as conscious from the start (e.g., Buehler, 1908; Lewin, 1935; Gick & Lockhart, 1995). These theorists frequently draw analogies to vision, in which people become aware of the identity of a newly presented object almost instantly. Other psychologists have depicted insights as arising initially at an unconscious level, and only later becoming conscious (e.g., Maier, 1931; Karmiloff-Smith, 1992; Kuhn *et al.*, 1995). These theorists postulate a process of representational redescription, in which insight initially arises at an unconscious level and only later becomes conscious. The issue is difficult to resolve empirically, because of the difficulty of independently measuring conscious and unconscious understanding simultaneously. Simply put, how can we know that people have had an insight if they do not tell us that they had it?

Studies in which gesture and speech on a single trial are used as indicators of understanding provide relevant evidence. When people are asked to explain their reasoning on mathematics and science problems, they frequently produce hand gestures that reflect more advanced understanding than is evident in their verbal explanations of how they solved the problem (Perry, Church, & Goldin-Meadow, 1988; Albali, 1999; Goldin-Meadow & Alibali, 2002). For example, when 5-year-olds see water poured from a typical glass into a taller, thinner one, and are asked whether the amount of water is the same or different than before, they typically say that the tall thin glass has more water and make vertical hand gestures to illustrate their point. However, the hand gestures of some of the children also indicate attention to the beakers' cross-sectional area, and children who produce such hand gestures are more likely to learn from instruction than are children whose hand gestures only indicate attention to the relative heights of the liquid columns (Church & Goldin-Meadow, 1986). This finding may mean only that gestures are a particularly sensitive index of early understanding. However, it may have a more general meaning: the gestures may reflect an unconscious understanding of the problem, and this unconscious understanding may represent a way station on the road to conscious understanding.[2]

The issue of whether insights arise abruptly or gradually also has a long and controversial history within psychology. Theorists who view insight as a conscious process usually also believe that insights occur abruptly (e.g., Kohler, 1925; Duncker, 1945; Perkins, 1995). Theorists who view insight as involving unconscious as well as conscious processes usually believe that the insights emerge gradually (e.g., Yaniv & Meyer, 1987; Ohlsson, 1992; Isaak & Just, 1995). However, other combinations of views have also been voiced. For example, Simonton (1995) proposed that insights arise from a progression of unconscious images that culminate in a sudden conscious discovery, and Simon (1973) proposed that insights arise gradually through a logical, conscious, sequence of steps.

Perhaps the strongest evidence favoring the "insights arise abruptly" perspective comes from studies in which participants are periodically asked during the problem-solving process to rate how close they are to the solution. Such ratings have been obtained on both typical problem-solving tasks and on insight problems (i.e., problems whose solution seems to the investigators to require an insight). On the typical problems, ratings of the nearness of the solution rise gradually as the solution draws near. On insight problems, however, the ratings are flat until a solution (correct or incorrect) is found (Metcalfe, 1986; Davidson, 1995). In other words, participants are unaware that they are approaching a solution until they reach it. Although these data are striking, they also are hard to interpret, because of the uncertainties involved in using conscious reports to measure unconscious processes.

Proponents of the alternative, "insights arise gradually" perspective tend to view insight as a product of spreading activation. The strongest support for this stance comes from studies of incubation effects. For example, in Kaplan and Simon (1990), an experimenter presented an insight problem and later planted subtle hints regarding its solution in the subject's environment. These hints increased the likelihood of solution, but the solutions often were generated hours after the hints were encountered. Subjects generally were unaware of the hints they encountered. Kaplan and Simon interpreted the finding as indicating that between the time when people encountered the hint and when they found the solution, they gradually worked through the implications of the hint until the solution arose. Alternatively, however, the hint may have been insufficient to generate the solution until some additional information was encountered in the

[2]Here and elsewhere in this article, we use inability to verbalize as the operational definition of unconscious. There clearly are limits to this operational definition—for example, people who cannot speak can still be conscious—but the criterion seems useful for the populations and tasks of interest in the present context.

environment (as in Archimedes entering the bath), at which point the solution suddenly emerged (Schooler, Fallshore, & Fiore, 1995).

2. Advantages of the Inversion Task

Siegler and Stern (1998) noted that inversion problems presented an opportunity to address these two issues in an unusually direct way. Whether strategies were discovered gradually or abruptly could be examined through trial-by-trial assessments of strategy use. These assessments could reveal whether there were intermediate forms that incorporated parts of the insight before the shortcut emerged. They also could indicate how consistently the shortcut strategy was used following its discovery.

The inversion problem also was useful for studying whether strategies could be discovered unconsciously. The reason was that it allowed independent measurement of conscious and unconscious versions of the insight underlying the shortcut strategy. Conscious use of the shortcut could be assessed through immediately retrospective verbal reports. Young school-age children typically report their use of arithmetic strategies quite accurately, as indicated by converging evidence from accuracy and solution time patterns (Siegler, 1987, 1989).

What made the inversion problem special was that an implicit measure of strategy use also could be obtained: the child's solution time. Ordinarily, the solution time on a given trial is insufficient to infer the strategy that was used on the trial. However, the solution time on a given trial is considerably more useful for inferring strategy use on inversion problems. The reason is that solving the problem via computation generates much slower solution times than solving it by using the shortcut. Consistent with this view, solution times on inversion problems in Siegler and Stern (1998) were bimodally distributed; 92% of times were either fast (4 sec or less) or slow (8 sec or more), with roughly equal percentages in the two categories. Converging evidence from overt behavior in the experiment supported the view that the fast times reflect use of the insight and the slow times the use of computation. Overt computational activity was observed on 80% of trials classified as computation vs. 0% of trials on which children were classified as using the shortcut.

Obtaining both the verbal report and the solution time on each trial made it possible to assess whether children ever used the shortcut strategy unconsciously and, if so, whether use of the unconscious shortcut occurred especially often just before discovery of the conscious version of the shortcut. The operational definition of unconscious strategy use was a solution time too fast to have been generated by young children via computation and a verbal report that did not indicate use of the shortcut. Thus, the second graders in Siegler and Stern (1998) were classified as using the shortcut on each trial on which they generated a time of 4 sec or less and said they used the shortcut, as using the computation strategy on each trial on which they generated a solution time of more than 4 sec and said they computed the answer, and as using the unconscious shortcut on each trial on which their solution time was 4 sec or less but they claimed to have computed the answer.

3. *The Unconscious Activation Hypothesis*

Based on previous research showing unconscious influences on other types of thinking, Siegler and Stern formulated the *unconscious activation hypothesis*: increasing activation of a strategy leads it first to be used at an unconscious level, and then, as the activation increases further, people use it consciously. The straightforward implication of this hypothesis was that the unconscious shortcut would emerge before the conscious version of the strategy.

Two experimental conditions were created to test the unconscious activation hypothesis. Half of the children were randomly assigned to the *blocked problems condition*, in which they were presented inversion problems on 100% of trials. The other children were assigned to the *mixed problems condition*, in which they were presented inversion problems on 50% of trials and standard problems on 50%. The unconscious activation hypothesis predicted that presenting children inversion problems on 100% of trials would lead to a more rapid increase in activation of the shortcut, which in turn would lead to (a) more rapid discovery of the shortcut and unconscious shortcut strategies (discovery after fewer inversion problems); (b) a shorter gap between discovery of the unconscious shortcut and discovery of the shortcut; (c) more consistent use of the shortcut on inversion problems after it was discovered; and (d) greater generalization of the strategy to novel problems of similar appearance, such as $A - B + B$ and $A + B + B$. The hypothesis also predicted that children in both groups would generate the unconscious shortcut strategy before the conscious version of the shortcut.

To test these hypotheses, German second graders were presented eight sessions, one session per week. Session 1 was a pretest, consisting of 10 inversion and 10 standard problems; children who used the shortcut strategy on any trial were eliminated from further participation. In Sessions 2, 3, 4, and 6, children who were randomly assigned to the blocked problems condition received 20 inversion problems. Children who were assigned to the mixed problems condition received 10 inversion problems interspersed with 10 standard problems in those sessions. In Sessions 1, 5, and 7, children in both groups received 10 standard and 10 inversion problems; the idea was to compare the two groups' performance on identical problems at several points during learning. The problems presented in Session 8 were also the same for children in both groups, but they included generalization problems as well as inversion and standard ones. The generalization problems included some that superficially resembled the inversion problems but on which the principle did not apply (e.g., $A + B + B$)

and other problems that differed superficially from the original problems but on which the principle did apply (e.g., A − B + B).

Each of the predictions of the unconscious activation hypothesis was borne out. Relative to children in the mixed problems condition, children in the blocked problems condition discovered the shortcut strategy earlier and used it more often (Figures 1 and 2). Almost 90% of the children discovered the unconscious version of the shortcut before the conscious version. The gap between discovery of the unconscious shortcut and the shortcut was also smaller in the blocked problems condition. Moreover, children in the blocked problems condition generalized the shortcut more often, both correctly and incorrectly.

Examination of strategy use just before and after discovery of the unconscious shortcut and shortcut strategies provided particularly direct support for the unconscious activation hypothesis. Figure 3 illustrates the circumstances surrounding the first use of the unconscious shortcut of children in the blocked problems condition. The trial labeled "0" on the *X*-axis is the trial on which that child first used the unconscious shortcut. Thus, by definition, 100% of children used the unconscious shortcut on Trial 0. The −1 trial for a given child is whichever trial immediately preceded that child's Trial 0; the +1 trial is whichever trial immediately followed the child's Trial 0; and so on.

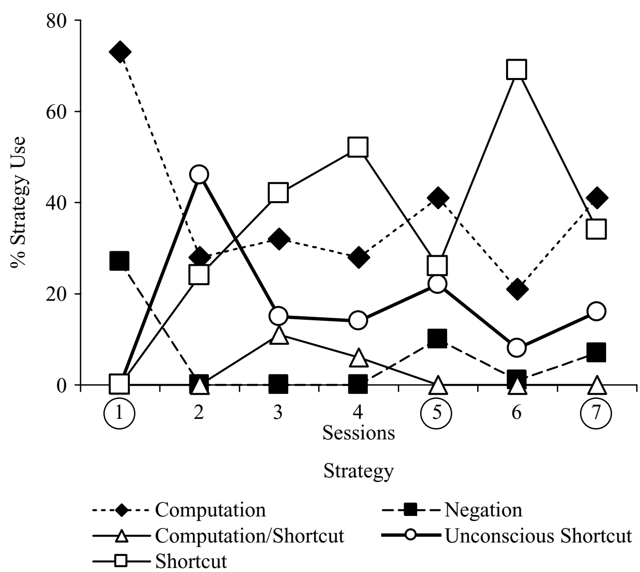

Fig. 1. Changes in children's strategy use over seven sessions: blocked problems condition (data from Siegler & Stern, 1998).

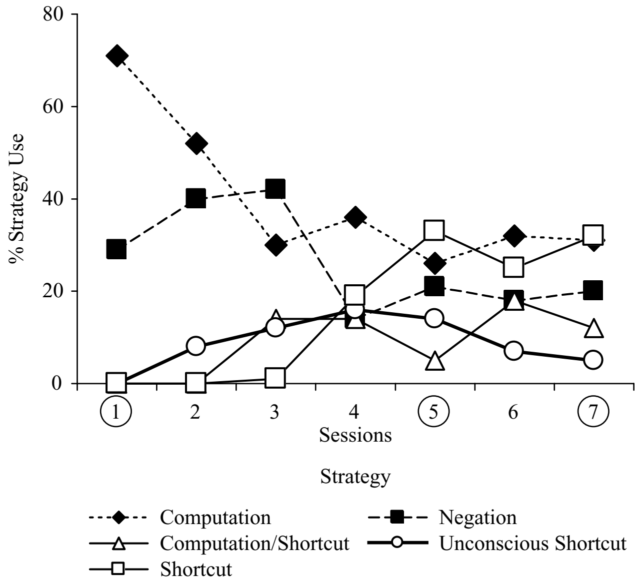

Fig. 2. Changes in children's strategy use over seven sessions: mixed problems condition (data from Siegler & Stern, 1998).

As shown in Figure 3, just before their first use of the unconscious shortcut, children in the blocked problems condition consistently used the computation strategy. After the initial use of the unconscious shortcut, most children continued to use that approach over the next three trials. By the fourth trial after the initial use of the unconscious shortcut, half of the children reported using the shortcut.

Figure 4 shows a parallel analysis centered on children's first use of the shortcut. On the three trials immediately preceding its first use, roughly 80% of children in the blocked problems condition used the unconscious shortcut (as opposed to less than 10% use of this strategy for the study as a whole). After children began to report using the shortcut, they continued to use it quite consistently within that session. However, when they returned a week later for the next session, fewer than 35% used the shortcut on any trial before Trial 5. Thereafter, more children rediscovered the shortcut, and by the end of the session, more than 90% of them were again using it.

Changes in solution times from the trials just before the first use of the unconscious shortcut to the trial on which it was first used suggested that the unconscious shortcut represented a sudden, qualitative shift in thinking. On the three trials immediately before the first use of the unconscious shortcut,

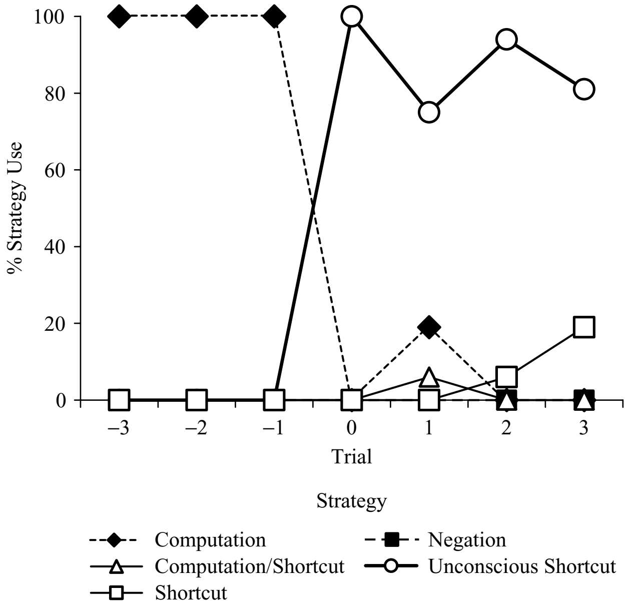

Fig. 3. Children's strategy use on trials immediately before and after first use of unconscious shortcut strategy in blocked problems condition (data from Siegler & Stern, 1998). Note that on this and the other backward trials graphs, the 0 trial indicates each child's first use of the relevant strategy, the − 1 trial is the trial before that, etc.

solution times averaged 12 sec; on its first use, the mean solution time was 2.7 sec. Solution times on subsequent unconscious shortcut trials (and on shortcut trials as well) continued to average between 2 and 3 sec in all sessions. In contrast, no change at all in solution times was evident between use of the unconscious shortcut and use of the shortcut. Thus, although children who used the unconscious shortcut did not report doing anything different, they had already had the insight at the level of nonverbal behavior.

The lack of reporting of the insight on unconscious shortcut trials could not be attributed to the children being generally inarticulate, to the insight being difficult to put into words, or to children's perceptions of social desirability preventing them from reporting an approach that they knew they were using. If those were the reasons for children initially not reporting the shortcut, why would the same children have almost invariably reported using it 3–5 trials later in the same session? Further supporting the view that use of the shortcut was at first unconscious, when children discovered the shortcut in the session following the one in which they initially used it, most again used the unconscious version just before beginning to report its use.

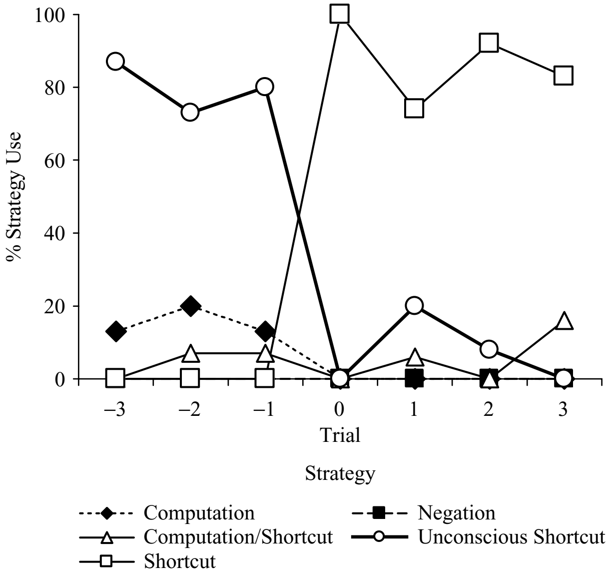

Fig. 4. Children's strategy use on trials immediately before and after first use of shortcut strategy in blocked problems condition (data from Siegler & Stern, 1998).

4. *Findings Constraining Models of Discovery*

The findings about conscious and unconscious strategy discoveries, and the findings relevant to the spreading activation hypothesis, were the highlights. However, a number of other interesting findings also emerged; together, these findings provided the data that a satisfactory model of strategy discovery needed to generate:

(1) Children used five strategies during the experiment. Four have already been noted: the computation, negation, unconscious shortcut, and shortcut strategies. The fifth was the computation/shortcut strategy. This approach involved a solution time greater than 4 sec, but a report of the shortcut strategy. Children generated the computation/shortcut by starting to add A and B, answering "A" before the addition was complete, and saying that they had generated the answer via the shortcut strategy. The strategy was uncommon in the blocked problems condition (3% of trials), but was somewhat more common in the mixed problems condition (9% of trials).

(2) Over sessions, children answered increasingly quickly and accurately, and they shifted their strategy use from predominant use of the computation and negation strategies to increased use of the shortcut strategy.

(3) Almost all children in both the blocked and the mixed problems conditions discovered both the unconscious shortcut and shortcut strategies.

(4) Among children in the blocked problems condition, most first used the computation strategy, then negation, then the unconscious shortcut, and then the shortcut. Half of the children used the computation/shortcut strategy after the shortcut; the others never used it. The order of first use was similar for children in the mixed problems condition, except that most of them used the computation/shortcut strategy before the shortcut.
(5) Relative to children in the mixed problems condition, children who received 100% inversion problems discovered the unconscious shortcut and shortcut strategies earlier, used them more often after their discovery, and generalized them to a greater extent, both appropriately (e.g., to problems of the form $A - B + B$) and inappropriately (e.g., to problems of the form $A - B - B$) in the final session.
(6) Children in both conditions continued to use a variety of strategies in each session after they discovered the shortcut. In the first few trials of each new session, use of the computation and negation approaches predominated. After between 4 and 6 trials, children usually resumed their reliance on the shortcut strategy. The shift occurred somewhat faster in later sessions, but use of the other strategies persisted in all sessions.
(7) Use of the unconscious shortcut was particularly frequent just before the first use of the shortcut in the blocked problems condition.
(8) Children in the mixed problems condition used the negation and computation/shortcut strategies more often than did children in the blocked problems condition.
(9) When children in the mixed and blocked problems groups received a mix of standard and inversion problems (Sessions 1, 5, and 7), shortcut strategy use was similar in the two conditions.

C. MODELING STRATEGY CHOICE AND DISCOVERY ON THE INVERSION PROBLEM

Our goal in modeling these findings was to expand Shrager and Siegler's (1998) SCADS model in ways that reflect generally accepted psychological principles and that would generate the above-listed findings on the inversion problem. Because the new model (SCADS*) has not been reported previously, we present it in some detail here. We first describe the new model's innovations, then describe how these innovations are integrated with its predecessor, SCADS, to yield a functioning simulation, and then describe the results that SCADS* generates.

*1. Innovations Within SCADS**

The basic modeling philosophy embodied in SCADS and its predecessors was to keep the simulations as lean as possible, that is, to include only those mechanisms necessary to generate the relevant data. The reason was not that we believed that the mechanisms embodied within the model were the only mechanisms relevant to strategy choice and discovery but rather that constructing the most parsimonious model that generated the relevant findings would highlight the contributions of the mechanisms that were most important to the outcomes. This philosophy also underlies SCADS*. However, the inversion problem is considerably more complex than the addition task modeled in SCADS, the number of strategies modeled is greater, and performance under two experimental conditions rather than one is simulated. This greater complexity required several additions to the original model. In particular, SCADS* adds the following mechanisms to those included in SCADS:

(1) Controlled attention
(2) Interruption of procedures
(3) Verbalization
(4) Priming
(5) Forgetting
(6) Dynamic feature detection

In this section, we describe these additions.

a. Controlled Attention. Although the capacity to shift attention from one focus to another clearly exercises pervasive effects on human cognition, it was not included in SCADS, because the simulation could generate the relevant data without it. Modeling performance and learning on the inversion problem, however, did require postulating a simple mechanism for shifting attention.

Strategies within SCADS* include two types of operations: arithmetic operations, such as adding A + B, and attention shifts, such as moving the focus of attention from A to B or A to C. The attention shifts are essential, because use of the shortcut strategy requires a different sequence of attentional foci than does the computation strategy. Use of the computation strategy to solve inversion problems that are written from left to right generally involves: attending to the leftmost number (A), attending to the plus between the digits, attending to the middle number (B), adding A + B, attending to the rightmost number (C), etc. However, the shortcut strategy can only be executed by focusing attention on B and C before any arithmetic operation is performed; without such a focus of attention, there is no way to know whether the shortcut is applicable. Thus, controlled attention is essential if SCADS* is to use the shortcut strategy as well as computation.

Attentional shifts are modeled within SCADS* as requiring considerably less time than arithmetic operations. This feature of the model is based on empirical findings that second graders' counting to solve arithmetic problems takes roughly 1 sec/count (Siegler, 1987), whereas shifts of attention occur far more quickly

(Palmer, 1999). This feature is important in determining which strategies are subject to interruptions and when the interruptions occur, as described in Section III.b.

b. Interruption of Procedures. Siegler and Stern (1998) concluded that children used five strategies to solve inversion problems: computation, negation, computation/shortcut, unconscious shortcut, and shortcut. Our modeling efforts, however, changed our perspective on the number of distinct strategies among which children choose. SCADS* generates the same five patterns of behavior as the children in Siegler and Stern (1998), but at the level of mechanisms, there are only two distinct strategies: computation and shortcut.

Within SCADS*, the negation and computation/shortcut approaches, though behaviorally distinct from each other and from the other strategies, reflect interruptions of the computation strategy by the shortcut at different points in the computation process, rather than constituting pre-assembled strategies that could be chosen as a whole at the outset of a problem. The capacity to interrupt ongoing procedures has emerged as an important characteristic of human cognition in several recent theories (Gross *et al.*, 1999; Lieberman, 2000), and it appears to be important for solving inversion problems as well. For example, within SCADS*, the negation strategy comes about through the shortcut strategy interrupting execution of the computation strategy immediately after A and B have been added.

SCADS* incorporates several assumptions regarding when interruptions occur. One is that only relatively slow operations, such as adding and subtracting, can be interrupted. This means that the computation strategy, which includes the relatively prolonged operation of counting, can be interrupted, whereas the shortcut, which does not include any time consuming operation, cannot. Another assumption is that available mental resources increase within a trial as execution of a strategy approaches completion, thus leading to more interruptions occurring toward the end of strategy execution, when the necessary resources are available (for empirical data supporting this assumption, see Kotovsky, Hayes, and Simon (1985)). Yet another influence on when interruptions occur within SCADS* is that interrupting execution of a strategy demands cognitive resources, resulting in there being insufficient resources to support more than one interruption per trial.

Interruption of procedures, like the capacity for controlling attention, is particularly important in accounting for how strategies (defined behaviorally) are generated and in accounting for the order in which the strategies are generated. SCADS* starts with computation and then generates other strategies in the order from latest to earliest interruption of processing within a trial. Thus, negation is generated earlier in the sequence of trials than the computation/shortcut, because it involves an interruption later within the trial on which it occurs, when cognitive resources are at their peak for that trial.

c. Verbalization. Ability to verbally describe a strategy, like all aspects of SCADS*, depends on the relevant activation. The verbalization activation for any given strategy is the product of the maximum time to execute operations within a strategy and the number of times the strategy has been executed. If this product exceeds a threshold that varies randomly from trial-to-trial, the strategy can be verbalized; otherwise, it cannot. The logic underlying this formula is that operations that take longer to execute are easier to describe verbally, and that repeatedly executing a procedure also increases ease of verbalization.

This feature of SCADS* is particularly important for explaining the existence of both conscious (verbalizable) and unconscious (nonverbalizable) versions of the shortcut strategy. The shortcut is used unconsciously at first because the values of both variables that contribute to its verbalization activation—maximum duration of operations and number of uses—are very low. As number of uses of the shortcut increases, its verbalization activation becomes high enough to exceed many thresholds, thus allowing the strategy to be verbalized. In contrast, computation, negation, and the computation/shortcut all can be verbalized from their initial use onward, because all include lengthier operations that allow the verbalization activation to exceed the relevant threshold from the beginning.

d. Priming. SCADS* incorporates two types of priming: priming from the previous trial and priming from other visual locations. First, consider priming from the previous trial. SCADS* incorporates the assumption that the triggering conditions of strategy use include where the person is looking (for empirical support for this assumption, see Richardson and Spivey (2000)). Thus, the strength of each strategy depends on whether the person is looking at the leftmost, middle, or rightmost number when execution of the strategy that eventually solves the problem begins. Launching execution of the strategy that eventually solved the problem on the previous trial from a specific location results in a boost to the strategy's strength at that location. This leads to the shortcut being primed first when attention is on the rightmost number, which is crucial in production of negation.

The second type of priming involves diffusion of priming from its original location to other locations. The amount of diffusion varies with the distance between the locations. Thus, although the shortcut generally is first used when attention is focused on the rightmost number (as part of the negation strategy, at a behavioral level), priming gradually extends leftward to the other numbers. This type of priming has a major influence on the strategies generated after negation. As priming of the shortcut spreads from right to left, the computation strategy is interrupted earlier. Eventually, the spread of priming makes it possible for the shortcut to be selected at the beginning of a trial, before any arithmetic operations have been conducted, giving rise to the behavior classified as the shortcut strategy.

Both types of priming contribute to the shortcut strategy being used earlier and more often in the blocked problems condition. Over trials within each session, the shortcut becomes increasingly primed. This priming initially leads to use of strategies such as negation, which involve interruption of computation; later, the priming leads to use of the unconscious shortcut and the shortcut, as the relevant activations increase at locations farther to the left. In contrast, in the mixed problems condition, the shortcut is less often applicable on previous trials, so that it does not receive as much priming on the new problem.

e. Forgetting. The forgetting vector is the rate at which the strength of each strategy decays from session to session (i.e., over the week-long intervals between successive sessions). This mechanism operates in much the same way as the priming mechanism described above; indeed, it could be described as the decay of priming. However, forgetting from the previous session is based on a summary of each strategy's activation over all trials within the session, rather than on the last trial of the session alone. Forgetting leads to a decrease in the strength of the most effective strategies at the end of the previous session and to an increase in the strength of the least effective strategies; thus, it involves a kind of regression to the mean of strategy strengths, in which differences among the strengths of strategies become less pronounced. The previous session's activations also decrease in influence over the course of each new session, as the influence of priming from preceding trials within the current session increases.

The forgetting mechanism is particularly important for explaining the steep decline in use of the shortcut from the end of one session to the beginning of the next. In the blocked problems condition, almost all children used the shortcut strategy on the last few trials of each session after the first one. However, forgetting leads SCADS*, like the children in the study, to rarely use the shortcut at the beginning of each new session. Also like the children, the time it takes SCADS* to resume using the shortcut decreases over sessions, because the activation of the shortcut increases over sessions. In other words, although forgetting reduces the activation of the shortcut from one session to the next, the activation remains higher that it was at the beginning of earlier sessions.

f. Dynamic Feature Selection. SCADS*, like its predecessor SCADS, begins each run able to encode many features. These include features that are relevant for solving inversion problems, such as the identity of each number and whether any pair of numbers in the problem is equal, and also features that are irrelevant, such as the relative magnitudes of the three numbers and their color and size. For each perceptual feature and each strategy, the system keeps track of two proportions. One proportion is the number of trials on which the feature was detected and on which the strategy generated unusually good performance (i.e., performance that was correct and at least 50% faster than usual) relative to the total number of trials on which the feature was present. The second counter tracks the number of trials on which the feature was absent and the strategy generated unusually good performance relative to the number of trials on which the feature was absent. If the absolute value of the difference between these two proportions is sufficiently large for several trials, presence or absence of the feature begins to be used in calculating the strength of the strategy and therefore in choosing which strategy to use. This aspect of SCADS* is particularly important for discovery of strategies other than computation, all of which depend on detection of the equality between the B and C terms within the problem. When B = C, these strategies are useful; when B ≠ C, the strategies are slightly harmful (because they require an unnecessary attention shift).

*2. Functioning of SCADS**

These innovations—controlled attention, interruption of procedures, verbalization, priming, forgetting, and dynamic feature selection—were integrated into SCADS in a way that produced strategy choice and discovery much like children's.

a. Overview. SCADS* begins its run with two types of knowledge. It knows how to add and subtract and thus can perform the computation strategy. It also knows that N − N = 0 and can perform that calculation very quickly; empirical studies of age peers of the children who participated in Siegler and Stern (1998) have shown that they solve N − N problems faster than any other subtraction problem (e.g., Woods, Resnick, & Groen, 1975), which suggests that they know the N − N rule.

Consistent with convention, at the outset of a run, SCADS* begins each trial attending to the leftmost number in the problem and proceeds rightward from there. As the system gains experience, it learns that the last two numbers are identical either on all trials (blocked condition) or on a substantial percentage of trials (mixed condition). Despite having this knowledge, the model initially continues to execute the computation strategy from left to right on each trial, because its attention starts there and it does not have enough resources to interrupt execution of the computation strategy once it starts.

With practice solving the three-term problems, the system soon gains enough resources to interrupt the computation strategy after A and B have been added but before the problem has been solved. Immediately after this interruption, SCADS* makes a new strategy choice, which is often to solve the problem as "B − B = 0; 0 + A = A." Having begun the trial by adding A and B and having ended it with this new strategy choice, the system has generated the behavior corresponding to the negation strategy (though from the system's point of view, it solved the problem via the shortcut).

With further practice, the system begins to attend to the B and C terms before any addition is attempted, to check if they are equal on the new problem. This at first usually gives rise to the unconscious shortcut, because the verbalization strength of the shortcut is low. As the verbalization strength of the shortcut increases, the process gives rise to the conscious version of the shortcut, which is chosen increasingly as the system gains experience solving inversion problems. Use of the shortcut increases more rapidly in the blocked problems condition than in the mixed problems condition, due to the shortcut always being applicable in the blocked problems condition. However, forgetting between sessions leads to other strategies also being used, particularly at the beginning of each new session.

b. Strategy Selection. With this overview as background, we can consider in greater detail the way that SCADS* selects among strategies. Strategy selection within SCADS* proceeds in basically the same way as in SCADS, but with additional inputs and greater flexibility. As with SCADS, the strategy choice process depends on the strength of each strategy relative to the strengths of its competitors. A strategy's strength is a function of the accuracy and speed of the solutions that it has produced, and also of its novelty. Strategies that have produced faster and more accurate performance than alternative approaches are selected increasingly often. Also as with SCADS, strengths are based on a combination of the strategy's global strength (summed across all problems), its featural strength (summed across problems with similar features), and its local strength (unique to the particular problem). Another carryover from SCADS is that newly generated strategies are awarded novelty points, a kind of strength that gives new strategies a chance to show their effectiveness. Thus, the basic strategy choice mechanism of SCADS and SCADS* is the same.

Because of the greater variety of influences on the strengths of strategies, however, strategy choice winds up being more complex within SCADS* than within its predecessor. This greater complexity begins with the model's encoding of problem features. SCADS* encodes each problem in terms of the feature detectors that are active at the time. These always include the identities of the numbers and arithmetic operations and include on a probabilistic basis other features, such as whether any numbers in the problem are identical, the colors and sizes of the numbers, and whether all numbers are odd or all are even. The problem as encoded activates the system's memory of the strength of each strategy when those features are present.

New to SCADS*, the strengths on which strategy choices are based also vary with the focus of visual attention. For example, the computation strategy has a stronger activation when attention is focused on the leftmost number than elsewhere, because that is the location from which the strategy has most often been launched. Also new to SCADS*, a strategy's strength is influenced by priming from the previous trial, from the previous session, and from other foci of visual attention. Priming from the previous trial is particularly important in allowing the model to learn from its experience in the immediate situation; priming from the previous session is particularly important in allowing the strength of strategies to reflect a larger database of relatively recent performance. As the system progresses through a given session, the weighting of experience in that session increases, and the weighting of experience in the previous session decreases.

After a strategy is selected, its execution requires cognitive resources. This limits the resources that are available for other types of strategy choices, such as deciding to interrupt execution of an ongoing strategy. Again as in SCADS, the more often a given strategy has been executed, the greater the resources available for these choices. Unlike in SCADS, calculation of the number of previous uses of a strategy is based in large part on uses that start from the given location.

Because the strength of each strategy varies with the focus of attention, the strategy selected following an interruption often differs from the strategy that was initially selected. In particular, the strength of the shortcut strategy, and therefore its likelihood of being chosen, tends to be greater at the middle and rightmost numbers, because the shortcut begins with determining that those two numbers are equal. This results in SCADS* frequently beginning to execute the computation strategy when attention is focused on the leftmost number but then interrupting the computation with the shortcut, once attention moves to the rightmost number.

When SCADS* interrupts execution of a strategy and conducts a second strategy choice, it sometimes chooses the same strategy as it originally did. When this occurs, execution of the original approach simply continues. However, if a different strategy is selected, the model abandons execution of the original procedure and begins again. When the strategy changes, calculations of speed and accuracy of the new strategy begin anew, and the newly chosen strategy is strengthened or weakened depending on the performance it generates from the time of the shift. Thus, if the model adds $A + B$ in the first 6 sec, then interrupts the computation and chooses the shortcut, and then generates the answer via the shortcut in the next 3 sec, the shortcut is credited with solving the problem in 3 sec.

The effects of interruptions on subsequent strategy choices depend on the type of problem being solved and on which strategy is chosen. Three types of cases arise. In the first case, the interruption occurs on an inversion problem and the shortcut strategy is chosen to complete the problem. Under these circumstances, the shortcut strategy is strengthened and chosen more often on future problems. The reason is that the shortcut generated the answer to the problem, and the speed and accuracy of the solution will be higher than the average for problems where $B = C$ (because the average for such problems includes answers generated by the slower and less accurate computation strategy). Now consider a second case: the interruption occurs on an inversion problem, and the computation strategy is chosen to complete the problem. Under these circumstances, the computation strategy is weakened. The reason is that the interruption slightly increases the time to solve the problem relative to that which occurs when computation is used without an interruption, and even without the interruption, computation is the slowest approach for solving inversion problems. Finally, consider a third case: the interruption occurs on a standard problem, and the system chooses computation. In this case, the interruption has little effect, because the shortcut cannot be used on standard problems.

c. Strategy Discovery. SCADS* generates the same four new approaches as the children in Siegler and Stern (1998): negation, computation/shortcut, unconscious shortcut, and shortcut. Because the discovery process is complex, this discussion is aimed at conveying the logic and key concepts rather than all of the details of implementation.

As noted previously, SCADS* solves problems by using the computation strategy until it frees sufficient cognitive resources to allow interruptions. Such interruptions give rise on a probabilistic basis to a new strategy choice or to an attempt to discover a new strategy. When the model attempts to discover a new strategy, it generates a short chain of visual attention and arithmetic operations and inserts it at the point in the strategy at which the interruption occurred. The effect is to change the order in which attention is focused on the numbers within the problem and therefore the order in which the arithmetic operations are performed. This feature of the model is consistent with our informal observation that after children become proficient in basic arithmetic, they fairly often depart from the usual arithmetic algorithm to see if another ordering of operations would make the problem easier.

SCADS* then applies two mechanisms that were also part of the SCADS discovery process. First, SCADS' redundancy elimination mechanisms are applied to the expanded sequence to eliminate unneeded actions. Then, SCADS' goal sketch filters are applied to guarantee that the new strategy uses all numbers in the problem once and only once and that an answer is generated. Many proposed strategies are never tried, because they violate the goal sketch, or because redundancy elimination results in a strategy identical to one already in the repertoire. However, some strategies survive the scrutiny and are tried.

Now we can proceed to a more detailed account of how SCADS* discovers the four strategies of interest. The first strategy discovered by the simulation (as by children) is negation. This strategy emerges when SCADS* is attempting to execute the computation strategy on an inversion problem, has added the leftmost and middle numbers, attends to the rightmost number, enters its value into working memory, and then interrupts the procedure and chooses the shortcut. Interruptions are particularly likely at this point because cognitive resources increase during the course of strategy execution and are at their maximum here, just before execution of the computation strategy is completed. The interruption initially leads to creation of a procedure that, on inversion problems, is identical to negation but that on other problems produces a variant of computation. (Within the notation in the following equations, different strategies are separated by periods, whereas the steps within a single strategy are separated by semicolons.)

(1) Do A + B. Interrupt. Do B − C; add A to the result of B − C; advance A as the answer.

SCADS* soon learns that this procedure is much faster and more accurate when the final two numbers are equal. It therefore creates the following procedure, which corresponds precisely to negation.

(2) Do A + B. Interrupt. Check if B = C; if so, do B − C; add A to the result of B − C; advance A as the answer.

Because these new strategies are generated when attention is focused on C, their strength is initially largely confined to the point in strategy execution when the model attends to C. However, as the system gains experience with inversion problems, the cognitive resources available for discovery increase, which leads to it becoming able to interrupt the computation strategy earlier in its execution. Moreover, the strength of the shortcut strategy, as expressed in Equations 2 and 3, increasingly diffuses to locations B and A. These two trends sometimes lead to generation of the computation/shortcut strategy, in which the computation strategy is interrupted before completion of the initial addition operation, and A is stated as the answer.

Either before this point or soon after, the increasing strength of the shortcut when attention is at location C leads to the shortcut being chosen at the beginning of trials and attention immediately being shifted to locations B and C to check if the values are equal. The result is a form of the shortcut:

(3) Check if B = C; if so, do B − C; add A to the result; advance A as the answer.

Next, the system deletes the subtraction operation B − C and the addition of the result to A, which are unnecessary for solving the problem. This gives rise to the purest form of the shortcut:

(4) Check if B = C; if so, advance A as the answer.

SCADS* also produces an overly general form of the shortcut:

(5) Check if any two numbers are equal; if so, advance A as the answer.

This version of the shortcut produced many of the overgeneralizations in Session 8.

The first uses of the shortcut are unconscious, because its verbalization strength is weak. After a number of uses of the unconscious shortcut, however, the verbalization strength of the shortcut increases sufficiently for the system to report using it.

d. SCADS' Performance.* To test the fit between the children's and the model's performance, we presented the model with experimental conditions like those that the children received. The version of the model corresponding to the mixed problems condition received 10 inversion and 10 standard items, randomly interspersed, in each of the first seven sessions, just as children did. The version of the model corresponding to the blocked problems condition received the same items as those presented in the mixed problems condition in Sessions 1, 5, and 7. In Sessions 2, 3, 4, and 6, however, this version of the model received 20 inversion problems and no standard problems; again, this paralleled the experience that children in the blocked problems condition received in these sessions. The first 10 problems presented in the blocked problems condition in these sessions were identical to the 10 inversion problems presented in the mixed problems condition in the sessions.

In both conditions, the statistics on the model's performance were based solely on performance on the first 10 inversion problems encountered (the only inversion problems encountered in the mixed problems condition and in Sessions 1, 5, and 7 in the blocked problems condition.) In Session 8, the model was presented the set of generalization problems that children encountered in their Session 8. Between each consecutive pair of sessions, the forgetting function was applied, to simulate the changes in priming and activation that were hypothesized to occur in the week-long period between the sessions.

The model's strategy use was classified on the basis of overt behavior and solution times, just as that of the children was. Thus, when the procedures corresponding to Equations 1 and 2 were applied to inversion problems, the response was classified as the negation strategy. Similarly, when the procedures corresponding to Equations 3, 4, and 5 were used, the response was classified as use of the shortcut strategy.

To model the variability that different children brought to the experimental situation, a number of parameters within SCADS* varied randomly, within limits, from run to run. Thus, each of the 50 runs can be thought of as representing a different child with different characteristics. One type of variability involved efficiency of learning; on different runs, each strategy's speed and accuracy improved more or less rapidly. The rapidity with which attentional resources were freed also varied randomly from run to run, as did the thresholds for interrupting strategy execution and feature detection. No effort was made to optimize these variations in values; as with SCADS, the goal was to build a model that could generate the main qualitative features of children's performance, and to be quantitatively reasonable, rather than to vary parameters to obtain the best possible fit to the data (and perhaps to take advantage of chance in doing so).

SCADS* generated all nine of the major features of performance shown by the children in Siegler and Stern (1998) and described previously. We organize the description of the models' performance around these features:

(1) *Children used five strategies in the experiment: computation, negation, computation/shortcut, unconscious shortcut, and shortcut*. SCADS* generated all five of these strategies and rarely used strategies that the children did not use. However, it did sometimes generate a sixth strategy that was not noted among the children: $A - C + B$. This strategy does not have any particular advantage over the computation strategy—it executes the same operations in a slightly different order and produces the same solution times. Children may have produced the strategy occasionally but the difference between it and the computation strategy was not evident in their explanations or solution times.

(2) *Over sessions, children became faster and more accurate and relied increasingly on the shortcut strategy*. SCADS* generated improvements in performance over sessions that paralleled those shown by the children in Siegler and Stern (1998). The main source of the improvement was generation of faster and more accurate strategies and increasing reliance on those more effective approaches. As shown in Figures 5 and 6, SCADS*' reliance on the computation strategy greatly decreased over sessions in both conditions. Under blocked problems conditions, reliance on computation decreased from 72% in Session 1 to 26% in Session 7; under mixed problems conditions, use of computation decreased from 73 to 26% over the same period. The corresponding decreases in the children's data were from 73 to 41% in the blocked problems condition and from 71 to 31% in the mixed problems condition. Because the computation strategy involves a much larger number of counting operations than any other strategy, it is the slowest and most error prone approach in the simulation, as in the children's data, which is why its use declines sharply once faster and more accurate approaches are available.

Conversely, from Session 1 to 7, SCADS*' use of the two fastest and most accurate strategies, the shortcut and unconscious shortcut, greatly increased. The increase was from 2 to 50% under blocked problems conditions and from 4 to 34% under mixed problems conditions. The corresponding increases in the children's data were from 0 to 50% in the blocked problems condition and from 0 to 37% in the mixed problems condition. The reason why these strategies are the fastest and most accurate within SCADS* is that they do not require any counting, and application of the rule "$N - N = 0$" is quick and invariably accurate. Presumably, the same factors led to their being the fastest and most accurate in the children's data as well.

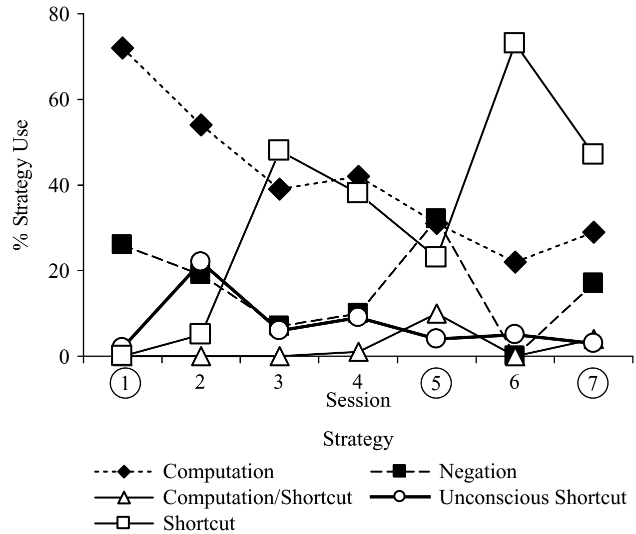

Fig. 5. Changes in SCADS' strategy use over seven sessions: blocked problems condition.*

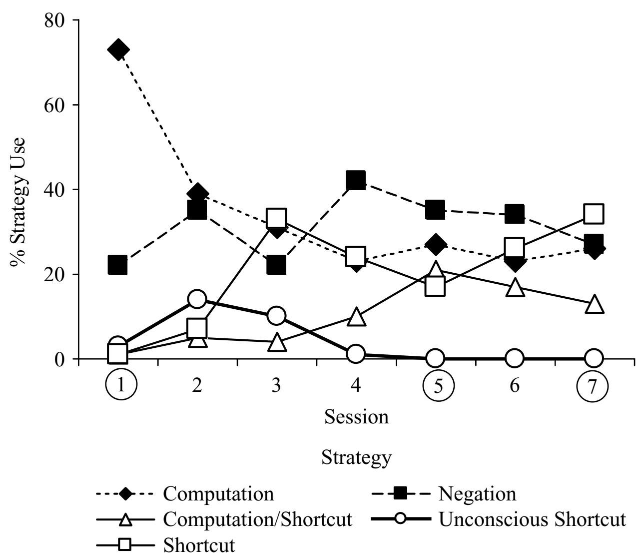

Fig. 6. Changes in SCADS' strategy use over seven sessions: mixed problems condition.*

(3) *Almost all children in both conditions discovered the shortcut and unconscious shortcut strategies*. Again, SCADS* behaved similarly. SCADS* discovered the shortcut approach on 100% of runs, and it discovered the unconscious shortcut on 81% of runs. These percentages were similar to those of children, though SCADS* discovered the shortcut about 10% more often than the children in Siegler and Stern (100 vs. 91%) and discovered the unconscious shortcut about 10% less often (81 vs. 93%).

(4) *Among children in the blocked problems condition, the first uses of the four relatively frequent strategies (more than 5% of trials) were usually in the same order: computation, negation, unconscious shortcut, and then shortcut. Among children in the mixed problems condition, the order of use was similar except that most used the computation/shortcut before they used the shortcut*. As noted previously, SCADS* started its run only knowing the computation strategy. At the level of process, it then discovered the shortcut strategy. However, at the level of behavior, which was the only level at which the children's strategies could be coded, the model generally discovered the strategies in the same order as the children did. After using the computation strategy to solve several problems, SCADS* almost always (96% of runs) generated the negation strategy. As in the children's data, SCADS* usually generated this strategy toward the end of Session 1. Also as in the children's data, the strategies immediately preceding the unconscious shortcut were always the computation or negation strategies (Figure 7). As with the children, the unconscious shortcut usually was discovered in Session 2. The next discovery most often was the shortcut, which the model also usually discovered in Session 2. As shown in Figure 8, discovery of the shortcut was typically preceded by use of the unconscious shortcut, as in the children's data.

One difference between the children's performance and that of SCADS* involved the precursors of the shortcut. When presented the blocked problems, SCADS* fairly often (30% of runs) generated the shortcut without having first generated the unconscious shortcut. The exact threshold that the verbalization strength needed to exceed was sometimes relatively low, so that fairly small amounts of verbalization activation were sufficient to exceed it and thus generate the shortcut without having previously generated the unconscious shortcut. This finding suggests that the variability of this threshold may have been greater than optimal, at least at the low end.

(5) *Presentation of the blocked problems resulted in the shortcut strategy being discovered earlier, used more often, and generalized more broadly than did presentation of the mixed problems*. These effects all follow from the same cause; the shortcut strategy becomes more activated when it applies to all problems. Several specific mechanisms contribute to the effect. One involves attention. Presentation of inversion problems on all trials leads SCADS* to learn more quickly to check if B and C are equal before adding A and B.

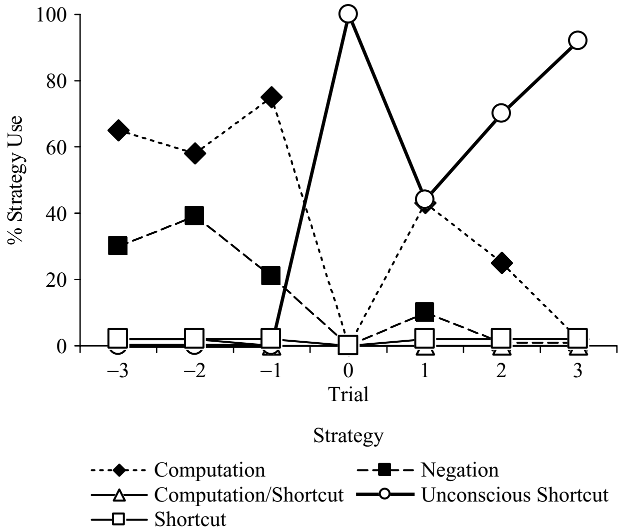

Fig. 7. SCADS'* strategy use on trials immediately before and after first use of unconscious shortcut strategy in blocked problems condition.*

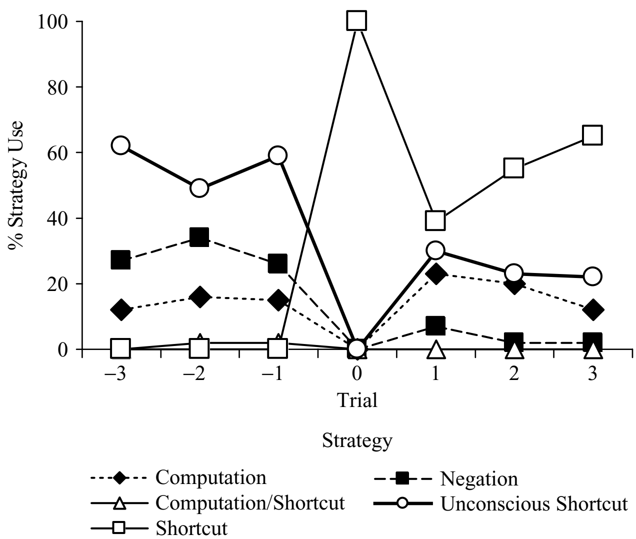

Fig. 8. SCADS'* strategy use on trials immediately before and after first use of shortcut strategy in blocked problems condition.*

In effect, under blocked problems conditions, SCADS* is consistently reinforced for shifting attention to the B and C terms at the beginning of the problem, whereas under mixed problems conditions, reinforcement is intermittent.

Priming also plays an important role in the greater use and earlier discovery of the shortcut strategy in the blocked problems condition. The more often the shortcut leads to problem solutions, the more likely it is to receive priming from the previous problem. The shortcut is primed even when the system's behavior corresponds to negation or the computation/shortcut, because within those strategies, the answer is produced by the shortcut (after it has interrupted execution of computation). In contrast, in the mixed problems condition, the shortcut receives less priming, because it cannot be used effectively on half of the trials and because its competitor, the computation strategy, receives greater priming, because it is the only way to solve half of the problems.

Finally, overgeneralization occurs through a combination of the greater activation of the shortcut leading to attention shifting to the B and C terms earlier in each trial, and to the coding of the equality of B and C sometimes being limited to identity of the absolute value of the numbers rather than identity of their signed values. In Session 8, after presentation of the blocked problems, SCADS* generalized correctly on 66% of problems on which the inversion principle applied, such as $A - B + C$; it generalized inappropriately on 56% of problems on which the principle did not apply, such as $A - B - B$. The shortcut strategy was used less often, but more selectively, following presentation of the mixed problems. It was used on 52% of problems where it was applicable and 32% of problems where it was not. The corresponding percentages for the children were 65 and 66% for children in the blocked problems condition, and 38 and 28% for children in the mixed problems condition. Thus, SCADS* generated slightly more appropriate generalization patterns than did the children, but like them showed inadequate discrimination between problems where the inversion principle did and did not apply.

(6) *Strategy use remained variable after discovery of the shortcut, even in the blocked problems condition*. Like the children, SCADS* continued to use the other four strategies on many trials after the shortcut was discovered (usually in Session 2). Indeed, as with the children, under mixed problems conditions SCADS* never used the shortcut on more than 50% of trials in any session, and under blocked problems conditions, it did so only in Session 6. Comparing the children's strategy use (Figures 1 and 2) to SCADS* strategy use (Figures 5 and 6) shows the degree of variability of both in all sessions.

Again, several mechanisms produce this outcome. Forgetting and loss of priming between sessions reduced the activation of the shortcut, which resulted in the computation strategy predominating at the beginning of each new session. SCADS*, like children, usually generated the negation strategy, and often the unconscious shortcut, before the shortcut in any given session. Gradual increases in long-term activation of the shortcut resulted in it being used more often on the early trials of later sessions, but even then it was not usually used at the outset of the session. Thus, SCADS* provided an explanation for why, despite the clear advantages of the shortcut strategy in speed, accuracy, and elegance, children continued to use other strategies along with it in all sessions of the study.

(7) *Use of the unconscious shortcut was particularly frequent in the blocked problems group in the trials immediately before the first use of the shortcut*. One of the most striking findings in Siegler and Stern (1998) was that children in the blocked problems condition used the unconscious shortcut on 80% of trials on the three items immediately preceding the first use of the shortcut strategy. This rate of use was far higher than the 9% of trials in all sessions on which children in the blocked problems condition used that strategy or the 3% of trials on which children in the mixed problems condition did.

As shown in Figure 8, SCADS* produced a similar effect. When it was presented the blocked problems, it used the unconscious shortcut on almost 60% of items on the three trials just before it first generated the shortcut. This was far greater than the 7% of trials in all sessions on which SCADS* used the unconscious shortcut in the blocked problems condition and the 4% of trials on which it generated this strategy in the mixed problems condition. As noted earlier, the hypothesized asynchrony between the rise in activation of the shortcut and the rise of the strategy's verbalization strength generated this effect.

(8) *Children in the mixed problems condition used the negation and computation/shortcut strategies more often than did children in the blocked problems condition*. SCADS* used negation and the computation/shortcut on 41% of trials in the mixed problems condition; they used these strategies on 18% of trials in the blocked problems condition. Similarly, the children in Siegler and Stern (1998) in the mixed problems condition used the negation and computation/shortcut strategies on 36% of trials, vs. 10% among those in the blocked problems condition. The smaller trial-to-trial priming of the shortcut strategy in the mixed problems condition was a major factor in producing SCADS*' greater reliance on the strategies that involved interruption of computation (negation and the computation/shortcut). Under mixed problems conditions, there were often sufficient resources available to allow interruptions, and the activation of the shortcut was often strong enough at the middle and rightmost locations for it to be chosen when attention was focused there, but not strong enough for it to be chosen when attention was focused on the leftmost location.

(9) *When children in the mixed and blocked problems groups received standard as well as inversion problems (Sessions 1, 5, and 7), shortcut strategy*

use was similar in the two conditions. SCADS* generated this effect, like a number of others involving changes in strategy use over sessions, through a combination of forgetting and priming. The strong trial-by-trial priming of strategy use that led to increasing use of the shortcut strategy over trials in the blocked problems condition did not operate when both groups received the mixed problems. In addition, forgetting reduced the activation of the shortcut from the level reached by the end of the preceding session. Long-term changes in activation led to SCADS* using the shortcut somewhat more often by Session 7 in the blocked condition—47% of trials in the blocked problems condition vs. 34% in the mixed problems condition. This difference was small, however, relative to the differences in Session 6, the final session in which the two groups received different problems. There, the blocked problems elicited 73% use of the shortcut, vs. 26% when SCADS* received the mixed problems. The corresponding statistics for the children were 69 vs. 25% use of the shortcut in Session 6 and 34 vs. 32% in Session 7. Again, SCADS* produced somewhat greater generalization than the children did, but like them used the shortcut far more often when presented 100% inversion problems.

IV. Conclusion

SCADS* demonstrates that diverse phenomena concerning strategy choice and discovery can be captured within a single computational model. Accounting for these diverse phenomena required adding a number of basic cognitive processes to the original SCADS structure. From one perspective, this could be viewed as a step backward: both models account for a wide range of phenomena, but SCADS* requires a wider range of processes to do so. In this sense, SCADS* is less parsimonious than its predecessor. However, the inversion task is considerably more complex than simple addition, and SCADS* learns under different experimental conditions (blocked and mixed) and on different types of problems (standard and inversion). It is unclear how a model that did not include priming, forgetting, controlled attention, and the other new features of SCADS* could have generated the types of learning shown by the children in this more complex and varied situation. In addition, these mechanisms clearly are involved in many aspects of human cognition, and their inclusion seems likely to be useful for modeling children's learning in a wide range of situations. New simulations of strategy choice and discovery in other contexts will enable us to test this belief.

A basic tenet of the triangulation strategy is that the three types of investigation are mutually enriching. Examples of this mutual enrichment arose in both of the examples discussed in this chapter. In the series of investigations of addition, a microgenetic study revealed that the preschoolers never used strategies that violated the basic principles of addition. This gave rise to the goal sketch hypothesis, which was tested in a cross-sectional study. The combined results gave rise to the goal sketch mechanism that played an important role not only within the SCADS model of addition but also within the SCADS* model of inversion.

Research on the inversion problem illustrated other advantages of the triangulation strategy. Cross-sectional studies indicated that children use three strategies: computation, shortcut, and negation. A microgenetic study revealed the existence of two other strategies, the computation/shortcut and unconscious shortcut, which were too short-lived to emerge in cross-sectional investigations. The modeling efforts that led to SCADS*, however, suggested that only computation and the shortcut were distinct strategies that children chose at the outset of a problem. The negation and computation/shortcut approaches appeared to arise through interruption of computation by the shortcut, and the unconscious shortcut appeared to arise through an asynchrony between the activation of the shortcut strategy itself and that of its verbalization. This conclusion raises the question of whether interrupting execution of ongoing strategies plays a role in discovery of other strategies as well, a question that can be addressed only through further cross-sectional and microgenetic studies.

The inversion example also demonstrated how age-related, microgenetic, and formal modeling studies can together address large scientific and philosophical issues. Whether insights arise suddenly or gradually, and whether they arise consciously or unconsciously, are subjects of longstanding debate but little empirical investigation. Applying the triangulation strategy demonstrated how progress on these issues could be made. Cross-sectional studies indicated that the inversion task was potentially revealing about the issues; the microgenetic study yielded data relevant to them; the computer simulation specified mechanisms that could generate unconscious as well as conscious strategy discoveries and that could generate a qualitative change through a series of small advances. Indeed, SCADS* raised new questions concerning when the discovery actually takes place on this task: when the system begins the problem by adding $A + B$ but solves it via the shortcut; when it uses the shortcut from the beginning of the trial but does so unconsciously; or when it uses the shortcut consciously from the beginning of the problem? As is often the case when mechanisms are specified, what looks from the outside like an abrupt qualitative change looks from the inside like the endpoint of a sequence of small steps forward.

The field of cognitive development has frequently been criticized for insufficient focus on change mechanisms (e.g., Flavell, 1984; Klahr & MacWhinney, 1998; Miller, 2002). These critiques seem well founded; there is certainly an imbalance between the huge number of changes with age and experience that developmentalists have uncovered and the tiny proportion of those changes that they have explained mechanistically. Greater adherence to the triangulation strategy may help remedy this imbalance.

ACKNOWLEDGEMENTS

Thanks are due to the National Institutes of Health Grant HD 19011 and the Office of Educational Research and Instruction Grant R305H020060, both of which supported this research.

REFERENCES

Albali, M. W. (1999). How children change their minds: Strategy change can be gradual or abrupt. *Developmental Psychology*, *35*, 127–145.

Ashcraft, M. H. (1987). Children's knowledge of simple arithmetic: A developmental model and simulation. In J. Bisanz, C. J. Brainerd, & R. Kail (Eds.), *Formal methods in developmental psychology* (pp. 302–338). New York: Springer.

Bisanz, J., & LeFevre, J. (1990). Strategic and nonstrategic processing in the development of mathematical cognition. In D. F. Bjorklund (Ed.), *Children's strategies; contemporary views of cognitive development* (pp. 213–244). Hillsdale, NJ: Erlbaum.

Bryant, P., Christie, C., & Rendu, A. (1999). Children's understanding of the relation between addition and subtraction: Inversion, identity, and decomposition. *Journal of Experimental Child Psychology*, *74*, 194–212.

Buehler, K. (1908). Tatsachen und Probleme ze einer Psychologie der Denkvorgaenge. [Facts and problems of a psychology of thinking]. Teil II: Ueber Gedankenzusammenhaenge. *Archiv fuer die gesamte Psychologie*, 12, 1–23.

Church, R. B., & Goldin-Meadow, S. (1986). The mismatch between gesture and speech as an index of transitional knowledge. *Cognition*, *23*, 43–71.

Davidson, J. (1995). The suddenness of insight. In R. J. Sternberg & J. E. Davidson (Eds.), *The nature of insight* (pp. 125–155). Cambridge, MA: MIT Press.

Duncker, K. (1945). On problem solving. *Psychological Monographs*, *58*(5), Whole No. 270.

Flavell, J. H. (1984). Discussion. In R. J. Sternberg (Ed.), *Mechanisms of cognitive development* (pp. 187–209). New York: Freeman.

Fuson, K. C. (1982). An analysis of the counting-on solution procedure in addition. In T. P. Carpenter, J. M. Moser, & T. A. Romberg (Eds.), *Addition and subtraction: A cognitive perspective* (pp. 67–82). Hillsdale, NJ: Erlbaum.

Geary, D. C. (1994). *Children's mathematical development: Research and practical implications*. Washington, DC: American Psychological Association.

Geary, D. C., & Wiley, J. G. (1991). Cognitive addition: Strategy choice and speed-of-processing differences in young and elderly adults. *Psychology and Aging*, *6*, 474–483.

Gick, M. L., & Lockhart, R. S. (1995). Cognitive and affective components of insight. In R. J. Sternberg & J. E. Davidson (Eds.), *The nature of insight* (pp. 197–228). Cambridge, MA: MIT Press.

Goldin-Meadow, S., & Alibali, M. W. (2002). Looking at the hands through time: A microgenetic perspective on learning and instruction. In N. Granott & J. Parziale (Eds.), *Microdevelopment: Transition processes in development and learning* (pp. 80–105). Cambridge, UK: Cambridge University Press.

Groen, G. J., & Parkman, J. M. (1972). A chronometric analysis of simple addition. *Psychological Review*, *79*, 329–343.

Gross, H.-M., Heinze, A., Seiler, T., & Stephan, V. (1999). Generative character of perception: A neural architecture for sensorimotor anticipation. *Neural Networks*, *12*, 1101–1129.

Gruber, H. E. (1981). On the relation between "Aha! experiences" and the construction of ideas. *History of Science*, *19*, 41–59.

Isaak, M. I., & Just, M. A. (1995). Constraints on thinking in insight and invention. In R. J. Sternberg & J. E. Davidson (Eds.), *The nature of insight* (pp. 281–325). Cambridge, MA: MIT Press.

Kaplan, C. A., & Simon, H. A. (1990). In search of insight. *Cognitive Psychology*, *22*, 374–419.

Karmiloff-Smith, A. (1992). *Beyond modularity: A developmental perspective on cognitive science*. Cambridge, MA: MIT Press.

Klahr, D., & MacWhinney, B. (1998). Information processing. In W. Damon (Series Ed.) & D. Kuhn, & R. S. Siegler (Vol. Ed.), *Handbook of child psychology: Vol. 2: Cognition, perception & panguage* (5th ed., pp. 631–678). New York: Wiley.

Klein, J. S., & Bisanz, J. (2000). Preschoolers doing arithmetic: The concepts are willing but the working memory is weak. *Canadian Journal of Experimental Psychology*, *54*, 105–116.

Kohler, W. (1925). *The mentality of apes*. New York: Liveright.

Kotovsky, K., Hayes, J. R., & Simon, H. A. (1985). Why are some problems hard? Evidence from Tower of Hanoi. *Cognitive Psychology*, *17*, 248–294.

Kuhn, D., Garcia-Mila, M., Zohar, A., & Anderson, C. (1995). Strategies of knowledge acquisition. *Monographs of the Society for Research in Child Development*, Serial No. 245.

LeFevre, J. A., Sadesky, G. S., & Bisanz, J. (1996). Selection of procedures in mental addition: Reassessing the problem-size effect in adults. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*, 216–230.

Lemaire, P., & Siegler, R. S. (1995). Four aspects of strategic change: Contributions to children's learning of multiplication. *Journal of Experimental Psychology: General*, *124*, 83–97.

Lewin, K. (1935). *A dynamic theory of personality*. New York: McGraw-Hill.

Lieberman, P. (2000). *Human language and our reptilian brain: The subcortical bases of speech, syntax, and thought*. Cambridge, MA: Harvard University Press.

Maier, N. R. F. (1931). Reasoning in humans: II. The solution of a problem and its appearance in consciousness. *Journal of Comparative Psychology*, *12*, 181–194.

Metcalfe, J. (1986). Premonitions of insight predict impending error. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *12*, 623–634.

Miller, P. (2002). *Theories of developmental psychology* (4th ed.). New York: Worth.

Newell, A. (1990). *Unified theories of cognition*. Cambridge, MA: Harvard University Press.

Ohlsson, S. (1992). Information-processing explanations of insight and related phenomena. In M. Keane & K. Gilhooley (Eds.), *Advances in the psychology of thinking* (Vol. 1, pp. 1–44). London: Harvester-Wheatsheaf.

Palmer, S. (1999). *Vision science: Photons to phenomenology*. Cambridge, MA: MIT Press.

Perkins, D. N. (1995). Insight in minds and genes. In R. J. Sternberg & J. E. Davidson (Eds.), *The nature of insight* (pp. 495–533). Cambridge, MA: MIT Press.

Perry, M., Church, R. B., & Goldin-Meadow, S. (1988). Transitional knowledge in the acquisition of concepts. *Cognitive Development*, *3*, 359–400.

Rasmussen, C., Ho, E., & Bisanz, J. (2003). Use of mathematical principle of inversion in young children. *Journal of Experimental Child Psychology*, *85*, 89–102.

Richardson, D. C., & Spivey, M. J. (2000). Representation, space, and Hollywood Squares: Looking at things that aren't there anymore. *Cognition*, *76*, 269–295.

Schooler, J. W., Fallshore, M., & Fiore, S. M. (1995). Epilogue: Putting insight into perspective. In R. J. Sternberg & J. E. Davidson (Eds.), *The nature of insight* (pp. 559–587). Cambridge, MA: MIT Press.

Shrager, J., & Siegler, R. S. (1998). SCADS: A model of children's strategy choices and strategy discoveries. *Psychological Science*, *9*, 405–410.

Siegler, R. S. (1987). The perils of averaging data over strategies: An example from children's addition. *Journal of Experimental Psychology: General*, *116*, 250–264.
Siegler, R. S. (1989). Hazards of mental chronometry: An example from children's subtraction. *Journal of Educational Psychology*, *81*, 497–506.
Siegler, R. S., & Crowley, K. (1991). The microgenetic method: A direct means for studying cognitive development. *American Psychologist*, *46*, 606–620.
Siegler, R. S., & Crowley, K. (1994). Constraints on learning in non-privileged domains. *Cognitive Psychology*, *27*, 194–227.
Siegler, R. S., & Jenkins, E. A. (1989). *How children discover new strategies*. Hillsdale, NJ: Erlbaum.
Siegler, R. S., & Shipley, C. (1995). Variation, selection, and cognitive change. In T. Simon & G. Halford (Eds.), *Developing cognitive competence: New approaches to process modeling* (pp. 31–76). Hillsdale, NJ: Erlbaum.
Siegler, R. S., & Shrager, J. (1984). Strategy choices in addition and subtraction: How do children know what to do? In C. Sophian (Ed.), *The origins of cognitive skills* (pp. 229–293). Hillsdale, NJ: Erlbaum.
Siegler, R. S., & Stern, E. (1998). A microgenetic analysis of conscious and unconscious strategy discoveries. *Journal of Experimental Psychology: General*, *127*, 377–397.
Simon, H. A. (1973). Does scientific discovery have a logic? *Philosophy of Science*, *40*, 471–480.
Simonton, D. K. (1995). Foresight in insight? A Darwinian answer. In R. J. Sternberg & J. E. Davidson (Eds.), *The nature of insight* (pp. 465–494). Cambridge, MA: MIT Press.
Starkey, P., & Gelman, R. (1982). The development of addition and subtraction abilities prior to formal schooling in arithmetic. In T. P. Carpenter, J. M. Moser, & T. A. Romberg (Eds.), *Addition and subtraction: A cognitive perspective* (pp. 99–116). Hillsdale, NJ: Erlbaum.
Stern, E. (1992). Spontaneous use of conceptual mathematical knowledge in elementary school children. *Contemporary Educational Psychology*, *17*, 266–277.
Stern, E. (1993). What makes certain arithmetic word problems involving comparison of sets so difficult for children? *Journal of Educational Psychology*, *85*, 7–23.
Van Lehn, K. (1990). *Mind bugs: The origins of procedural misconceptions*. Cambridge, MA: MIT Press.
Wittgenstein, L. (1969). *On certainty*. New York: Harper & Row.
Woods, S. S., Resnick, L. B., & Groen, G. J. (1975). Experimental test of five process models for subtraction. *Journal of Educational Psychology*, *67*, 17–21.
Yaniv, I., & Meyer, D. E. (1987). Activation and metacognition of inaccessible stored information: Potential bases of incubation effects in problem solving. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *13*, 187–205.

OUT-OF-SCHOOL SETTINGS AS A DEVELOPMENTAL CONTEXT FOR CHILDREN AND YOUTH

Deborah Lowe Vandell, Kim M. Pierce and Kimberly Dadisman
DEPARTMENT OF EDUCATIONAL PSYCHOLOGY, UNIVERSITY OF WISCONSIN, MADISON, WI 53706, USA

Since the publication of Bronfenbrenner's seminal volume, *The Ecology of Human Development: Experiments by Nature and Design* (1979), developmental researchers have focused on understanding the effects of two contexts—the family and the school—on child developmental outcomes. These contexts are no doubt important environments for children and youth (see Eccles, Wigfield, & Schiefele, 1998; Parke & Buriel, 1998; Collins *et al.*, 2000). Since the 1990s, however, investigators have been increasingly aware of the role of out-of-school settings in the lives of children (Carnegie Corporation, 1992; Vandell & Posner, 1999; Lerner, 2005; Mahoney *et al.*, 2005). This awareness is evident in recent reports published by the National Academy of Science that evaluated research findings related to the effects of the out-of-school context on developmental

Advances in Child Development and Behavior
R Kail (Editor)

outcomes of children and youth, and in increased public and private expenditures to support after-school programs and activities (Committee on Community-Level Programs for Youth, 2002; Committee on Family and Work Policies, 2003).

The impetus to understand the out-of-school context in relation to child developmental outcomes has been sparked by a convergence of factors. One factor is the substantial numbers of households in which mothers are employed. More than 75% of families in the United States with children ages 6–17 years have mothers in the labor force (Kleiner, Nolin, & Chapman, 2004). In the case of single mothers, 79% are employed, an increase of 11 percentage points since 1994. Because parents' work hours are typically longer than the school day, there often is a gap between parents' work schedules and their children's school schedules. Families have turned to a variety of arrangements, including after-school programs and structured activities, to "cover" these periods. In 2001, for example, more than 50% of households with children in kindergarten through eighth grade used nonparental care arrangements after school for an average of about 9 h a week (Kleiner, Nolin, & Chapman, 2004). A question raised by parents, policy makers, and researchers is whether some arrangements are more supportive (or are more detrimental) for children than others.

A second factor contributing to the interest in the out-of-school context has been concerns about poor academic achievement in many US schools, particularly those serving low-income children of color. After-school programs are being viewed as a venue for improving math and reading skills although there is not an agreement about the types of programs and activities that might best accomplish this objective. Some organizations and experts have emphasized tutoring, homework help, and direct instruction (see http://www.ed.gov/programs/21stcclc/index.html), whereas others (see http://www.afterschoolalliance.org/) have argued for a broader range of enrichment activities.

A third factor contributing to the increased interest in the out-of-school context is concern about health and safety risks when children and youth lack adult supervision during the hours after school. Juvenile arrests are more heavily concentrated in the time between the end of the school day and the dinner hour than any other time period, including evenings and weekends (Snyder & Sickmund, 1999). In addition, reports of sexual intercourse by adolescents are higher during the hours between 3 p.m. and 6 p.m. than any other time period (Cohen *et al.*, 2002).

Yet another factor contributing to interest in the out-of-school context is the recognition that children have more "free time" than previously (Larson & Verma, 1999). In the last 150 years, long hours of obligatory labor have given way to substantial free time (Kleiber & Powell, 2005). In colonial America, for example, children spent upwards of 10 h a day in household or income-generating labor (Johnson, 2002), whereas youth in contemporary America spend less than an hour a day in these activities (Larson & Verma, 1999). Time in obligatory labor has been replaced, in part, by schooling and homework, but the after-school hours represent a considerable block of time in which children can engage in structured and unstructured activities either in the presence of adults or away from adult supervision.

This convergence of factors has contributed to interest in the out-of-school context by scholars from different academic disciplines and subdisciplines, including developmental psychology, community psychology, sports psychology, sociology, education, leisure studies, and youth development. Our goal in this chapter is to make developmental psychologists aware of this broader literature. First, we consider the conceptual models that are guiding the research, and the methodological challenges of conducting this research. The remainder of the chapter is then devoted to consideration of two particular settings—after-school programs and structured extracurricular activities. Because of space limitations, we do not review the research that considers other important out-of-school settings represented by self-care (see Committee on Family and Work Policies, 2003), summer camps (see Hattie *et al.*, 1997), adolescent employment (see Mortimer & Finch, 1996), media use (see Huston & Wright, 1998), and virtual environments created by computers (see Subrahmanyam *et al.*, 2000).

I. Conceptual Frameworks

Several theoretical frameworks have been used to inform out-of-school research. Four theories, *ecological systems theory*, *stage-environment fit theory*, *flow theory,* and *attachment theory* have identified some general issues to consider and overarching frameworks. Other work has focused on particular developmental processes occurring within out-of-school settings, including the roles and functions of relationships with peers.

A. ECOLOGICAL SYSTEMS THEORY

Vandell and Posner (1999) used ecological systems theory (Bronfenbrenner, 1979; Bronfenbrenner & Morris, 1998) to develop a research agenda to guide the study of after-school environments. Vandell and Posner argued that understanding the effects of the out-of-school context can be advanced by conceptualizing after-school programs, extracurricular activities, and self-care as specific *microsystems,* which are represented as “patterns of activities, roles, and interpersonal relations experienced by the developing person in a given setting with particular physical and material characteristics” (Bronfenbrenner, 1979, p. 22). As part of this after-school research agenda, investigators were called to describe children's experiences in particular settings in terms of the quality of interactions with adults, quality of interactions with peers, and patterns of activities. Characterizations of a particular after-school microsystem also were viewed as including descriptions of physical and psychological resources.

In addition to developing richer descriptions of particular after-school microsystems, a second part of the proposed research agenda was consideration of linkages between children's after-school microsystems, or the study of *mesosystems* in Bronfenbrenner's terminology. For example, a child may be home alone after school before he goes to soccer practice or to music lessons. Or, a child may attend an after-school program for 3 days each week and then be cared for by an older sibling for two afternoons. Research is needed to ascertain how children and families coordinate multiple arrangements and if effects of different patterns of care vary as a function of children's developmental level or competencies. Research has only begun to emerge that considers patterns or combinations of after-school settings in relation to child developmental outcomes (e.g., Mahoney, Lord, & Carryl, in press).

A third part of the research agenda proposed by Vandell and Posner (1999) is the consideration of linkages between contexts at least one of which does not include the developing child, or *exosystems* in Bronfenbrenner's formulation. For example, linkages between parents' employment and the out-of-school context are evident, with maternal employment impacting participation in extracurricular activities that are supported by tuition and fees. Other linkages between the family and out-of-school contexts are evident in conditional or moderated effects with child developmental outcomes. For example, the benefits of after-school programs appear to be greater for children who are living in poverty than for children in more privileged circumstances (Committee on Family and Work Policies, 2003).

Although ecological systems theory has provided a general set of issues to consider with respect to the out-of-school context, it does not articulate developmental processes that may be of particular relevance for the after-school hours. Other work, however, has begun to identify these processes.

B. STAGE-ENVIRONMENTAL FIT

Jacquelynne Eccles' conceptualization of stage-environmental fit also has helped to guide research examining the out-of-school context. Eccles has argued that child functioning is influenced by the *fit* between developmental stage and the social environment (Eccles *et al.*, 1993; Eccles, 1999). A good fit between the individual and the particular environmental context is posited to result in more positive developmental outcomes, whereas a poor fit is posited to increase the likelihood of negative developmental outcomes. A good fit is achieved when the demands of the environment are appropriate to the child's emotional, cognitive, social, and psychological needs.

In her initial work, Eccles employed the stage-environment fit perspective to examine the effects of the school context. She observed a disjunction between young adolescents' developmental needs and capacities and the structure and organization of their middle schools. In a developmental period in which youth are seeking increased opportunities for decision-making, autonomy, and participation, their middle school environments offer fewer opportunities than grade schools do. The organization of the middle school, in which students change classes every period, reduces opportunities for close teacher–child relationships even though supportive relationships with adults continue to be important for young adolescents. As a consequence, according to stage-environment fit theory, some of the problems associated with adolescence are not inherent in being an adolescent, but a reflection of the placement of adolescents in an environment that is unsuitable to their developmental needs.

In her subsequent research, Eccles extended stage-environment theory to the study of the out-of-school context (Eccles, 1999, in press; Committee on Community-Level Programs for Youth, 2002). Because after-school programs and extracurricular activities have more freedom than traditional schools to provide social environments that fit the developmental needs of children and adolescents, Eccles argued that the out-of-school setting may be particularly important for the promotion of positive developmental outcomes.

Reports from adolescents suggest that structured out-of-school activities are a better fit than other contexts. For example, Hansen, Larson, and Dworkin (2003) asked 450 high school students to rate their opportunities for learning in youth activities, at school, and while hanging out with friends. The students reported more opportunities to try new things, to set goals, to push themselves, to learn time management, to learn to control their temper, to improve their athletic skills, to learn to work with others and learn to compromise, to have opportunities to be in charge, and to develop linkages with other adults in the community at youth activities than occurred either at school or while hanging out with friends. The adolescents also noted some differences in the opportunities afforded by particular activities. Students who participated in service and civic activities reported more opportunities to learn about helping others than did students involved in arts and sports activities. Students who participated in service and civic activities reported more linkages with the community than those in academic clubs. Students involved in sports were more likely than those in service, arts, or academic clubs to report negative peer influences and adult leaders who were controlling and manipulative. These interviews underscore the potential developmental functions of the out-of-school context, at least for adolescents.

Other theoretical work has identified psychological processes that are salient for children in the out-of-school context. Three processes warrant particular attention: (a) engagement, effort, and motivation, (b) quality of relationships with adults, and (c) the quality of peer networks.

C. FLOW

Flow theory (Csikszentmihalyi, 1990) describes prototypical experiences of intrinsic motivation that are reported by respondents to be deeply involving and enjoyable, that engender full concentration, and that present a perceived balance between challenge and skill. This combination of experiences is then posited to propel or push development forward. Larson (2000) extended consideration of flow to the out-of-school context in general and voluntary structured activities in particular. In a series of studies in which youth reported their activities and experiential states at randomly sampled moments, Larson determined adolescents' psychological states at school, at home, and during structured voluntary activities. During classwork and homework, adolescents reported high levels of concentration and challenge, but low levels of intrinsic motivation. While watching television and while hanging out friends, students reported low concentration and effort but high intrinsic motivation. Larson noted that it was primarily in voluntary structured activities that youth reported high levels of intrinsic motivation and high levels of effort and concentration, i.e., flow. Larson has posited that this combination of high effort, concentration, and intrinsic motivation fosters positive youth development, particularly the development of *initiative*, defined as motivation from within to initiate and sustain effort towards a challenging goal over time.

Vandell and colleagues (in press) obtained similar results in a sample of low-income middle school students ($N = 199$) whose experiences were sampled 35 times during 1 week in the fall and 35 times during 1 week in the spring (over 12,000 signals). Substantial differences were found in the adolescents' activities, feeling states, and motivation at after-school programs vs. elsewhere. Students were more likely to engage in academic/arts enrichment, organized sports and physical activities, and volunteer service at programs vs. elsewhere, and they reported being more engaged, exerting more effort and concentration, and caring more about their activities at programs than elsewhere.

Another developmental process that warrants further attention is Larson's (2000) postulated "arc of effort." Many structured activities involve a clearly specified product or culminating event—a musical event or show, an end-of-season sports competition, a publication. Preparation for the culminating event occurs over a fairly extended period of time and involves concentrated and coordinated efforts. Systematic research is needed that considers the effect of out-of-school experiences that incorporate these arcs of effort to foster the development of initiative.

D. ATTACHMENT THEORY

Attachment theory (Bowlby, 1969; Pianta, 1999) emphasizes the importance of sensitive and supportive relationships with parents and teachers in providing a secure base from which children learn and develop. Supportive relationships with adults is emerging as an essential element in after-school settings as well. For example, in analyses of a subsample of 137 children in the NICHD Study of Early Child Care who attended after-school programs, Vandell, Pierce, and Lee (2005) found relationships with after-school program staff to predict child developmental outcomes over and above family and school factors. Less conflictual relationships with program staff predicted higher reading and math achievement, and fewer problem behaviors according to both mothers and teachers at the end of first grade, even after controlling for children's prior functioning in these domains, family factors, and instructional quality of the children's first grade classrooms.

Similar findings were obtained in a second study that considered three aspects of the after-school environment—quality of interactions with program staff, quality of peer interactions at the programs, and organization of program activities (Pierce, Hamm, & Vandell, 1999). All three areas were significant predictors of children's behavior and performance in their first grade classrooms, controlling for other program elements and family factors. Children who experienced more positive interactions with program staff were reported by first grade teachers to have fewer internalizing and externalizing behavior problems, whereas children who experienced more negative interactions with program staff had lower math and reading performance, controlling for quality of peer interactions and program activities. Relations with adult staff also appear important for older children and adolescents. Mahoney, Schweder, and Stattin (2002) found that students reported less depressed mood when they attended programs in which there was high support from the activity leader.

E. FRIENDSHIPS, AFFILIATION, AND PEER NETWORKS

Supportive relationships with peers and positive peer networks also are known to play a central role in the healthy development of children and adolescents, whereas hostile relationships (bullying, victimization, mutual antipathies) and deviant peer networks place young people at risk for behavioral, emotional, and academic problems (Hartup, 1996; Rubin, Bukowski, & Parker, 1998; Kiesner, Poulin, & Nicotra, 2003; Brown, 2004). Peers have often been studied in the school context, but there is growing evidence that peer networks are unique predictors of child emotional well-being in both in- and out-of-school contexts.

Research by Eccles and colleagues (Eccles & Barber, 1999; Barber, Eccles, & Stone, 2001) has highlighted a synergistic connection between extracurricular activities, peer group composition, and identity development. These investigators observed that adolescents who participated in certain extracurricular activities had more academic friends and fewer friends who skipped school or used drugs. For example, students who participated in prosocial activities identified themselves and their peer group as academically oriented and relatively unlikely to engage in risky behaviors, which in turn was associated with better academic achievement and a greater likelihood of attending college.

Peer relationships in after-school programs serving elementary school children appear to serve similar affiliative functions. Sandstrom and Coie (1999) found that rejected children who regularly attended programs became less rejected over time, a change that was not observed in rejected children who did not attend programs. These investigators speculated that the programs offered these children opportunities to improve their social standing through interaction with competent peers who model appropriate social behavior.

Other out-of-school contexts, however, provide opportunities for contacts with deviant peers and are associated with increased antisocial behavior and poorer academic performance. Osgood and colleagues (Osgood *et al.*, 1996) focused specifically on processes that may explain reported relations between unsupervised peer contact and antisocial behaviors. These authors argued that the mechanisms of the association are explained by routine activity theory, which posits that unstructured socializing with peers in the absence of authority figures presents opportunities for deviance. The theory has several components. First, it is easier to engage in deviant acts, and these acts are more rewarding, in the presence of peers. Second, the absence of authority figures reduces the potential for social control of deviant acts. Third, the lack of structure leaves time available for deviant behavior.

Findings consistent with routine activity theory are seen in research conducted by Mahoney and his colleagues (Mahoney & Stattin, 2000; Mahoney, Stattin, & Magnusson, 2001). These investigators contrasted young adolescents who participated in a specific unstructured activity—hanging out with peers at youth recreation centers in which there were few rules and little adult guidance—with adolescents who did not attend the centers. Adolescents who hung out at the youth centers exhibited more antisocial behavior (e.g., shoplifting, getting drunk, destroying things, fighting, skipping school) than youth who did not go to the centers, and cumulative juvenile offense frequency at age 13 was greater for boys who attended the youth centers more frequently. Juvenile offending that persisted into adulthood was particularly marked for youth who hung out with peers at the centers on a frequent basis. In fact, youth center participation was the single best predictor of criminality at age 30, controlling for SES and child behavior and competence at age 10. The investigators speculated that participation in the youth centers increased antisocial behavior and criminal offending through peer socialization. The adolescents who hung out at the youth centers reported an older peer network composed of students who did poorly in school and were out on the town at night more often.

Thomas Dishion's work (see Dishion, McCord, & Poulin, 1999) suggests that even structured after-school programs can have unintended negative consequences. In a series of studies, high-risk youth were observed to reinforce each other's problem behaviors: deviant youth who participated in an intervention program that included activities with peers did not post the same reductions in problem behaviors as deviant youth who participated in a family-based intervention. Similar unintended consequences in an intensive group-tutoring program for low-achieving high school students were reported by Catterall (1987), who observed that mutual bonding among the low-achieving students appeared to increase (not decrease) their feeling of alienation at school. Thus, a critical issue for future research is determining effective strategies for developing positive (and not negative) peer networks in the after-school programs.

II. After-School Programs

In this section, we consider one particular type of out-of-school arrangement—after-school programs. After-school programs are organized group activities that occur on a regular basis, typically 4 or 5 days a week. A primary goal of after-school programs traditionally has been to provide supervision to children of working parents (Capizzano, Tout, & Adams, 2000; Committee on Work and Family Policies, 2003; Kleiner, Nolin, & Chapman, 2004). In 2002, programs served as the primary after-school arrangement for 21% of children ages 6–9 years and 14% of children ages 10–12 years whose mothers were employed (Kleiner, Nolin, & Chapman, 2004).

The roles and functions of after-school programs, however, are in some flux. During the 1990s, the vast majority of families using programs paid fees or tuition for their children to attend (Smith, 2000, 2002). Initiatives such as the federally funded 21st Century Community Learning Centers (CLC), however, are free of charge and have substantially expanded the numbers of low-income children being served by programs. A related change is in the location of programs. The majority of the contemporary programs are now being housed in schools (Kleiner, Nolin, & Chapman, 2004), a shift from the 1990s when 28% were school-based and the majority were housed at child care centers, community centers, and churches (Seppanen *et al.*, 1993). Changes also are occurring in program activities. Traditionally, programs offered an array of activities (sports, arts and crafts, games, music, drama, theme-based activities, homework time). However, programs such as the 21st Century CLCs have adopted a decidedly academic focus and emphasize preparation for high-stakes tests, tutoring and remediation, and academic skill development. Researchers have only just begun to study the effects of these changes in program goals and philosophy.

A. WHO ATTENDS PROGRAMS?

A number of family and child factors are associated with children's participation in after-school programs. Consistent findings reported in several national surveys (Capizzano, Tout, & Adams, 2000; Smith, 2002; Kleiner, Nolin, & Chapman, 2004) are that children are more likely to attend programs when mothers are employed and are employed for more hours, when mothers work traditional schedules rather than nontraditional schedules, and when children reside in single-parent households rather than two-parent households. In the 2001 National Household Education Survey (Kleiner, Nolin, & Chapman, 2004), participation in programs was not related to family income or parental education although such relations were evident in the 1990s.

The most consistent child characteristic associated with program participation is that attendance is highest for children in the early primary grades, with steady decreases thereafter. Other child characteristics such as ethnicity and child adjustment predict program participation in some studies but not others. These differences in child demographics are consistent with the changing mission of after-school programs. Programs are increasingly serving children of color and children with academic deficits.

B. DO AFTER-SCHOOL PROGRAMS AFFECT CHILD DEVELOPMENTAL OUTCOMES?

Since 1995, there has been a virtual explosion in research reports that consider possible effects of after-school programs on children's development. In spring 2004, a website supported by the Harvard Family Research Project (http://www.gse.harvard.edu/hfrp/eval.html) provides links to 64 program descriptions and 104 program evaluations. In addition, the number of articles published in peer-reviewed scientific journals and in edited volumes has increased (see Committee on Community-Level Programs for Youth, 2002; Committee on Family and Work Policies, 2003; Mahoney *et al.*, 2005). These research studies reflect four methodological strategies—ethnographies and other narrative descriptions, correlational designs, quasi-experimental designs, and random assignment experiments.

1. Program Descriptions

One set of studies provides descriptions of specific programs located in particular schools, libraries, or communities (McLaughlin, 2000; Garner, 2002). These reports do a wonderful job of conveying the richness of programming. One report, for example, describes a program in a remote rural community in which middle school students serve as the local computer consultants who design webpages and digitize materials for the local library (Garner, Zhao, & Gillingham, 2002). Another describes an inner-city dance program (Vandell *et al.*, 2004). Still other descriptions portray programs that focus on leadership skills and community service (Youniss *et al.*, 2002). Readers interested in an introduction to the array of after-school program models are referred to these reports.

These descriptive accounts reflect the features of community-based programs identified by the Committee on Community-Level Programs for Youth (2002) as important for positive youth development: physical and psychological safety; appropriate structure; supportive relationships; opportunities for belonging; positive social norms; support for efficacy and mattering; opportunity for skill building; and integration of family, school, and community efforts. These descriptions also are consistent with the successful programs identified in Fashola's (2002) review of 34 programs that served at-risk students.

McLaughlin's (2000) 10-year study of exemplary community-based after-school programs (McLaughlin, Irby, & Langman, 1994) also provides rich descriptions of programs that created intentional learning environments in which adults made ongoing efforts to make activities accessible and challenging for all attendees. Activities were part of concentrated programs that aimed to deepen skills and competence through intense engagement in a specific area. Materials were adapted to the interests and strengths of the attendees. Youth had a central role in designing activities and establishing and enforcing rules in the group. Adult leaders were able to embed opportunities within the program activities to build a range of academic competencies and life skills. Programs often were organized around culminating events. McLaughlin argues that participation in these programs had positive effects on the youth who attended the programs. Compared to the "typical" responses reported in the National Educational Longitudinal Survey (NELS), program participants were more likely to obtain good grades, to expect to graduate from high school, and to expect to go to college. They also were more likely than typical NELS respondents to report feeling good about themselves and to indicate higher levels of efficacy.

Although descriptive studies are suggestive, they are limited in some important respects. They do not systematically test the effects of variations in program offerings nor do they address the possibility of selectivity bias (i.e., program participants differ *a priori* from nonparticipants in ways that account for ostensible program effects). Program descriptions also leave unanswered the question of how representative these programs are of after-school programs in the United States. Finally, descriptive studies do not systematically test for effects of varying quality programs. It seems unlikely that poorer quality programs will have the same effects as programs that are of higher quality. Attending poor-quality programs might even have detrimental effects.

2. Correlational Studies

A second set of research studies has used correlational designs to test hypotheses about relations between program characteristics and child developmental outcomes. Researchers using these designs face methodological challenges because children's participation in particular arrangements is influenced by family circumstances and beliefs, by child characteristics and interests, and by school and community resources (Vandell & Shumow, 1999). Research on the out-of-school context has lagged behind the research examining the effects of early child care in the assessment of and control for selection factors. Much of the early research examining out-of-school contexts did not control for selection. Subsequent research, however, has sought to control for family and child selection.

In one study of 150 predominantly middle class children who attended after-school programs on a daily basis (Pierce, Hamm, & Vandell, 1999), variations in program quality were examined in relation to child developmental outcomes after controlling for parenting practices, parental income and education, and children's functioning at the beginning of the school year. Boys who attended after-school programs in which there were more positive emotional climates were reported by their first grade teachers to exhibit fewer problem behaviors at school in comparison to boys who attended programs with less positive climates. A more negative emotional climate in the after-school programs, in contrast, predicted poorer academic performance at school. Additionally, boys who attended programs rated as fostering autonomy and choice among activities had better social skills according to their first grade teachers. In a follow-up study of these same children, observed program quality in Grades 2, 3, and 4 was examined in relation to child developmental outcomes (Vandell & Pierce, 2001). Controlling for family background factors and prior child functioning, Vandell and Pierce found that children who attended higher quality after-school programs (defined as positive interactions with staff, positive interactions with peers, opportunities for autonomy, and developmentally appropriate activities) obtained better work habits ratings and higher academic grades in Grades 2–4.

Relations between program quality and child developmental outcomes have been reported in other samples as well. Vandell, Shumow, and Posner (2005) described pervasive differences in program quality at two programs serving low-income children in terms of staff–child interaction, emotional climate, and provisions for child autonomy that were further moderated by child characteristics. Staff members at the poor-quality program were particularly negative when interacting with children with longstanding problem behaviors, whereas staff in the other program appeared particularly sensitive in their interactions with troubled youth. Across the school year, students who attended the poorer quality program showed declines in their report card grades and ratings of work habits and peer relations, whereas the performance of children in the more supportive program either improved or remained the same. Family and school environments were very similar for the two groups.

These associations between program quality and child developmental outcomes are troubling because the available evidence suggests that program quality in the United States is highly variable and much of the care is not good quality. The National Survey of Before- and After-School Care found substantial variability in the structural and caregiver characteristics of after-school programs (Seppanen *et al.*, 1993). In that survey, conducted in the early 1990s, staff turnover averaged 60% a year. Child:staff ratios ranged from 4:1 to 25:1. Education of the staff varied from less than a high school degree through a graduate degree. Subsequent reports from parent and student surveys (Kleiner, Nolin, & Chapman, 2004) suggest that programs continue to vary widely on these dimensions. In one study that conducted on-site visits and participant surveys at 15 programs (Gambone & Arbreton, 1997), 36% of those surveyed reported feeling less safe at the program than elsewhere, 39% felt never or almost never valued at the program, 35% felt there were no adults at the program to whom they could turn, and 40% reported never or almost never having input in program activities. These variations in program features coupled with the evidence of relations between program quality and child developmental outcomes underscore the need to include assessments of program quality in evaluations of program effectiveness.

3. *Quasi-Experimental Studies*

A third set of research studies have examined effects of after-school programs on children's developmental outcomes using quasi-experimental designs. In these studies, investigators contrast the treatment group (i.e., program children) and comparison groups composed of students who are not enrolled in after-school programs. Some of these studies have sought to determine whether the children in these groups are comparable in other respects, save differences in program participation. Some studies have included assessments of program quality, although many have not. The studies conducted by developmentalists (Posner & Vandell, 1994; Marshall *et al.*, 1997; Pettit *et al.*, 1997) have included an eclectic set of programs sponsored by various organizations or groups. Other quasi-experimental studies (Huang *et al.*, 2000; Grossman *et al.*, 2002; Welsh *et al.*, 2002) have been conducted as part of evaluations of particular after-school programs that were funded by the programs or by foundations as part of a delivery of services.

In a study conducted in nine schools serving low-income children who lived in unsafe neighborhoods, Posner and Vandell (1994) reported that third grade children who attended a variety of school- and community-based after-school programs had fewer antisocial behaviors and better reading and math grades, work habits, emotional adjustment, and peer relationships than children who were in self-care, sitter care, or parental care after school, controlling for child ethnicity, family income, and maternal education. Marshall and colleagues (1997) also detected relations between attending after-school programs and children's behavioral adjustment, in this case controlling for family structure, family income, maternal education and employment status, maternal depression, and quality of the neighborhood environment: low-income children attending programs had fewer internalizing behavior problems than did those children in unsupervised care.

As noted previously, other quasi-experimental studies are evaluations of particular programs. In a large-scale evaluation of 96 programs sponsored by The After-School Corporation (TASC) in New York City, Welsh *et al.* (2002) reported that low-achieving students, Black students, Hispanic students, and English Language Learners were more likely than other students to benefit from active participation in the TASC programs, as evidenced by greater gains in math achievement relative to their peers. Changes in reading and math achievement for highly active participants ($N = 12{,}973$), active participants ($N = 17{,}805$), nonactive participants ($N = 8104$) and nonparticipants ($N = 39{,}870$) were examined. Students who were active participants in TASC programs for more than a year showed significantly greater gains in math achievement than did similar nonparticipating classmates. To make the nonparticipant pool as similar as possible to TASC participants, the investigators controlled statistically for differences in demographic characteristics, grade level, and initial test scores. Contamination of the treatment and comparison groups was dealt with by setting (and applying) enrollment and attendance criteria.

A second large-scale evaluation (Grossman *et al.*, 2002) was conducted of the Extended Services School Initiative, comprised of 60 after-school programs in 15 states. All of the programs were school-based and offered a range of activities including academic enrichment, sports, and recreation. The majority of the participants were children of color; 72% were eligible for free or reduced lunch. Student outcomes were evaluated in a subset of 10 programs in six cities. Baseline and follow-up surveys were administered to students in Grades 4–8, and program attendance data were collected. Controlling for baseline performance, students who regularly attended the after-school programs were less likely to skip school and more likely to report being "very proud to belong to their school."

A third report (Huang *et al.*, 2000) summarizes an evaluation of the L.A.'s BEST program, a 15-year collaboration between the Los Angeles public schools and the City of Los Angeles that primarily serves low-income students of color. Program activities are comprehensive and include homework assistance, library activities, field trips, performing arts, recreational activities, and educational enrichment activities. The evaluation included 4312 students in Grades 2–5. Reading and math grades and test scores were tracked over a 5-year period. Longer term program participation was associated with higher scores on standardized tests of mathematics, reading, and language arts, controlling for gender, ethnicity, and language status. Program participation also was associated with subsequent gains in school attendance, controlling for gender, ethnicity, income, and language status. A path analysis indicated that higher levels of participation in the L.A.'s BEST program resulted in better school attendance that, in turn, related to higher scores on standardized tests of academic achievement.

The highest profile and largest evaluation to date involved the 21st Century CLC. CLCs were created by the US Congress in 1994 to make greater use of school buildings when school is not in session. Initially funded with $750,000 as a demonstration project, program funding increased to $40 million in 1998, $200 million in 1999, and $1 billion in 2002 when funds were provided to 7500 school-based programs in more than 1400 communities. Prior to 2002, the US Department of Education made competitive awards to designated agencies (primarily public schools) for 3 years, and grantees were not required to match federal funds with state or local funds. In January 2002, the No Child Left Behind Act converted the CLC to a state formula grant. For the new program, each state is allocated funds and makes its own awards.

The national CLC evaluation (Dynarski *et al.*, 2003) was awarded in 1999 and data collection began in 2000, focusing on the first three cohorts of CLC programs. The evaluation of the middle school programs, conducted in 34 school districts and 62 programs, utilized a quasi-experimental design to compare program students ($N = 1750$) with comparison students ($N = 2437$). Baseline data were collected for the two groups in fall 2000 and follow-up data were collected in spring 2001. Some effects associated with program participation were detected. In the first-year follow-up, grades in math were significantly higher for youth in the CLC participant group, with larger effects detected for Black and Hispanic students than for White students. Teachers reported less absenteeism in participants compared with nonparticipants. In other respects (feelings of school safety, goal setting, homework completion, other academic grades), however, the two groups did not differ.

This evaluation has been subject to intense scrutiny and several detailed critiques have been written (Bissell *et al.*, 2003; Mahoney & Zigler, 2003; Vandell, 2003). These critiques noted that the comparison group was substantially more advantaged than the treatment group: the mean reading achievement score for the treatment group at baseline was the 40th percentile, whereas the mean score for the comparison group was the 50th percentile at baseline, indicating that the comparison students were more academically competent than the treatment group. These substantial baseline differences make interpretation of the outcome findings problematic, because the researchers did not control for these baseline differences in the impact analyses, nor did they assess change in scores over time. A second concern is that these newly established CLC programs were not yet consistently implementing high-quality activities. The implementation findings that were included in the report suggest problems with program attrition, attendance, and staff turnover. Other evaluations (Huang *et al.*, 2000; Vandell & Pierce, 2001; Welsh *et al.*, 2002) have found that positive program effects are more likely to occur in established programs that have offered high-quality activities over time as opposed to programs that are just getting underway. A third concern is that the CLC programs were required to adopt an overly narrow academic focus that does not capitalize on the particular strengths of effective after-school programs.

4. Experimental Studies

There have been a handful of random assignment experimental studies of after-school programs (Ross *et al.*, 1992; Morrison *et al.*, 2000; Cosden *et al.*, 2001; Dynarski *et al.*, 2003). These studies have focused on the effects of academically oriented after-school programs on academic performance of low-income children. Experimental studies of comprehensive, project-oriented programs have not been conducted to date.

An experimental design was used to assess the impact of the Cooke Middle After School Recreation Program (Lauver, 2002). This recreation-based program was offered in the evening hours (5–7 p.m.) to complement a more academic-oriented program. The program was designed to provide opportunities for physical exercise and a safe environment. From the pool of interested students, some were randomly assigned to the program group ($N = 124$) and others to a control group ($N = 98$) that did not differ in student grade level, age, race, gender, family structure, educational attainment of parents, or home language. Program students reported spending more time on homework than did students in the control group, and were more likely to report that they wanted to attend college or job training after high school. Within the program group, those students who attended the program more often also attended school more and spent more time on homework. No differences in GPA or standardized test scores were detected between the two groups at the end of the first school year, however.

Other experimental studies of after-school programs have involved smaller samples of children and focused on academic skills. In a study of 35 educationally at-risk students who attended a 3-year homework club in Grades 4–6

(Cosden *et al.*, 2001), students who attended 77% or more of the homework sessions, which were offered 3–4 days a week by a credentialed teacher, obtained higher scores on reading, math, and language achievement tests than children who participated less consistently. In this experimental study, participation in alternative programs and selective attrition were critical issues: 31% of the control group participated in other after-school programs, and 58% of the program students dropped out of the target program.

By far, the largest experimental study is the elementary school component of the 21st Century CLC evaluation (Dynarski *et al.*, 2003). A random assignment experimental design was used to assess program effects in 18 programs in seven school districts. The treatment group consisted of 403 children who were enrolled in one of the CLCs and a control group was composed of 226 children on the program waiting lists. The selected programs were more urban and served a larger percentage of minority students than the average CLC elementary program. Thus, the evaluation can be viewed as a test of program impacts for an at-risk sample. Results indicated that the programs did not appear to improve student effort at school: there were no differences in homework completion between the treatment and control groups, and program participation had no effects on reading or math grades or reading test scores.

At first blush, these findings suggest no academic benefits of the CLC programs, at least after the first year. It should be noted, however, that concerns have been raised about the elementary evaluation (Bissell *et al.*, 2003; Mahoney & Zigler, 2003; Vandell, 2003). A substantial proportion of the programs (4 of 18 programs, or 22%) were targeted to adults in the community (another focus area in the CLC charge), and children attended the center only when they accompanied a parent or grandparent. It is not clear why these adult-focused programs were expected to impact child outcomes at the end of Year 1. Second, critics have questioned the emphasis on test scores and academic outcomes, especially after Year 1. A theory of change proposed by Vandell and colleagues (2004) postulates a cascade of effects in which programs first impact school attendance, work habits, and teacher–child relationships, which then affect academic performance and achievement. A third concern is that these CLC elementary programs were relatively new programs that had not yet stabilized in offering consistently high-quality programming. Finally, inspection of student performance at baseline and at the end of school year showed a widening of a performance gap that favored the treatment group. The math scores of the treatment group increased during the evaluation year, whereas the math scores in the control group decreased during the evaluation year. Unfortunately, the study authors did not test these changes in test scores to determine whether they were statistically significant.

Identification of factors that succeed in bringing children into programs and maintaining their participation is clearly needed. These factors may be the individual differences in children (or families) who are more motivated or academically inclined to seek out programs in their community, but program offerings and staff outreach at schools and in neighborhoods also may be important. A related issue concerns what is needed to sustain youth participation over time. If programs fail to offer activities that are interesting to children or if staff is detached or punitive, then children may well drop out. The research evidence pertaining to structured activities that is reviewed in Section III suggests some potentially effective strategies to be studied.

III. Structured Activities

A second type of out-of-school context is *structured activities and lessons*, which are organized activities that focus on particular content domains such as individual and team sports, a musical instrument, or community service. These activities can occur weekly, monthly, or even several times a week. Although working families sometimes use activities as a way to supervise their children during the out-of-school hours, their primary function is enrichment and enjoyment. According to the 2001 National Household Education Survey (Kleiner, Nolin, & Chapman, 2004), 38% of children in kindergarten through eighth grade participate in structured activities at least once a week, with sports reported as the most common activity.

A. WHO PARTICIPATES IN STRUCTURED ACTIVITIES?

Children's participation in structured activities is predicted by both family and child factors. Participation is higher in households with higher family income, higher parental education, two parents, and employed mothers (Hofferth *et al.*, 1991; Smith, 2002; Tout, Scarpa, & Zaslow, 2002; Kleiner, Nolin, & Chapman, 2004). In one national survey (Tout, Scarpa, & Zaslow, 2002), for example, nonpoor families had much higher rates of participation in extracurricular activities after school than poor families (90% of youth ages 6–17 vs. 65% of poor nonwelfare families and 59% of welfare families).

Processes within the family context also are associated with children's participation in structured activities. Parental modeling and instrumental and emotional support appear to be important. Youth spend more time in structured activities when their parents are more involved in community activities, and they spend less time in structured activities when their mothers spend more time watching television (Fletcher & Shaw, 2000). Additionally, parental warmth (Fletcher, Elder, & Mekos, 2000; Larson, Dworkin, & Gillman, 2001), parental encouragement of activities (Mahoney & Stattin, 2000), and specific instrumental support such as driving children to activities (Anderson *et al.*, 2003) are linked to youth being more involved in structured activities.

Consistent relations between child characteristics and structured activities have been found. For example, Kleiner, Nolin, and Chapman (2004) reported that boys are more likely to be involved in sports, whereas girls are more involved in clubs and lessons; and White adolescents have the highest participation rates in all categories of structured activities, whereas Hispanic youth show the lowest participation rates. Rates of participation also are higher in middle school than in elementary school, and surveys of adolescents suggest that high school students participate in more structured activities than middle school students. Parents report that 57% of youth ages 12–17 years participated in team sports and 60% participated in clubs or organizations after school or on weekends in the past year (Brick *et al.*, 1999). According to the National Longitudinal Survey of Youth, 60% of high school sophomores and 70% of seniors participate in at least one extracurricular activity (Zill, Nord, & Loomis, 1995).

Interviews with students about their reasons for participating (and not participating) in particular activities can inform practitioners and policy makers about why some structured activities are more successful than others in attracting and sustaining involvement. Studies of elementary (Casey, Ripke, & Huston, 2005) and middle school (Hultsman, 1992) students reveal that joining an activity is related to program fees and costs, parental support (and lack of support), transportation, cultural attitudes, and needs for older children to provide child care for younger siblings. These same factors were cited as reasons for dropping out of an activity. Additional factors associated with dropping out include having friends drop out of the activity and concerns about not being "good enough" to continue the activity. Still other factors have been reported by Butcher, Lindner, and Johns (2002) for dropping out of organized sports teams: an overemphasis on winning, a lack of fun, lack of interactions with friends outside the sport, and a desire to have fewer directions from adults.

As children move into middle school and high school, access to some structured activities depends on children's skill levels. Try-outs and auditions determine placement on sports teams and musical performance groups. Athletics, fine arts, and cheerleading are the most closed activities; newspaper and yearbook, the most open. The likelihood that students will participate in an activity in high school is greater if they have experience with the activity in middle school (McNeal, 1998).

Features of the school context also are linked to participation in structured activities. Students are more likely to participate in activities in smaller schools compared to larger schools, in rural and suburban schools vs. urban schools, and in private schools compared to public schools (Marsh, 1992). These associations likely reflect the availability of extracurricular activities at the school as well as greater encouragement of activity participation in smaller schools.

In summary, there appear to be substantial differences in participation in structured activities, with low-income youth appearing to have much less access to these activities. Although some of these variations may reflect differences in cultural values and interests, parents and youth cite concerns about cost, transportation, and safety as the primary barriers. In Section III.B, we consider implications of a lack of access to activities in terms of developmental outcomes.

B. IS PARTICIPATION IN STRUCTURED ACTIVITIES RELATED TO CHILD DEVELOPMENTAL OUTCOMES?

1. Social Competencies

A number of investigators have examined participation in structured activities in relation to social competencies and behavioral outcomes. Consistent with the proposition that structured activities provide connections with socially and academically competent peers, Sandstrom and Coie (1999) detected relations between participation in structured activities and changes in peer relationships over a 2-year period. Rejected children who participated in extracurricular activities in fourth grade, compared to rejected children who did not, obtained higher social preference scores in Grade 5, controlling for Grade 4 social preference scores. The authors speculated that contact with peers in structured activities outside the classroom provided rejected children with opportunities to learn social norms and skills, and perhaps to showcase strengths that are not apparent in the classroom setting.

Posner and Vandell (1999) reported relations between structured activities and socioemotional outcomes in a sample of low-income urban children. Controlling for child and family factors and teacher-rated emotional adjustment in third grade, cumulative amount of time spent in extracurricular enrichment activities at after-school programs in Grades 3–5 predicted positive changes in children's emotional adjustment in Grade 5.

Similar findings were obtained in a study of eighth grade Swedish youth who participated in structured activities (Mahoney & Stattin, 2000). The structured activities were based in the community and included sports, music, theater/fine arts, hobbies, church, scouting, and politics. The activities met at least once a week and involved peers of the same age and an adult leader. Students who participated in a structured activity engaged in less antisocial behavior such as shoplifting, getting drunk, destroying things, fighting, and skipping school compared to the students who were not involved in a structured activity, controlling for parental education.

In other work, Mahoney (2000) considered relations between participation in structured activities during middle and high school and outcomes in young adulthood (ages 20 and 24). Four groups of youth were identified based on cluster analyses of physical maturity, aggression, popularity, academic competence, and SES. Two of these groups were high in competence, a third group displayed a moderate level of competence, and the fourth group was composed of students who evidenced multiple risks (physically mature, aggressive, less popular, less academically competent, lower SES). Students in the high-risk group were more likely than students in other configurations to show antisocial patterns in young adulthood. However, participation in school-based extracurricular activities in at least 1 year during Grades 6–10 was associated with reduced rates of school dropout and criminal arrest among the high-risk students. There were no differences within the high-risk group on the indicators of competence that could account for the activity-adjustment link.

Other research has examined the amount of time spent in structured activities. Pettit *et al.* (1997) found that low-income children who attended activity-oriented programs evinced more positive developmental outcomes than children not in these programs. In this study, children who had engaged in organized activities in first, third, and fifth grades displayed fewer externalizing behavior problems in Grade 6 according to their teachers, controlling for child gender, family SES, and child externalizing behavior in kindergarten. The obtained relation was curvilinear: small to moderate amounts of organized activities were associated with fewer problem behaviors than large amounts of time in these activities.

The available, albeit limited, experimental data are consistent with the correlational findings. Jones and Offord (1989) reported positive effects of structured activity participation in a primary intervention program for low-income youth ages 5–15. The treatment consisted of an after-school program located at a public housing project. The program provided youth with opportunities to participate in various skill development activities such as sports, guitar, ballet, baton, scouting, and other activities, directed by highly skilled adults. The researchers reported significantly less antisocial behavior (fewer legal charges against youth, fewer security incidents in the housing complexes) among youth in the treatment group.

2. Academic and Cognitive Outcomes

Other investigators have reported relations between structured activities and academic and cognitive outcomes. Findings from the NICHD Study of Early Child Care, a large multi-site prospective longitudinal study, suggest that structured activities confer positive academic advantages in the primary grades (NICHD Early Child Care Research Network (ECCRN), 2004). Children who consistently participated in structured activities during kindergarten and first grade obtained higher standardized math scores at the end of first grade compared to children who never or only sporadically participated, controlling for children's prior academic achievement, ethnicity, and gender as well as numerous family factors. Almost all of these activities had a nonacademic focus—organized team sports such as soccer, individual sports such as karate and swimming, music lessons, and clubs such as Daisy Scouts. For the most part, the children were not "over-programmed"; they typically participated in a single activity for fewer than 3 h each week.

Casey *et al.* (2005) reported similar relations in a sample of low-income children ages 9–13. Children who participated in structured activities (coached sports, religion classes, clubs, or youth groups) had higher achievement test scores and higher ratings of social behaviors from their teachers than did children who did not participate in structured activities. These academic findings, as well as those of the NICHD ECCRN (2004), are intriguing because the activities were not explicitly academic in nature.

In a longitudinal study of over 10,000 youth, Marsh (1992) examined changes in academic and socioemotional outcomes from Grade 10 to Grade 12. An activity score was created based on the number and level of involvement in structured activities at school and in the community. In analyses that controlled for background variables and Grade 10 adjustment, Marsh found that youth who engaged in more activities in Grade 10 showed positive changes in Grade 12 in terms of social self-concept, academic self-concept, educational aspirations, GPA, absenteeism from school, time spent on homework, taking advanced courses, and being on the academic track. Students from lower SES families appeared to derive the greatest benefits from structured activities. Marsh speculated that activity participation increases commitment to school by enhancing academic self-concept, which in turn mediates positive effects on other academic outcomes.

There is further evidence that structured activities may be a protective factor for students who are at risk of dropping out of high school. Utilizing a person-oriented approach that placed students into competence clusters, Mahoney and Cairns (1997) found that at-risk students (boys and girls who were aggressive and unpopular, and girls who evidenced poor achievement), compared to students who were more competent, were more likely to drop out of high school by Grade 11 if they were involved in only one activity during middle school and no activities during high school, controlling for parental education.

Research by Cooper and colleagues (1999) considered the effects of five types of after-school activities (school-based extracurricular activities, community-based structured activities, homework, television watching, and paid employment) on adolescents' academic performance. They tested the proposition that time spent in nonacademic activities such as structured sports, lessons, and groups displaces time for academic activities such as homework, resulting in poorer academic outcomes. Results indicated, however, that more time in extracurricular activities, other structured groups, and homework (and less time in working at a job or watching television) was associated with higher test scores and grades controlling for child gender, grade level, ethnicity, and receipt of free or reduced-price lunch. In these cross-sectional data, however, more academically competent adolescents may have had greater access to and interest in structured activities.

Other studies conducted with middle and high school populations have controlled for adolescents' prior adjustment, resulting in more confidence that participation in structured activities explains adjustment outcomes. For example, Jordan and Nettles (2000) utilized NELS 88 data to examine time spent in structured community-based activities such as lessons, sports, and community service during Grade 10 and outcomes in Grade 12. Controlling for child gender, family SES, school contextual variables (racial composition, size, proportion of students living in poverty), and prior socioemotional (self-concept) and academic (Grade 10 reading test scores) adjustment, Jordan and Nettles found that the amount of time spent in activities during Grade 10 predicted composite math and reading achievement scores in Grade 12 as well as greater self-reported preparation for class. Similarly, Darling, Caldwell, and Smith (2005) found that students who consistently participated in high school extracurricular activities (sports, performing arts, leadership groups, interest clubs), across two school years had higher grades, greater educational aspirations, and better attitudes toward school compared to students who did not participate in the activities, after controlling for adjustment in the first year and youth and family factors (grade, sex, ethnicity, parent education).

Finally, time spent in structured activities in middle and high school also has been associated with educational outcomes in young adulthood. Mahoney, Cairns, and Farmer (2003) examined educational status at age 20 relative to consistency of activity participation in Grades 7–8 (early adolescence) and Grades 9–10 (middle adolescence). In regressions that controlled for child gender, SES, interpersonal competence (low aggression and high popularity) in both middle school and high school, and educational aspirations at age 18, consistency of participation in both middle school and high school (none, 1 year, or 2 years at each school level) predicted enrollment in college at age 20. This association was strongest for the students characterized by low interpersonal competence, for whom college attendance was unlikely if they had participated in structured activities for less than 2 years and more likely if they had participated in all 4 years of Grades 7–10. Path analyses indicated that consistency of activity participation in both early and middle adolescence was associated positively with interpersonal competence in middle adolescence, educational aspirations at age 18, and educational status at age 20. The model was an adequate fit for the students in the high-competence group in Grades 7–8, but an exceptional fit for the low-competence group, indicating that at-risk students may derive particular benefits from participation in structured activities.

Investigators also have begun to examine differential effects associated with specific types of activities. In a cross-sectional study of fourth grade students, Fletcher, Nickerson, and Wright (2003) reported that participation in community-based clubs was associated with higher grades and higher teacher ratings of academic competence, whereas participation in organized sports was associated with greater social competence as rated by teachers, controlling for gender, ethnicity, and SES. Because children's functioning prior to Grade 4 was not assessed, one cannot rule out the possibility that children who are academically and/or socially skilled may elect to participate in different activities.

Morris and Kalil (in press) utilized cluster analysis in their examination of structured activity data collected as part of the Canadian Self-Sufficiency Project, an experimental antipoverty demonstration project conducted with low-income families of children ages 6–12 years. Five clusters were formed based on children's after-school activities: (a) high clubs (18% of the sample), (b) high sports (20%), (c) high sports and clubs (14%), (d) high sports, clubs, and lessons (16%), and (e) low sports, clubs, and lessons (32%; no-activity group). Research staff administered a measure of math skills to the children, and parent reports of child academic achievement, behavior problems, and prosocial behavior were obtained. In the substantive analyses, the investigators controlled for a number of characteristics that were related to the activity participation clusters: child gender and age, parental education, family income, and the number of young children in the household. Children who did not participate in structured activities had the poorest adjustment, and children who were involved in all three types of activities evidenced the best all-around adjustment as reflected by better math skills, fewer behavior problems, and more prosocial behaviors.

Other investigators have examined the developmental consequences of participation in specific types of activities during high school. Sports, in particular, have received a great deal of attention. One consistently reported negative outcome is alcohol use: participation in high school sports is associated with larger increases in alcohol consumption across two school years (Eccles & Barber, 1999; Darling, Caldwell, & Smith, 2005) and with greater use of alcohol into young adulthood (Barber, Eccles, & Stone, 2001), compared to participation in other activities or no activity participation. Mixed results have been reported regarding academic outcomes. Controlling for prior adjustment, Darling, Caldwell, and Smith (2005) found that sports participants did not fare as well in terms of grades, educational aspirations, and attitudes toward school as students who participated in nonsport activities at the school, although the sports participants did perform better than students who were not involved in any structured school-based activities. Schreiber and Chambers (2002), using NELS 88 data and controlling for gender and SES, found that participation in a greater number of school-based sports was associated with lower achievement test scores in Grade 8 and in Grade 10, whereas participation in nonsport school-based activities (e.g., band, math club, drama club, newspaper, debate) was associated with higher achievement scores in both grades.

Others have reported more positive links between sports and academic outcomes. For example, participation in school-based sports at the middle and high school levels is associated with a reduced likelihood of dropping out of school (McNeal, 1995; Mahoney & Cairns, 1997). For low-income urban youth in middle or high school, participation in sports (as well as other school-based activities) is associated with higher grades, controlling for gender, race, and school level (Pedersen & Seidman, 2005). And, Eccles and Barber (1999) reported positive changes from Grade 10 to Grade 12 in GPA as well as a greater likelihood of full-time enrollment in college after graduation for working class and middle class students who participated in sports during their sophomore year, controlling for Grade 10 GPA, gender, maternal education, and intellectual aptitude. These positive effects persisted into young adulthood, with Grade 10 sports involvement predicting completion of more years of post-high school education and higher college graduation rates (Barber, Eccles, & Stone, 2001).

A study by Guest and Schneider (2003) indicates that the effects of participation in high school sports may depend on the school context. In their study, students in Grades 10 and 12 who participated in sports attained a higher GPA and held greater educational expectations compared to students who did not participate in any extracurricular activities, controlling for gender, ethnicity, grade level, parent education, course sequence in math, and delinquency. However, the effects of sports involvement differed by the school context. Being seen by others in the school as athletic, which was linked to sports involvement, had a positive association with GPA in schools where a smaller proportion of students go on to college after high school graduation, but a negative association with GPA in schools where more than 80% of the student body attends college.

Other researchers have examined several specific types of activities in addition to sports as they relate to youth adjustment outcomes in high school. Eccles and Barber (1999) examined changes in adjustment associated with participation in particular types of school- and community-based structured activities, including prosocial activities (volunteering and community service), performing arts (band, dance, drama), leadership activities (student government, pep club, cheerleading), and academic clubs. The researchers found that participation in any of these activity types during Grade 10, controlling for gender, maternal education, and intellectual aptitude, was associated with a higher cumulative GPA in Grade 12. Involvement in leadership activities and academic clubs was associated with a greater likelihood of attending college full time after high school graduation. Students who were involved in prosocial activities reduced their use of alcohol and marijuana between Grades 10 and 12, controlling for substance use in Grade 10. The performing arts were associated with one negative outcome, an increase in the number of days that students skipped school during Grade 12 relative to Grade 10.

In a follow-up study conducted 6 years later, Barber, Eccles, and Stone (2001) found that the positive effects of Grade 10 activity participation were maintained into young adulthood. Participation in prosocial activities, leadership activities, and performing arts during Grade 10 was associated with higher rates of graduation from college, controlling for maternal education and student math and verbal ability. Socioemotional outcomes were examined also. Students who had participated in prosocial activities reported higher self-esteem 6 years later compared to participants in the other activities, and they continued to use less alcohol and marijuana than their high school peers. Students who had participated in the performing arts experienced the most negative outcomes: compared to students who participated in the other types of activities, performing artists reported greater use of alcohol 4 years after high school graduation, more suicide attempts, and more visits to a psychologist.

Eccles also has examined adjustment outcomes associated with participation in structured activities through cluster analysis. Bartko and Eccles (2003) performed a cluster analysis of extent of involvement (ranging from less than once per month to every day) in 11 types of school- and community-based activities during Grade 12. The analysis resulted in six activity clusters: *sports* (high involvement in sports, low involvement in other activities), *school* (high involvement in school-based clubs, homework, reading for pleasure; low involvement in other activities), *volunteer* (high involvement in volunteering and community service, low involvement in other activities), *high involved* (high involvement across activity types, especially community-based clubs but also sports, volunteering, homework, reading for pleasure, and religious groups), *work* (employed, low involvement in structured activities), and *uninvolved* (low involvement in structured activities, not employed). The authors then compared the activity clusters on a number of academic, behavioral, and psychological outcomes in Grade 12, controlling for gender and parental education. Students in the school and high-involved clusters had higher GPAs, and students in the uninvolved cluster had lower GPAs, than students in the sports, volunteer, and work clusters. Problem behavior was more evident in the sports, work, and uninvolved clusters and less evident in the school and high-involved clusters. Greater levels of externalizing behaviors were associated with membership in the work and uninvolved clusters. Internalizing behavior scores were higher for the uninvolved cluster, and lower in the sports and high-involved clusters, compared to the other activity clusters. Levels of depression were higher among the uninvolved students than among students who participated in activities or worked.

The overall picture gained from the work by Bartko and Eccles (2003), consistent with Morris and Kalil's (in press) findings in middle childhood, is that adolescents who are involved in a variety of structured activities evidence the best all-around adjustment, whereas adolescents who are not involved in any structured activities evidence the poorest adjustment. High school students who participate in school-based clubs also fare well. Sports participation to the exclusion of other types of activities is associated with both positive and negative adjustment indicators. It is important to note, however, that the associations obtained in this study could be due to selection factors, most notably prior adjustment that may have affected participation in the activities.

IV. Summary and Conclusions

Since the 1990s, there has been a growing recognition of the importance of the out-of-school context for children and adolescents. Fueled in part by family demographics that include substantial numbers of employed mothers and single mothers, in part by concerns about poor academic performance and problem behaviors, and in part by intensified efforts to find ways to promote positive youth development, researchers and practitioners have focused their attention on two particular out-of-school settings: after-school programs and structured activities.

The research findings pertaining to full-time (i.e., 5 days a week) after-school programs are mixed, which may reflect the substantial heterogeneity of the programs in terms of children being served, the types of activities offered, and the training and background of the staff. The federal funding of the 21st Century CLCs and various state and local initiatives has increased the numbers of low-income and English-learning students participating in after-school programs. A substantial number of programs are becoming more school-like. The available research suggests that (under some conditions) attending after-school programs is linked to improved social and academic outcomes. Children are more likely to show academic and social benefits when staff–child relationships are positive and nonconflictual, when programs offer a variety of age-appropriate activities from which children can select those of interest, and when children attend on a regular basis. The research findings about voluntary structured activities are more straightforward. Participation in these activities has been consistently linked to positive academic and social developmental outcomes in numerous studies. What appears to be key is that the activities are voluntary, are characterized by sustained engagement and effort, and provide opportunities to build or develop skills.

Although the available research has begun to inform our understanding of the out-of-school context, further research is sorely needed. First, there is a need for research to identify the social, cognitive, and linguistic processes by which participation in programs and structured activities influences child and youth developmental outcomes. For example, researchers need to consider the competitiveness of sport activities in relation to children's social and emotional functioning.

Researchers also might examine after-school experiences as settings in which complex thought processes can develop. Heath (1999) has conducted initial work in the area of language development by obtaining language samples during voluntary structured activities and analyzing their content. In the initial samples, students engaged in few sustained conversations on a topic and they frequently changed topics. After 3–4 weeks at the program, however, Heath noted substantial changes in the students' conversations and language. The use of conditionals (should, would, could) increased. She also noted increases in strategies to obtain clarifications from others and increases in the use of shifted registers and genres. Heath's (1999) linguistic analyses in conjunction with research that considers social and motivational processes underscore the broader point that the out-of-school context is complex and multi-layered and likely to be of substantial importance in the lives of children and youth.

Research is needed to identify other important developmental processes in programs and structured activities. A promising procedure for identifying these processes is experience sampling (Csikszentmihalyi & Larson, 1987; Larson, 1989). Experience sampling methodology allows researchers to collect systematic data about an individual's activities, thoughts, and affective states by obtaining reports from participants at multiple randomly sampled points in time. Participants are signaled to provide a report in a variety of ways, such as with beepers or alarm watches. This record of experiences is not usually captured by other data collection methods. For example, program observations provide data on observed activities, interactions, and program climate, but do not offer insights into students' feelings and experiences within the after-school environment. Questionnaire and survey data are retrospective, asking respondents to recall past experiences and feelings regarding their after-school activities. Experience sampling could be used to examine any number of processes in after-school programs and structured activities.

A better understanding of the effects of program content also is needed. Whether after-school programs should focus exclusively on enrichment activities or exclusively on academic activities, or include both enrichment and academic components, is the subject of heated debate. Some after-school scholars (Halpern, 1999; Heath, 1999; Eccles, in press) have argued forcefully that a focus on academics undermines the unique strengths and role of programs, and that programs should emphasize extracurricular enrichment activities. Others (Noam, 2004) have supported the move by policy makers and educators to make programs more academic, with an emphasis on homework help, tutoring, and preparation for academic achievement tests. The effects of different approaches to after-school programming have not been evaluated systematically. Research that describes, compares, and then tests effects of different program content models is needed to determine which types of programs are successful in attracting and keeping students (a necessary condition for programs to effect change), and to determine whether different types of programs are differentially associated with improvements in student outcomes such as school attendance, academic achievement, social competencies, and behavioral adjustment. A related question is whether structured activities that are obligatory or required have the same effects as voluntary structured activities do.

Researchers also should further examine the impact of different attendance patterns on child developmental outcomes. We do not have solid information about optimal intensity and duration of attendance in terms of outcomes. There are suggestions in the literature that long-term, frequent attendance at programs is associated with positive outcomes for low-income children. Research needs to examine whether these results hold for middle-income children and youth as well.

Finally, experimental studies should be conducted in which children and adolescents are randomly assigned to after-school programs and structured activities. All of the research to date on structured activities, and most of the research on after-school programs, has been nonexperimental, so questions about selection bias remain. Experimental studies in which children and adolescents are randomly assigned to participation in programs and activities would be a valuable next step in understanding relations between participation and child and youth outcomes. Such research should not be conducted until we have more information about the components of high-quality programming in terms of program content and developmental processes, however.

REFERENCES

Anderson, J. C., Funk, J. B., Elliott, R., & Smith, P. H. (2003). Parental support and pressure and children's extracurricular activities: Relationships with amount of involvement and affective experience of participation. *Applied Developmental Psychology*, *24*, 241–257.

Barber, B. L., Eccles, J. S., & Stone, M. R. (2001). Whatever happened to the jock, the brain, and the princess? Young adult pathways linked to adolescent activity involvement and social identity. *Journal of Adolescent Research*, *16*, 429–455.

Bartko, W. T., & Eccles, J. S. (2003). Adolescent participation in structured and unstructured activities: A person-oriented analysis. *Journal of Youth and Adolescence*, *32*, 233–241.

Bissell, J. S., Cross, C. T., Mapp, K., Reisner, E. R., Vandell, D. L., Warren, C., *et al.* (2003). *Statement from members of the scientific advisory board for the 21st Century Community Learning Center evaluation*, May 10.

Bowlby, J. (1969). *Attachment and loss*. New York: Basic Books.

Brick, P. D., Kenney, G., McCullough-Harlin, R., Rajan, S., Scheuren, F., Wang, K., *et al.* (1999). *1997 NSAF survey methods and data reliability: Methodology report no. 1*. Washington, DC: Urban Institute.

Bronfenbrenner, U. (1979). *The ecology of human development: Experiments by nature and design*. Cambridge, MA: Harvard University Press.

Bronfenbrenner, U., & Morris, P.A. (1998). The ecology of developmental processes. In W. Damon (Series Ed.) & R. M. Lerner (Vol. Ed., *Handbook of child psychology: Vol. 1. Theoretical models of development* (5th ed., pp. 993–1028). New York: Wiley.

Brown, B. B. (2004). Adolescents' relationships with peers. In R. M. Lerner & L. Steinberg (Eds.), *Handbook of adolescent psychology* (2nd ed., pp. 363–394). New York: Wiley.

Butcher, J., Lindner, K. J., & Johns, D. P. (2002). Withdrawal from competitive youth sport: A retrospective ten-year study. *Journal of Sport Behavior*, *25*, 145–163.

Capizzano, J., Tout, K., & Adams, G. (2000). *Child care patterns of school-age children with employed mothers*. *Occasional paper 41*. Washington, DC: The Urban Institute.

Carnegie Corporation. (1992). *A matter of time: Risk and opportunity in the out-of-school hours*. New York: Carnegie Corporation.

Casey, D. M., Ripke, M. N., & Huston, A. C. (2005). Activity participation and the well-being of children and adolescents in the context of welfare reform. In J. L. Mahoney, R. W. Larson, & J. S. Eccles (Eds.), *Organized activities as contexts of development: Extracurricular activities, after school and community programs* (pp. 65–84). Mahwah, NJ: Erlbaum.

Catterall, J. S. (1987). An intensive group counseling drop-out prevention intervention: Some cautions on isolating at-risk adolescents within high schools. *American Educational Research Journal*, *24*, 521–540.

Cohen, D. A., Farley, T. A., Taylor, S. N., Martin, D. H., & Schuster, M. A. (2002). When and where do youths have sex? The potential role of adult supervision. *Pediatrics*, *110*, 66–72.

Collins, W. A., Maccoby, E. E., Steinberg, L., Hetherington, E. M., & Bornstein, M. H. (2000). Contemporary research on parenting: The case for nature and nurture. *American Psychologist*, *55*, 218–232.

Committee on Community-Level Programs for Youth. (2002). *Community programs to promote youth development*. Washington, DC: The National Academies Press.

Committee on Family and Work Policies. (2003). *Working families and growing kids: Caring for children and adolescents*. Washington, DC: The National Academies Press.

Cooper, H., Valentine, J. C., Nye, B., & Lindsay, J. J. (1999). Relationships between five after-school activities and academic achievement. *Journal of Educational Psychology*, *91*, 369–378.

Cosden, M., Morrison, G., Albanese, A. L., & Macias, S. (2001). When homework is not home work: After-school programs for homework assistance. *Educational Psychologist*, *36*, 211–221.

Csikszentmihalyi, M. (1990). *Flow*. New York: Harper Row.

Csikszentmihalyi, M., & Larson, R. (1987). Validity and reliability of experience sampling method. *Journal of Nervous and Mental Disease*, *175*, 526–536.

Darling, N., Caldwell, L. L., & Smith, R. (2005). Participation in school-based extracurricular activities and adolescent adjustment. *Journal of Leisure Research*, *37*, 51–76.

Dishion, T. J., McCord, J., & Poulin, F. (1999). When interventions harm: Peer groups and problem behavior. *American Psychologist*, *54*, 755–764.

Dynarski, M., Moore, M., Mullens, J., Gleason, P., James-Burdumy, S., Rosenberg, L., *et al.* (2003). *When schools stay open late: The national evaluation of the 21st Century Community Learning Centers program*. Princeton, NJ: Mathematica Policy Research.

Eccles, J. S. (1999). The development of children ages 6 to 14. *The Future of Children*, *9*, 30–44.

Eccles, J. S. (in press). Extracurricular and other after school activities for youth. *Review of Research in Education*.

Eccles, J. S., & Barber, B. L. (1999). Student council, volunteering, basketball, or marching band: what kind of extracurricular involvement matters? *Journal of Adolescent Research*, *14*, 10–43.

Eccles, J. S., Midgley, C., Wigfield, A., Buchanan, C. M., Reuman, D., Flanagan, C., & Mac Iver, D. (1993). Development during adolescence: The impact of stage-environment fit on adolescents experiences in schools and families. *American Psychologist*, *48*, 90–101.

Eccles, J. S., Wigfield, A., & Schiefele, U. (1998). Motivation to succeed. In W. Damon (Series Ed.) & N. Eisenberg (Vol. Ed.), *Handbook of child psychology: Vol. 3. Social, emotional, and personality development* (5th ed., pp. 1017–1096). New York: Wiley.

Fashola, O. S. (2002). *Building effective afterschool programs*. Thousand Oaks, CA: Corwin.

Fletcher, A. C., Elder, G. H. Jr., & Mekos, D. (2000). Parental influences on adolescent involvement in community activities. *Journal of Research on Adolescence*, *10*, 29–48.

Fletcher, A. C., Nickerson, P., & Wright, K. L. (2003). Structured leisure activities in middle childhood: links to well-being. *Journal of Community Psychology*, *31*, 641–659.

Fletcher, A. C., & Shaw, R. A. (2000). Sex differences in associations between parent behaviors and characteristics and adolescent social integration. *Social Development*, *9*, 133–148.

Gambone, M. A., & Arbreton, A. J. A. (1997). *Safe havens: The contribution of youth organization to healthy adolescent development*. Philadelphia, PA: Public/Private Ventures.

Garner, R. (Ed.) (2002). *Hanging out: Community-based after-school programs for children*. Greenwood: Westport, CT.

Garner, R., Zhao, Y., & Gillingham, M. (2002). An alternative to self-care in a small Midwestern town. In R. Garner (Ed.), *Hanging out: Community-based after-school programs for children* (pp. 1–18). Westport, CT: Greenwood.

Grossman, J. B., Price, M. L., Fellerath, V., Jucovy, L. Z., Kotloff, L. J., Raley, R., *et al.* (2002). *Multiple choices after school: Findings from the extended-service schools initiative*. Philadelphia, PA: Public/Private Ventures.

Guest, A., & Schneider, B. (2003). Adolescents' extracurricular participation in context: The mediating effects of schools, communities, and identity. *Sociology of Education*, *76*, 89–109.

Halpern, R. (1999). After-school programs for low-income children: Promise and challenges. *The Future of Children*, *9*, 81–95.

Hansen, D. M., Larson, R. W., & Dworkin, J. B. (2003). What adolescents learn in organized youth activities: A survey of self-reported developmental experiences. *Journal of Research on Adolescence*, *13*, 25–55.

Hartup, W. W. (1996). The company they keep: Friendships and their developmental significance. *Child Development*, *67*, 1–13.

Hattie, J., Marsh, H. W., Neill, J. T., & Richards, G. E. (1997). Adventure education and outward bound: Out-of-class experiences that make a lasting difference. *Review of Educational Research*, *67*, 43–87.

Heath, S. B. (1999). Dimensions of language development: Lesson from older children. In A. S. Masten (Ed.), *Cultural processes in child development: The Minnesota symposium on child psychology* (Vol. 29, pp. 59–75). Mahwah, NJ: Erlbaum.

Hofferth, S., Brayfield, A., Deich, S., & Holcomb, P. (1991). *National childcare survey, 1990*. Report No. 91-5. Washington, DC: Urban Institute.

Huang, D., Gribbons, B., Kim, J. S., Lee, C., & Baker, E. L. (2000). *A decade of results: The impact of the LA's BEST after school enrichment program on subsequent student achievement and performance*. Los Angeles, CA: UCLA Center for the Study of Evaluation.

Hultsman, W. Z. (1992). Constraints to activity participation in early adolescence. *Journal of Early Adolescence*, *12*, 280–299.

Huston, A. C., & Wright, J. C. (1998). Mass media and children's development. In W. Damon (Series Ed.) & I. E. Sigel & K. A. Renninger (Vol. Eds.), *Handbook of child psychology: Vol. 4. Child psychology in practice* (5th ed., pp. 999–1058). New York: Wiley.

Johnson, C. D. (2002). *Daily life in colonial New England*. Westport, CT: Greenwood.

Jones, M. B., & Offord, D. R. (1989). Reduction of antisocial behavior in peer children by nonschool skill-development. *Journal of Child Psychology and Psychiatry*, *30*, 737–750.

Jordan, W. J., & Nettles, S. M. (2000). How students invest their time outside of school: Effects on school-related outcomes. *Social Psychology of Education*, *3*, 217–243.

Kiesner, J., Poulin, F., & Nicotra, E. (2003). Peer relations across contexts: Individual-network homophily and network inclusion in and after school. *Child Development*, *74*, 1328–1343.

Kleiber, D. A., & Powell, G. M. (2005). Historical change in leisure activities during after-school hours. In J. L. Mahoney, R. W. Larson, & J. S. Eccles (Eds.), *Organized activities as contexts of development: Extracurricular activities, after school and community programs* (pp. 23–44). Mahwah, NJ: Erlbaum.

Kleiner, B., Nolin, M. J., & Chapman, C. (2004). *Before- and after-school care, programs, and activities of children in kindergarten through eighth grade: 2001*. National Center for Education Statistics: US Department of Education. Institute of Education Sciences, NCES, 2004–2008.

Larson, R. (1989). Beeping children and adolescents: A method for studying time use and daily experience. *Journal of Youth and Adolescence*, *18*, 511–530.

Larson, R. W. (2000). Toward a psychology of positive youth development. *American Psychologist*, *55*, 170–183.

Larson, R., Dworkin, J., & Gillman, S. (2001). Facilitating adolescents' constructive use of time in one-parent families. *Applied Developmental Science*, *5*, 143–157.

Larson, R. W., & Verma, S. (1999). How children and adolescents spend time across the world: Work, play, and developmental opportunities. *Psychological Bulletin*, *125*, 701–736.

Lauver, S. C. (2002). *Assessing the benefits of an after-school program for urban youth: An impact and process evaluation*. Philadelphia, PA: Author.

Lerner, R. M. (2005). Promoting positive youth development through community and after-school programs. In J. L. Mahoney, R. W. Larson, & J. S. Eccles (Eds.), *Organized activities as contexts of development: Extracurricular activities, after school and community programs* (pp. ix–xii). Mahwah, NJ: Erlbaum.

Mahoney, J. L. (2000). School extracurricular activity participation as a moderator in the development of antisocial patterns. *Child Development*, *71*, 502–516.

Mahoney, J. L., Lord, H., & Carryl, E. (in press). An ecological analysis of after-school program participation and the development of academic performance and motivational attributes for disadvantaged children. *Child Development*.

Mahoney, J. L., & Cairns, R. B. (1997). Do extracurricular activities protect against early school dropout? *Developmental Psychology*, *33*, 241–253.

Mahoney, J. L., Cairns, B. D., & Farmer, T. W. (2003). Promoting interpersonal competence and educational success through extracurricular activity participation. *Journal of Educational Psychology*, *95*, 409–418.

Mahoney, J. L., Larson, R. W., Eccles, J. S., & Lord, H. (2005). Organized activities as contexts for child and adolescent development. In J. L. Mahoney, R. W. Larson, & J. S. Eccles (Eds.), *Organized activities as contexts of development: Extracurricular activities, after school and community programs* (pp. 3–22). Mahwah NJ: Erlbaum.

Mahoney, J. L., Schweder, A. E., & Stattin, H. (2002). Structured after-school activities as a moderator of depressed mood for adolescents with detached relations to their parents. *Journal of Community Psychology*, *30*, 69–86.

Mahoney, J. L., & Stattin, H. (2000). Leisure activities and adolescent antisocial behavior: The role of structure and social context. *Journal of Adolescence*, *23*, 113–127.

Mahoney, J. L., Stattin, H., & Magnusson, D. (2001). Youth recreation centre participation and criminal offending: A 20-year longitudinal study of Swedish boys. *International Journal of Behavioral Development*, *25*, 509–520.

Mahoney, J. L., & Zigler, E. F. (2003). *The national evaluation of the 21st-Century Community Learning Centers: A critical analysis of first-year findings*. New Haven, CT: Department of Psychology, Yale University, July 3.

Marsh, H. W. (1992). Extracurricular activities: Beneficial extension of the traditional curriculum or subversion of academic goals? *Journal of Educational Psychology*, *84*, 553–562.

Marshall, N. L., Coll, C. G., Marx, F., McCartney, K., Keefe, N., & Ruh, J. (1997). After-school time and children's behavioral adjustment. *Merrill-Palmer Quarterly*, *43*, 497–514.

McLaughlin, M. W. (2000). *Community counts: How youth organizations matter for youth development*. California: Public Education Network.

McLaughlin, M., Irby, M., & Langman, J. (1994). *Urban sanctuaries: Neighborhood organizations in the lives and futures of inner-city youth*. San Francisco: Jossey-Bass.

McNeal, R. B. Jr. (1995). Extracurricular activities and high school dropouts. *Sociology of Education*, *68*, 62–81.

McNeal, R. B. Jr. (1998). High school extracurricular activities: Closed structures and stratifying patterns of participation. *Journal of Educational Research*, *91*, 183–191.

Morris, P., & Kalil, A. (in press). Out of school time use during middle childhood in a low-income sample: Do combinations of activities affect achievement and behavior? In A. Huston & M. Riphe (Eds), *Middle childhood: Contexts of development*. New York: Cambridge University Press.

Morrison, G., Robertson, L., Harding, M., Weissglass, T., & Dondero, A. (2000). The protective function of after-school programming and parent education and support for students at risk for substance abuse. *Evaluation and Program Planning*, *23*, 365–371.

Mortimer, J. T., & Finch, M. D. (Eds.) (1996). *Adolescents, work, and family: An intergenerational developmental analysis*. Thousand Oaks, CA: Sage.

National Institute of Child Health and Human Development Early Child Care Research Network. (2004). Are child developmental outcomes related to before- and after-school care arrangements? Results from the NICHD study of early child care and youth development. *Child Development*, *75*, 280–295.

Noam, G. (2004). *After-school worlds: Creating a new social space for development and learning*. San Francisco: Jossey-Bass.

Osgood, D. W., Wilson, J. K., O'Malley, P. M., Bachman, J. G., & Johnston, L. D. (1996). Routine activities and individual deviant behavior. *American Sociological Review*, *61*, 635–655.

Parke, R., & Buriel, R. (1998). Socialization in the family: Ethnic and ecological perspectives. In W. Damon (Series Ed.) & N. Eisenberg (Vol. Ed.), *Handbook of child psychology: Vol. 3. Social, emotional, and personality development* (5th ed., pp. 463–552). New York: Wiley.

Pedersen, S., & Seidman, E. (2005). Contexts and correlates of out-of-school activity participation among low-income urban adolescents. In J. L. Mahoney, R. W. Larson, & J. S. Eccles (Eds.), *Organized activities as contexts of development: Extracurricular activities, after school and community programs* (pp. 85–110). Mahwah, NJ: Erlbaum.

Pettit, G. S., Laird, R. D., Bates, J. E., & Dodge, K. A. (1997). Patterns of after-school care in middle childhood: Risk factors and developmental outcomes. *Merrill-Palmer Quarterly*, *43*, 515–538.

Pianta, R. C. (1999). *Enhancing relationships between children and teachers*. Washington, DC: American Psychological Association.
Pierce, K. M., Hamm, J. V., & Vandell, D. L. (1999). Experiences in after-school programs and children's adjustment in first grade classrooms. *Child Development*, *70*, 756–767.
Posner, J. K., & Vandell, D. L. (1994). Low-income children's after-school care: Are there beneficial effects of after-school programs? *Child Development*, *65*, 440–456.
Posner, J. K., & Vandell, D. L. (1999). After-school activities and the development of low-income urban children: A longitudinal study. *Developmental Psychology*, *35*, 868–879.
Ross, J. G., Saavedra, P. J., Shur, G. H., Winters, F., & Felner, R. D. (1992). The effectiveness of an after-school program for primary grade latchkey students on precursors of substance abuse. *Journal of Community Psychology*, 22–38, Special issue: Programs for change: Office for Substance Abuse Prevention demonstration models.
Rubin, K. H., Bukowski, W. & Parker, J. G. (1998). Peer interactions, relationships and groups. In N. Eisenberg (Series Ed.) & W. Damon (Vol. Ed.), *Handbook of child psychology: Vol. 3. Social, emotional and personality development* (5th ed., pp. 619–700). New York: Wiley.
Sandstrom, M. J., & Coie, J. D. (1999). A developmental perspective on peer rejection: Mechanisms of stability and change. *Child Development*, *70*, 955–966.
Schreiber, J. B., & Chambers, E. A. (2002). After-school pursuits, ethnicity, and achievement for 8th- and 10th-grade students. *The Journal of Educational Research*, *96*, 90–100.
Seppanen, P., Love, J., deVries, D., Berstein, L., Seligson, M., Marx, F., *et al.* (1993). *National study of before- and after-school programs. Final report*. Washington, DC: US Department of Education, Office of Policy and Planning.
Smith, K. (2000). *Who's minding the kids? Child care arrangements: Fall 1995. Current Population Reports, P70-70*. Washington, DC: US Census Bureau.
Smith, K. (2002). *Who's minding the kids? Child care arrangements: Spring 1997. Current Population Reports, P70-86*. Washington, DC: US Census Bureau.
Snyder, H., & Sickmund, M. (1999). *Juvenile offenders and victims: 1999 national report*. Washington, DC: US Department of Justice, Office of Juvenile Justice and Delinquency Prevention.
Subrahmanyam, K., Kraut, R. E., Greenfield, P. M., & Gross, E. F. (2000). The impact of home computer use on children's activities and development. *The Future of Children: Children and Computer Technology*, *10*, 123–144.
Tout, K., Scarpa, J., & Zaslow, M. J. (2002, March). *Children of current and former welfare recipients: Similarly at risk*. Washington, DC: Child Trends.
Vandell, D. L. (2003). Playing by the rules: The 21st Century evaluation failed to meet established standards for research. *The evaluation exchange*.
Vandell, D. L., & Pierce, K. M. (2001, April). Experiences in after-school programs and child well-being. In J. L. Mahoney (Chair), *Protective aspects of after-school activities: Processes and mechanisms.* Paper Symposium conducted at the biennial meeting of the Society for Research in Child Development, Minneapolis, MN.
Vandell, D. L., Pierce, K. M., & Lee, D. (2005, April). Relationships with after-school program staff: Supporting (versus hindering) children's adjustment and achievement. In J. L. Mahoney & D. L. Vandell (Chairs), *Features of after-school programs that promote development: Type, quality and content*. Symposium conducted at the biennial meeting of the Society for Research in Child Development, Atlanta, GA.
Vandell, D. L., & Posner, J. (1999). Conceptualization and measurement of children's after-school environments. In S. L. Friedman & T. D. Wachs (Eds.), *Measuring environment across the life span: Emerging methods and concepts* (pp. 167–196). Washington, DC: American Psychological Association Press.

Vandell, D. L., Reisner, E. R., Brown, B. B., Pierce, K. M., Dadisman, K., & Pechman, E. M. (2004, February). *The study of promising after-school programs: Descriptive report of the promising programs*. Madison, WI: Wisconsin Center for Education Research, University of Wisconsin-Madison.

Vandell, D. L., Shernoff, D. J., Pierce, K. M., Bolt, D. M., Dadisman, K., & Brown, B. B. (in press). Activities, engagement, and emotion at after-school programs (and elsewhere). In H. B. Weiss, P. M. Little, & S. M. Bouffard (Eds.), New directions for youth development: Participation in youth programs: Enrollment, attendance, and engagement (No. 105).

Vandell, D. L., & Shumow, L. (1999). After-school child care programs. *Future of Children*, *9*, 84–90.

Vandell, D. L., Shumow, L., & Posner, J. (2005). After-school programs for low-income children: Differences in program quality. In J. L. Mahoney, R. W. Larson & J. S. Eccles (Eds.), *Organized activities as contexts of development: Extracurricular activities, after school and community programs* (pp. 437–456). Mahwah, NJ: Erlbaum.

Welsh, M. E., Russell, C. A., Williams, I., Reisner, E. R., & White, R. N. (2002). *Promoting learning and school attendance through after-school programs: Student level changes in educational performance across TASC's first three years*. Washington, DC: Policy Studies Associates.

Youness, E., Gillingham, M., Nowatzki, D., & Roche, E. (2002). Working at play in four Chicago computer clubs. In R. Garner (Ed.), *Hanging out: Community-based after-school programs for children* (pp. 113–136). Westport, CT: Greenwood.

Zill, N., Nord, C. W., & Loomis, L. S. (1995). *Adolescent time use, risky behavior, and outcomes: An analysis of national data*. Rockville, MD: Westat.

This Page Intentionally Left Blank

MECHANISMS OF CHANGE IN THE DEVELOPMENT OF MATHEMATICAL REASONING

Martha W. Alibali

DEPARTMENT OF PSYCHOLOGY, UNIVERSITY OF WISCONSIN-MADISON,
MADISON, WI 53706, USA

I. Introduction

How do children progress from less knowledge to more knowledge in a domain? Understanding this transition is at the heart of understanding all forms of knowledge change, including changes due to development, learning, and instruction. To deeply understand change, we must be able not only to characterize the shape of change, but also to explain how change occurs. In other words, we must be able to specify the mechanisms that give rise to change. Detailed knowledge about the mechanisms of change will allow us to predict who will change and when, enable us to set up conditions to foster or inhibit change, and guide us about how to intervene when change does not occur as desired.

Advances in Child Development and Behavior
R Kail (Editor)

Note that I refer here to *mechanisms* of change, in the plural rather than the singular. It is virtually certain that multiple change mechanisms apply in every domain. Different mechanisms may be responsible for changes observed in different individuals, at different times, and with respect to different specific content. The particular mechanisms at work in any particular individual's development depend on many factors, including characteristics of that individual and the individual's opportunities for learning, both formal and informal.

In this chapter, I focus on change mechanisms that apply in the development of mathematical reasoning. Mathematical reasoning is used in a variety of situations, including both formal mathematical tasks and everyday activities, such as shopping, managing finances, and cooking (e.g., doubling recipes, estimating quantities), so it is an important type of reasoning to understand in its own right. In addition, mathematical reasoning is a useful domain within which to study change mechanisms, because many changes in people's mathematical thinking and problem solving can be readily observed, and because such changes are amenable to experimental investigation. Furthermore, understanding change in the domain of mathematical reasoning may prove useful for designing optimal instruction and for understanding mathematics learning difficulties.

I focus largely on mathematical reasoning in the elementary and middle school years, although the mechanisms discussed are neither limited to these periods, nor limited to the domain of mathematics *per se*. Indeed, the same mechanisms may apply to other types of content (e.g., scientific reasoning, statistical reasoning) and with respect to earlier achievements, such as the development of counting, and later achievements, such as the acquisition of key concepts in calculus. However, most of the mechanisms discussed in this chapter have been studied most intensively in school-age children.

The purpose of this chapter is to review research regarding two classes of change mechanisms that have been proposed to apply in the domain of mathematical reasoning: (1) mechanisms that involve reciprocal relations between knowledge of problem-solving procedures and knowledge of concepts, and (2) mechanisms that involve expressing knowledge in speech and gestures. Of course, these two classes of mechanisms are not the only ones that have been proposed to apply in the development of mathematical reasoning. The present review is intended to be focused rather than exhaustive.

Mechanisms of change can be characterized at different grain sizes, and they may operate at different levels, ranging from the behavioral level (e.g., self-explaining) to the neural level (e.g., strengthening or weakening of neural connections). The mechanisms that are the focus of the current chapter are characterized at the behavioral level. For each class of mechanisms, I begin with global claims and proceed to more detailed hypotheses about the functioning of specific mechanisms.

II. Mechanisms that Involve Reciprocal Relations between Knowledge of Problem-Solving Procedures and Knowledge of Concepts

Mathematical knowledge consists of several different types of knowledge (e.g., Hiebert & LeFevre, 1986; Bisanz & LeFevre, 1990). Among these are knowledge of concepts (e.g., the principle that the two sides of an equation represent the same quantity) and knowledge of procedures for solving problems (e.g., procedures for isolating variables in algebraic equations). As many researchers have acknowledged, defining conceptual and procedural knowledge precisely is difficult, as is distinguishing them in practice. For present purposes, I adopt the definitions used by Rittle-Johnson and Alibali (1999). *Conceptual knowledge* involves understanding of principles that govern a domain and of relations among pieces of knowledge within a domain. *Procedural knowledge* involves knowledge of problem-solving procedures, or action sequences for solving problems.

Past research suggests that conceptual and procedural knowledge are intertwined in multiple and complex ways (e.g., Carpenter, 1986; Rittle-Johnson, Siegler, & Alibali, 2001; Canobi, Reeve, & Pattison, 2003). Indeed, gains in each type of knowledge can lead to gains in the other (Rittle-Johnson & Alibali, 1999). Thus, one class of mechanisms of change in mathematical reasoning involves reciprocal relations between knowledge of concepts and knowledge of problem-solving procedures. In the following subsections, I review evidence about how knowledge of concepts and knowledge of procedures influence one another. In each subsection, I first consider evidence that one type of knowledge may influence the other, and then consider potential mechanisms that may underlie that influence.

A. CONCEPTS INFORM PROCEDURE GENERATION

Several sources of evidence converge to suggest that gains in conceptual knowledge can influence procedural knowledge. First, a number of studies in various mathematical domains have shown that instruction that focuses on conceptual principles leads students to generate new problem-solving procedures. These include studies of decimal fractions (e.g., Hiebert & Wearne, 1989), multi-digit arithmetic (e.g., Fuson & Briars, 1990; Hiebert & Wearne, 1996; Blöte, Van der Burg, & Klein, 2001), and mathematical equivalence (e.g., Perry, 1991; Alibali, 1999). As one example, Rittle-Johnson and Alibali (1999) examined changes in the procedures that fourth- and fifth-grade students used to solve mathematical equivalence problems, which are equations with addends on both sides of the equal sign, such as $3 + 4 + 6 = 3 + \_$. Some of the students received a brief lesson focusing on the principle that the two sides of an equation represent the same quantity. Most students who received this lesson generated correct procedures for solving the problems after instruction, including procedures such as “find the sum of the left side, and then find a number that makes the sum on the right side the same as that on the left” (the *equalize* procedure) and “cancel the number that is the same on both sides, and add the remaining numbers” (the *cancel-and-group* procedure).

Second, some studies have shown that children with greater conceptual knowledge display greater gains in procedural knowledge after instruction. For example, Rittle-Johnson, Siegler, and Alibali (2001) assessed children’s conceptual understanding of decimal fractions before and after an intervention that included a brief lesson. They also assessed children’s procedural skill at placing decimal fractions on the number line before, during, and after the intervention. Children who had higher scores on the conceptual knowledge pretest made greater improvements in procedural knowledge from the pretest to the later segments of the study. Thus, amount of initial conceptual knowledge was associated with size of gains in procedural knowledge.

These studies raise the question of *how* conceptual knowledge leads to changes in procedural skill. By what processes might conceptual knowledge inform procedure generation? In considering this issue, Baroody (2003) distinguished between direct and indirect effects of conceptual knowledge on procedural advances. The direct pathway involves *conceptual instigation* of procedural innovations, such that innovations are directly motivated or triggered by new conceptual knowledge. The indirect pathway involves *conceptual support* for procedural advances that are motivated by other, non-conceptual factors. A wide variety of non-conceptual factors may compel children to change their problem-solving procedures, including a desire for cognitive economy, feedback about solution correctness, and outside intervention (e.g., seeing an adult or another child model a different procedure). The distinction between conceptual instigation and conceptual support is useful to bear in mind when considering the mechanisms that may underlie the effects of conceptual knowledge on procedural knowledge.

Conceptual knowledge may influence improvements in procedural knowledge through a number of possible mechanisms, either by directly instigating procedural innovations, or by supporting procedural advances driven by other, non-conceptual factors. In the following sections, I consider three potential mechanisms, and review existing empirical evidence for each: (1) gains in conceptual knowledge may lead to changes in problem representation, which in turn may enable the generation of new procedures; (2) gains in conceptual knowledge may lead children to realize that existing procedures are incorrect, and this may trigger procedure generation; and (3) gains in conceptual knowledge may guide children's evaluation of potential, alternative procedures, whether those procedures are self-generated or learned from outside sources.

1. Changes in Problem Representation

One direct pathway by which gains in conceptual knowledge may instigate procedural innovations is by causing changes in problem representation, which in turn may enable procedural innovations. Problem representation can be defined as "the internal depiction or recreation of a problem in working memory during problem solving" (Rittle-Johnson, Siegler & Alibali, 2001, p. 348). A new problem representation is created each time a problem is solved. Problem representations are sometimes inaccurate or incomplete because solvers may fail to represent certain problem features or may represent them inaccurately.

Past research from a variety of domains supports each of the links in the pathway from gains in conceptual knowledge to improved problem representation to gains in procedural knowledge. Problem solvers who have greater conceptual knowledge in a domain tend to form more accurate and more complete problem representations (e.g., Chi, Feltovich, & Glaser, 1981, physics problem solving), and experimental manipulations designed to improve solvers' problem representations lead to generation of new problem-solving procedures (e.g., Siegler, 1976, balance scale problem solving). Within mathematics, evidence from two domains, decimal fractions and mathematical equivalence, supports both of the links in this hypothesized pathway.

First, consider the evidence from the domain of decimal fractions. As described previously, Rittle-Johnson, Siegler, and Alibali (2001) assessed children's procedural skill at placing decimal fractions on a number line on a pretest, during an instructional intervention, and on a posttest. They also assessed children's conceptual understanding of decimal fractions at pretest and posttest. Rittle-Johnson *et al.* found that pretest to posttest gains in children's conceptual knowledge of decimal fractions were associated with improvements in their abilities to correctly place a decimal fraction on a number line.

Crucially, this relation was accounted for, in part, by children's representations of the decimal fractions, which were assessed based on explanations that the children provided during the intervention. On each of the 12 intervention trials, children solved a problem (e.g., indicated the location on a number line corresponding to a given decimal fraction), were told the correct answer, and then were asked to explain why that correct answer was correct. For each explanation, children were scored as having a correct representation if they conveyed either a *common-unit* understanding of the fraction, in which the fraction was represented in terms of its smallest unit (e.g., 0.625 is represented as 625 thousandths), or a *composite* understanding of the fraction, in which the fraction is represented as the sum of the individual column values (e.g., 0.625 is represented as the sum of 6 tenths, 2 hundredths, and 5 thousandths). One common incorrect representation of the fraction stemmed from an analogy to whole numbers (e.g., 0.625 is represented as "in the six hundreds").

Using mediation analyses, Rittle-Johnson *et al.* demonstrated that (1) conceptual knowledge at pretest was associated with frequency of correct problem representation during the intervention, (2) frequency of correct problem representation during the intervention was associated with gains in procedural knowledge from pretest to posttest and transfer test, and (3) the relation between initial conceptual knowledge and gains in procedural knowledge was diminished (although not completely eliminated) when frequency of correct problem representation was included in the regression model. Thus, problem representation accounted for at least part of the relation between initial conceptual knowledge and procedural knowledge gain. Children with greater conceptual knowledge were more likely to accurately represent the decimal fractions, and children who had better representations of the fractions displayed better learning of how to place the fractions on the number line.

The components of this pathway have also received empirical support in the domain of mathematical equivalence. Several studies have shown that instructional interventions that are designed to inculcate conceptual knowledge of mathematical equivalence also lead to improvements in problem representation. In one study (McNeil & Alibali, in press-b), children in a conceptual-instruction group received a brief intervention about the meaning of the equal sign, and they were also shown a correctly solved equivalence problem. Children in a control group were shown a correctly solved equivalence problem but received no other intervention. Children's problem representations were assessed both before and after the interventions by asking them to reconstruct equivalence problems that they viewed for 5 seconds each. As seen in Figure 1, among children who represented equivalence problems poorly on the pretest, children in the conceptual intervention condition made greater improvements in problem representation than did children in the control condition. In another study, Rittle-Johnson and Alibali (1999) provided children with a lesson that focused on the goal of making both sides of the problem equal, but that did not provide any guidance about how to achieve this goal. Among children who represented equivalence problems incorrectly on a pretest assessment, 58% of the children who received the conceptual lesson represented the problems correctly on a posttest assessment, compared to only 30% of children in a control condition who did not receive instruction. Thus, manipulations designed to promote conceptual knowledge also lead to gains in problem representation.

Evidence from studies of mathematical equivalence has also shown that improvements in problem representation can lead to generation of new procedures for solving equivalence problems. Alibali, McNeil, and Perrott (1998) presented some children with equivalence problems in which the equal sign was printed in red, and also told them to "notice where the equal sign

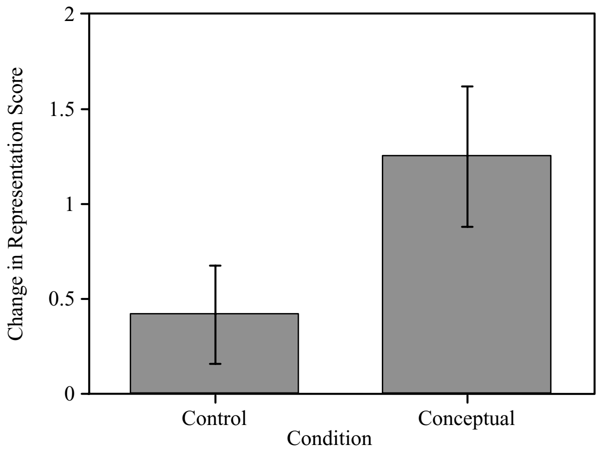

Fig. 1. Mean pretest to posttest improvements in representation scores for children who received a brief conceptual lesson and children who did not. Children who performed well on the pretest (i.e., who reconstructed at least two of three problems correctly) were excluded from the analysis. Children in the conceptual condition ($N = 12$) made greater improvements in problem representation than did children in the control condition ($N = 12$), $t(22) = 1.84$, $p < 0.05$. Data drawn from the study reported in McNeil and Alibali (in press-b).

is in the problem." This intervention was intended to highlight the position of the equal sign, which children often fail to represent correctly (McNeil & Alibali, in press-b). Children who received this intervention were more likely than children who did not to generate new procedures for solving the problems on a posttest. However, children expressed most of their new procedures in gestures and not in speech on the posttest. These new gestured procedures tended to occur in responses in which children expressed one (old) procedure in speech and a different (new) procedure in the accompanying gesture.[1] For example, as seen in Table I (third row), a child might express the *add-all* procedure in speech, but at the same time, express the *equalize* procedure (i.e., make both sides equal) in gesture. Responses in which speech conveys one procedure and gesture conveys another procedure have been termed "gesture–speech mismatches" in prior work (Perry, Church, & Goldin-Meadow, 1988). Furthermore, the procedures expressed uniquely in gesture in such mismatches have been shown in other research to reflect implicit, emerging procedures (e.g., Garber, Alibali, & Goldin-Meadow, 1998). Thus, these data suggest that changes in representation helped "get the ball rolling" in the process of generating new procedures.

[1]For details about how the system for coding gestures was developed, see Perry, Church, and Goldin-Meadow (1988).

TABLE I

Sample Gesture–Speech Match and Mismatch Responses for Mathematical Equivalence Problems

	Verbal explanation	Gestured explanation
	Gesture–speech matches	
1	I added 3 plus 9 plus 5 plus 3, and I got 20. (*add-all*)	Right hand point: left 3, 9, 5, right 3, solution. (*add-all*)
2	3 plus 9 plus 5 is 17, and 3 plus 14 is 17. (*equalize*)	Left hand point: sweep under $3 + 9 + 5$. Right hand point: sweep under $3 + \_$. (*equalize*)
	Gesture–speech mismatches	
3	3 plus 9 is 12, plus 5 is 17, plus 3 makes 20. (*add-all*)	Left hand point: sweep under $3 + 9$, $9 + 5$. Right hand point: sweep under $3 + \_$. (*equalize*)
4	I added 3 and 9 and 5 and got 17, and 3 plus 14 is 17. (*equalize*)	Right hand point: 9, 5, solution. (*group*)

All examples are for the problem $3 + 9 + 5 = 3 + \_$. Procedure codes are indicated in parentheses.

Note: Procedures in gesture and speech are coded independently, using systems developed by Perry, Church, and Goldin-Meadow (1988). For each response, the codes assigned to the verbal and gestured explanations are compared. If the same procedure is assigned to the verbal and gestured explanations, the response as a whole is considered a gesture–speech match. If different procedures are assigned to the verbal and gestured explanations, the response as a whole is considered a gesture–speech mismatch.

In a follow-up study (Alibali & McNeil, in preparation), children were again presented with equivalence problems in which the equal sign was printed in red, and they also received feedback that the procedures they had used to solve the pretest problems were incorrect. Compared to the initial study in which children did not receive such feedback, many more children in this study generated new procedures that they expressed in speech. In the red-equal-sign condition, nearly half of the children expressed new procedures in speech on the posttest. Moreover, most of the new procedures that children generated involved the equal sign in some way. The most commonly generated procedure was the *equalize* procedure (a correct procedure), and the second most commonly generated procedure was the *add-to-equal-sign* procedure (an incorrect procedure), both of which rely on representing the position of the equal sign.

Taken together, these two studies suggest that changes in representation are a key component of the process of procedure generation. When children notice new features of problems, but do not realize that their existing procedures are incorrect, they generate new procedures and express them in gesture.

When children receive feedback that their existing procedures are incorrect, they generate new procedures and express them in speech. Thus, changes in problem representation promote procedure generation. When children begin to represent new problem features, they become able to generate new procedures that depend on those features.

From this perspective, one can think more concretely about the processes by which gains in conceptual understanding lead to generation of new procedures. Imagine a fourth-grade girl who lacks a sophisticated concept of the equal sign, and who instead thinks that the equal sign means "put the answer." Many late elementary and middle school students hold this view (Kieran, 1981; Rittle-Johnson & Alibali, 1999; McNeil & Alibali, in press-a). When presented with an equivalence problem such as $3 + 4 + 5 = 3 + \_$, the girl is likely to ignore the position of the equal sign, and instead represent the problem as a typical addition problem, such as $3 + 4 + 5 + 3 = \_$ (McNeil & Alibali, 2004). She is also likely to solve the equivalence problem incorrectly, using a procedure such as *add-all*, and arriving at a solution of 15.

Suppose this girl then learns via instruction that the equal sign does not mean "put the answer" but instead means that the quantities on each side of it are the same. If she is then presented with an equivalence problem, she might begin to notice the equal sign in the problem and, for the first time, begin to represent the two "sides" of the problem. Once the "sides" of the problem are represented as problem features, a whole new set of procedures become possible. The girl may then generate the *equalize* procedure, which involves making both sides equal.

2. Realizing that Existing Procedures Are Incorrect

Another direct pathway by which gains in conceptual knowledge may instigate procedural innovations is via the realization that existing procedures are incorrect. If acquiring new conceptual knowledge leads children to realize that their existing procedures are faulty, they may be compelled to construct new procedures to "fill the gap." In this case, conceptual knowledge may both instigate construction of a new procedure and support children's efforts to construct a new procedure that is correct.

An example may clarify the processes involved in this pathway. Consider the girl in the previous example, who learned that the equal sign means that the quantities on either side of it are the same. The girl may attempt to link this new conceptual knowledge with her existing procedure for solving equivalence problems, namely, the *add-all* procedure. However, in doing so, she may realize that her existing procedure is incompatible with this meaning of the equal sign. This may lead her to conclude that her existing procedure is incorrect, and compel her to generate a new one.

In constructing a new procedure, the girl may also utilize her newly acquired conceptual knowledge to guide her representation of the problem, and she may begin to represent features of the problem that she had heretofore represented incorrectly or not represented at all, such as the position of the equal sign. Thus, realizing that her existing procedure is incorrect may compel her to change her representation of the problem, engaging the mechanism discussed previously.

Although this pathway from new conceptual knowledge to realizing an existing procedure is incorrect to generating a new procedure is a plausible one, to my knowledge, there is no empirical data to support the entire pathway. However, there is ample evidence for the second link in the pathway, namely that children change their procedures when they learn that their existing approaches are incorrect. In one of the experimental conditions in Alibali (1999), children were asked to solve equivalence problems in an intervention, and for problems that were solved incorrectly, children were told, "That's a good try, but it's not the right way to solve the problem." Thirty-one percent of the children in this condition generated new procedures for solving the problem on a posttest, compared to no children in a control condition in which children did not receive feedback about the correctness of their prior approach.

Some data can also be marshaled to support the first link in the pathway, namely, the link from gains in conceptual knowledge to realizing an existing procedure is incorrect. In Rittle-Johnson and Alibali's (1999) study of mathematical equivalence, children evaluated three correct and three incorrect procedures both before and after a brief conceptual lesson, or without instruction in a control group. All of the children used incorrect procedures at pretest. The data of interest in the present context are children's evaluations of the *add-all* procedure, the most common incorrect procedure for problems of the form used in the study ($a + b + c = a + \_$). Children rated the procedures on a scale from "not so smart" (0) to "very smart" (2). Ratings of the *add-all* procedure given by children in the conceptual-instruction group decreased substantially (from 1.32 to 0.53), whereas ratings of children in the no-instruction control group held steady (from 1.22 to 1.28). These data suggest that the conceptual instruction led children to re-evaluate the *add-all* procedure, which the majority of the children used at pretest, and realize that this procedure was not correct.

Although these data are suggestive, it would be more definitive to directly assess children's acquisition of the target concepts, and to examine the relation between acquisition of the concepts and evaluation of pretest procedures. If this pathway works as hypothesized, children who acquire the target concepts should evaluate their incorrect pretest procedures more negatively at posttest than children who do not acquire the target concepts. A reanalysis of data from Rittle-Johnson and Alibali (1999) supports this hypothesis. Children were classified as having gained conceptual knowledge if they scored higher on the posttest conceptual knowledge assessment than on the pretest conceptual knowledge assessment. As seen in Figure 2, children who gained conceptual

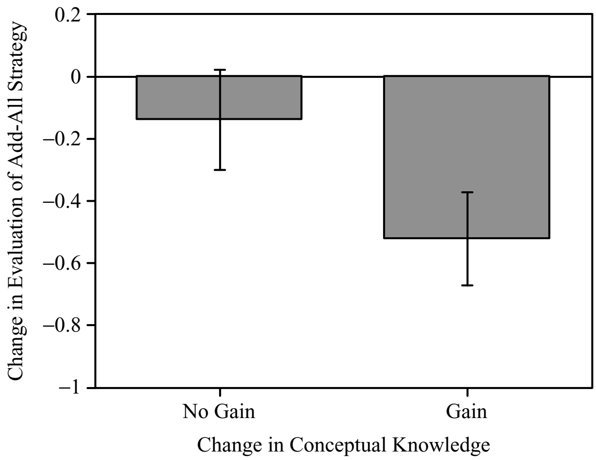

Fig. 2. Mean pretest to posttest changes in children's evaluations of the add-all *procedure for children who made gains in conceptual knowledge from pretest to posttest and children who did not make gains, collapsing across experimental conditions. Children who made gains in conceptual knowledge decreased their evaluations of the* add-all *procedure more than did children who did not make gains in conceptual knowledge, $t(46) = 1.71$, $p < 0.05$, one-tailed. Data drawn from Rittle-Johnson and Alibali (1999).*

knowledge over the course of the study decreased their evaluations of the *add-all* procedure more than did children who did not gain conceptual knowledge.

It would also be of value to address the entire pathway in a single study—ideally a microgenetic study that would examine conceptual knowledge, procedure evaluation, and procedure use at multiple time points. Such a study could address the temporal relations between gains in conceptual knowledge and changes in procedure evaluation.

3. *Guiding Evaluations of Alternative Procedures*

A third pathway by which conceptual knowledge may influence procedural knowledge is that gains in conceptual knowledge may guide children's evaluation of potential alternative procedures. Children may invent new procedures themselves, or they may learn new procedures from a variety of outside sources, such as observing other children, reading textbooks, learning from instruction, and so forth. Conceptual knowledge may inform children's decisions about whether to adopt these procedures, and may guide their evaluation of procedures invented to meet the demands of the current problem. Thus, in this pathway, conceptual knowledge is a support for procedural advances that are motivated by other, non-conceptual factors, such as modeling, instruction, or a desire for efficiency.

This pathway was first described by Siegler and Crowley (1994), who argued that children use conceptual knowledge, in the form of "goal sketches," to evaluate alternative procedures for solving problems. Goal sketches incorporate information about the goals that procedures within a domain must meet. In an experiment on children's procedures for solving simple addition problems (e.g., $4 + 8 = ?$), Siegler and Crowley found that 5-year-olds who did not yet spontaneously use the *count-from-larger-addend* procedure judged it to be smarter than an equally novel illegitimate procedure, and just as smart as their typical procedure of counting from one. The illegitimate procedure involved counting one of the addends twice, so it was inconsistent with the hypothesized goal sketch for simple addition. Thus, conceptual knowledge, in the form of a goal sketch, appeared to guide children's evaluation of the novel procedures.

Other findings support the idea that gains in conceptual knowledge influence children's evaluations of alternative procedures. As noted earlier, in Rittle-Johnson and Alibali's (1999) study of mathematical equivalence, children evaluated a set of six different procedures on a three-point scale both before and after a conceptual lesson, or without instruction in a control group. As seen in Figure 3, children in the conceptual-instruction group altered their evaluations in appropriate ways, decreasing their evaluations of incorrect procedures ($M = -0.32$) and increasing their evaluations of correct procedures ($M = +0.20$). In contrast, children in the control group did not change their evaluations of incorrect procedures ($M = +0.01$) and actually decreased their evaluations of correct procedures ($M = -0.20$). These data suggest that gains

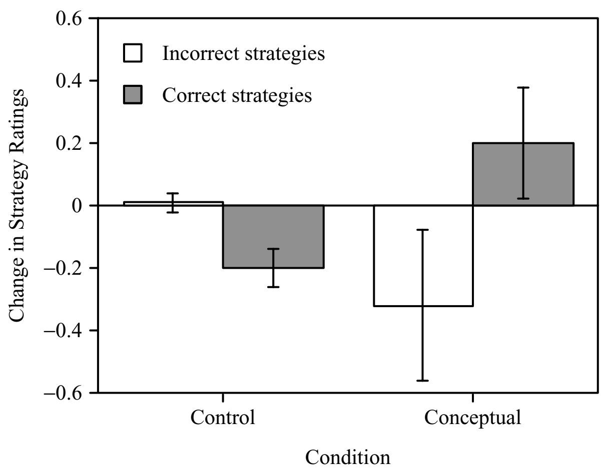

Fig. 3. Mean pretest to posttest change in children's evaluations of correct procedures (N = 3) and incorrect procedures (N = 3) for children in the conceptual-instruction group and the control group. Data drawn from Rittle-Johnson and Alibali (1999).

in conceptual knowledge due to the lesson helped children in the conceptual-instruction group to evaluate potential alternative procedures in sensible ways.

Changes in children's evaluations of alternative procedures presumably inform children's procedure choices. Contemporary models of procedure choice in problem solving, such as that proposed by Shrager and Siegler (1998) (see also Siegler & Shipley, 1995), hold that learners store information about the effectiveness and efficiency of various potential procedures and use this information to guide their procedure choices. Within such a framework, improvements in children's evaluations of alternative procedures, such as increases in evaluations of correct procedures and decreases in evaluations of incorrect procedures, are likely to lead to better choices among competing procedures.

4. Summary

Existing empirical data strongly support the possibility that gains in conceptual knowledge lead to changes in problem representation, which in turn enable the generation of new procedures. Less empirical data support the possibilities that (1) gains in conceptual knowledge lead children to realize that existing procedures are incorrect, and this impels them to generate new procedures, and (2) gains in conceptual knowledge guide children's evaluation of alternative procedures, both self-generated ones and ones learned from outside sources. However, the available data suggest that these are also viable pathways to account for the impact of gains in conceptual knowledge on procedural knowledge.

B. PROCEDURES PROVIDE THE BASIS FOR INFERRING CONCEPTS

Mounting evidence indicates that gains in procedural knowledge can influence conceptual knowledge. Several sources of evidence converge on this conclusion. First, several studies have shown that children demonstrate gains in conceptual knowledge after a procedural lesson. For example, Rittle-Johnson and Alibali (1999) provided third- and fourth-grade students with instruction about a correct procedure for solving equivalence problems (cancel like addends and group the remaining addends). Children's conceptual understanding of the equal sign symbol was assessed both before and after the lesson, using a battery of tasks designed to assess understanding both explicitly (e.g., tell what the equal sign means) and implicitly (e.g., rate other children's definitions of the equal sign, evaluate non-standard equations as making sense or not). Children who received the procedural lesson made greater gains in conceptual knowledge than did children in a control group who did not receive any lesson. At the individual level, 53% of children in the procedural instruction group improved their conceptual understanding from pretest to posttest, whereas only 38% of children in the control group did so.

Second, some evidence indicates that improvements in procedural knowledge are associated with gains in conceptual knowledge. As described previously, Rittle-Johnson, Siegler, and Alibali (2001) assessed children's procedural skill at placing decimal fractions on a number line on a pretest, during an instructional intervention, and on a posttest. They also assessed children's conceptual understanding of decimal fractions at pretest and posttest. In a regression analysis, they controlled for scores on the procedural knowledge pretest, and determined that procedural knowledge scores during the intervention and at posttest predicted improvements from pretest to posttest in children's conceptual knowledge. Procedural knowledge scores at both time points were positively associated with size of conceptual knowledge gains. Thus, learning to correctly place fractions on the number line was linked with improvements in children' conceptual knowledge of decimal fractions.

However, despite the evidence that gains in procedural knowledge *can* lead to improvements in conceptual knowledge, some evidence suggests that this causal pathway is less strong or less consistent than the reverse pathway (that from gains in conceptual knowledge to improvements in procedural knowledge). For example, Rittle-Johnson and Alibali (1999) found that more than 80% of children who received conceptual instruction generated correct procedures for solving problems at posttest, whereas only 53% of children who received procedural instruction displayed pretest to posttest gains in conceptual knowledge. Of course, one cannot be certain that the conceptual and procedural knowledge assessments were equally sensitive. However, the pattern of results suggests that the strength of influence of each type of knowledge on the other may be asymmetrical. In addition, many children in the procedural-instruction group were unable to adapt their newly learned procedures to solve transfer problems, suggesting that they had not acquired sufficient conceptual knowledge to guide their adaptation of the instructed procedure to novel contexts. Thus, although children did acquire conceptual knowledge from a procedural lesson, gains were modest and did not hold for all children.

Other evidence also suggests that the influence of gains in procedural knowledge on conceptual knowledge may be limited. In some domains (e.g., multi-digit subtraction, fraction multiplication, fraction division), people learn correct procedures but never fully understand the conceptual underpinnings of those procedures (Fuson, 1990; Ma, 1999). Furthermore, Byrnes and Wasik (1991) provided children a lesson on the least common denominator procedure for fraction addition, but did not observe gains in conceptual knowledge following the lesson. There are several potential explanations for this null result; however, one possibility is that gains in procedural knowledge do not promote gains in conceptual knowledge in all cases or for all children.

This work suggests that any purported mechanisms by which procedural knowledge may lead to gains in conceptual knowledge must be able to account for variability across children in whether or not the gains occur. In the following sections, I consider two potential mechanisms, and review existing empirical evidence for each: (1) learning new procedures may cause children to change their problem representations, and improved problem representation may in turn provide a basis for inferring concepts; and (2) learning new procedures may lead children to reflect upon why those procedures "work" or are effective, and this may provide a basis for inferring concepts.

1. Changes in Problem Representation

One possible mechanism by which gains in procedural knowledge may promote gains in conceptual knowledge is via changes in problem representation. Learning new procedures may cause children to improve their problem representations, and improved problem representation may in turn provide a basis for inferring concepts.

A hypothetical example may serve to clarify the processes involved in this pathway. Consider a fourth-grade boy who holds an operational concept of the equal sign, namely that the equal sign means "put the answer." Suppose that this boy learns that one correct way to solve an equivalence problem such as $3 + 9 + 5 = 3 + \_$ is to cancel the addend that appears on both sides of the equal sign (in this case, the 3), and add the remaining numbers (in this case, the 9 and 5). After learning this procedure, the boy may be more likely to notice the position of the equal sign in other equations that he encounters, because it is essential to attend to the equal sign in order to determine what the "sides" of the problem are. Thus, learning this new procedure may lead him to represent equivalence problems more accurately than he did before. In particular, he may more accurately represent the position of the equal sign.

If the boy begins to accurately represent the position of the equal sign in equivalence problems and other equations, he is likely to note that the equal sign does not always occur at the end of the problem. However, this fact is inconsistent with his conception of the equal sign as meaning "put the answer." Recognizing this inconsistency may lead him to consider other possible interpretations of the equal sign that are consistent with his newly learned procedure. In so doing, he may infer that the equal sign indicates that two quantities are the same—a conception that is compatible with his newly learned procedure. Thus, in this instance, the boy learned a new problem-solving procedure and consequently improved his representation of equations, and his improved representation provided the basis for inferring an improved concept of the equal sign.

No evidence to date addresses this pathway in its entirety. However, there is some evidence for the first link in the pathway, namely, the link from gains in procedural knowledge to improvements in problem representation. Alibali, Ockuly, and Fischer (2005) investigated whether learning new, correct procedures via instruction led to improvements in children's representations of equivalence problems. Third- and fourth-grade students completed reconstruction and recognition tasks to assess problem representation both before and after an intervention in which some children learned a correct procedure for solving equivalence problems. The intervention focused on two different correct procedures, *equalize* (make both sides equal) and *add–subtract* (add the left side and subtract the number on the right), which were presented in a 2×2 design, yielding four groups: *equalize* only, *add–subtract* only, both procedures, or neither procedure. As seen in Figure 4, children who learned the *equalize* procedure improved their problem representations more than children who did not learn the *equalize* procedure. Learning this procedure led children to represent the right side of the problems more accurately (i.e., at posttest, fewer children made errors such as reconstructing the target problem $3 + 9 + 5 = 3 + \_$ as $3 + 9 + 5 = 3$ or $3 + 9 + 5 = +3$).

Surprisingly, learning the *add–subtract* procedure did not lead to improvements in problem representation. This finding suggests a possible reason why not all procedural lessons lead to conceptual knowledge gains: namely, some procedures may be better at promoting improved problem representation than others. Indeed, the *equalize* procedure requires children to represent the *sides* of the equation, whereas the *add–subtract* procedure requires only that children represent the numbers on the left and the number on the right. The fact that the *equalize* lesson led to improvements in children's representations of the right side

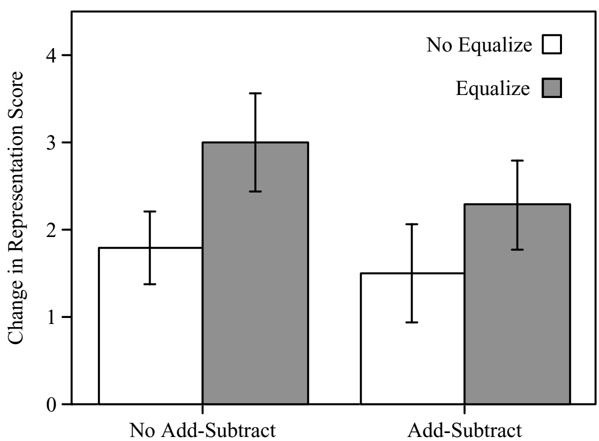

Fig. 4. Mean pretest to posttest improvements in representation scores for children who were taught the equalize *procedure, the* add–subtract *procedure, both procedures, or neither procedure (from Alibali, Ockuly, & Fischer, 2005). Children who learned the* equalize *procedure improved their problem representations more than children who did not learn the* equalize *procedure,* $F(1, 45) = 4.2$, $p < 0.05$.

of the equation is consistent with this view. Thus, some procedural lessons may be more likely to engage this pathway than others.

Unfortunately, there is not direct evidence yet for the second link in the pathway, namely, the link from improvements in problem representation to inferring correct concepts. To address this issue, studies are needed that involve manipulations to improve problem representations and measures of conceptual knowledge both before and after those manipulations. Because the process of inferring new concepts is likely to be gradual, it would seem wise to use sensitive measures of conceptual knowledge, such as rating or recognition measures (e.g., rating correct and incorrect definitions of the equal sign provided by other children) so that subtle changes in concepts can be detected. Such studies might also include probes about concepts in various contexts (e.g., asking about the meaning of the equal sign symbol in various contexts, such as an addition problem and an equivalence problem), because past research has shown that emerging conceptual knowledge may be displayed in some contexts but not others (McNeil & Alibali, in press-a).

2. Reflection about Why Procedures Work

Another possible mechanism by which learning new procedures may lead to gains in conceptual knowledge is by fostering reflection about why those procedures work. When children learn new procedures, they may choose to reflect about the rationales underlying those procedures, even if those rationales are not provided in instruction. What might compel such reflection? One possibility is that newly learned, correct procedures yield different solutions than previously used, incorrect procedures. Noting the differences in the solutions yielded by the old and the new procedures may lead children to reflect about the basis of the new procedures. Another possibility is that newly learned procedures may be more efficient than previously used procedures. If this is the case, then children may be able to implement the new procedures without using all of their available processing resources, and they may have resources available which they may allocate to consider the conceptual basis of the new procedures (*cf.* Shrager & Siegler, 1998). Reflecting about newly learned procedures may lead children to articulate principles or to form links between previously isolated pieces of knowledge, both forms of conceptual knowledge. Thus, learning new procedures may provoke reflection about why the procedures work, and this reflection may promote gains in conceptual knowledge.

Consider a girl who initially uses the *add-all* procedure to solve equivalence problems such as $3 + 9 + 5 = 3 + \_$. If the girl then learns the *cancel-and-group* procedure by observing a friend solving the same problem, the girl will note that the new procedure yields a different solution. She may also find that the *cancel-and-group* procedure is easier and less resource-intensive to implement, because it requires adding only two of the addends in the problem, rather than all four. Thus, the girl may have sufficient processing resources available to reflect on the basis of the new procedure, as well as reason to consider why her friend's procedure yields a different solution than her own prior procedure. She may reflect about the implications of canceling like addends, and realize that the new procedure serves to make both sides of the problem equal. This insight may then form the basis for inferring or articulating the principle that the two sides of an equation represent the same quantity.

Some evidence in support of this pathway can be gleaned from existing studies, most notably from McNeil and Alibali's (2000) investigation of the effects of externally imposed goals on children's learning from a procedural lesson. Past research (e.g., Dweck & Elliot, 1983; Licht & Dweck, 1984; Dweck & Leggett, 1988) had suggested that children's achievement goals influence their behavior in learning contexts. With this in mind, McNeil and Alibali (2000) compared the likelihood of making conceptual gains from a procedural lesson for three groups of children: (1) children who were provided a *learning* goal at the outset of the lesson, (2) children who were provided a *performance* goal at the outset of the lesson, and (3) children who were not provided any explicit goal at the outset of the lesson. The goal manipulations were adapted from Elliot and Dweck (1988). The performance goal manipulation focused on solving the problems correctly, and included the phrases, "the most important thing will be for you to try to solve the problems correctly," and "you will be tested on more problems so that I can see how well you can do." In contrast, the learning goal manipulation focused on understanding of the domain, and included the phrases, "the most important thing will be for you to think about the problems and understand them," and "you will do some more problems so you can see how much you've learned." Following the manipulation, all children received a lesson focusing on the cancel-and-group procedure.

Figure 5 presents the proportion of children in each of the three conditions whose scores on the conceptual assessments increased, stayed the same, or decreased from pretest to posttest. As seen in the figure, children in both goal conditions were more likely to increase their scores than children in the no goal (control) condition. However, the proportion of children who increased their scores was greatest in the group of children who received learning goals, which focused on understanding of the domain. The learning goal manipulation likely promoted conceptual gain because it encouraged children to try to understand—and therefore to reflect about—the taught procedure.

Individual children probably differ in their tendency to reflect upon the procedures they use—so this mechanism of change also provides a ready explanation for why some children acquire conceptual knowledge after a procedural lesson, and others do not. The goal manipulations in McNeil and Alibali (2000), and in particular, the learning goal manipulation, may have

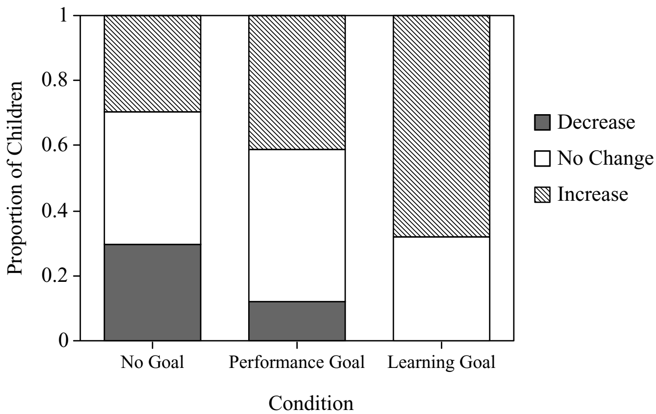

Fig. 5. Proportion of children in each of three goal conditions whose scores on the conceptual assessments increased, stayed the same, and decreased from pretest to posttest. From McNeil and Alibali (2000). Copyright © 2000 by the American Psychological Association. Reprinted with permission.

encouraged reflection among children who might not have reflected if left to their own devices.

3. *Summary*

Gains in procedural knowledge clearly *can* lead to improvements in conceptual knowledge; however, such improvements are not observed in all studies or among all children, so the mechanisms underlying this pathway must be able to explain such variability. Empirical evidence indicates that gains in procedural knowledge lead to improvements in problem representation, and it is plausible, though as yet undocumented, that improvements in problem representation may provide learners with a basis for inferring domain concepts. Also, some evidence, albeit limited, suggests that reflection about why procedures work can promote gains in conceptual knowledge. Individual learners may differ in their tendencies to reflect upon or to draw inferences based on new knowledge, so both of these purported mechanisms could account for individual differences in the likelihood of conceptual gains after learning new procedures.

These hypothesized individual differences are reminiscent of reported individual differences in students' tendencies to explain material to themselves as they study—termed "self-explanations" (Chi *et al.*, 1989). A large body of literature suggests that explanations—both when produced for the self and when produced for others—may instigate knowledge change. In Section III, I consider potential mechanisms of knowledge change that involve explanation, both in speech and in gesture.

III. Mechanisms that Involve Expressing Knowledge in Speech and Gesture

Several lines of research have suggested that expressing or explaining knowledge may lead to knowledge change. In the following subsections, I review research on two types of activities that have been linked to knowledge change in past research: (1) producing explanations, for example, when solving problems or studying worked problem examples, and (2) producing spontaneous gestures during explanations. In each subsection, I first review evidence linking the activity to knowledge change, and then consider potential mechanisms that may be responsible for the observed effects.

A. EXPLANATION AND KNOWLEDGE CHANGE

Explanation can be defined as "a family of activities aimed at making something more understandable" (Neuman, Leibowitz, & Schwarz, 2000). Explanations involve activities such as generating inferences, filling in details, articulating underlying principles, generating justifications for problem-solving steps, and so forth. Explanations typically go beyond what was directly stated in the material being explained.

A large body of literature has investigated the role of explanation in learning. The bulk of this literature has focused on two types of explanations: (1) instructional explanations, which are explanations provided by teachers, tutors, or other agents that provide instruction (such as computer-based learning environments), and (2) self-explanations, which are explanations generated by learners themselves. Because my focus in this chapter is on mechanisms that involve expression of knowledge, in this section I consider self-explanations and their role in cognitive change.

The study of self-explanations began with Chi and colleagues' seminal research on college students learning about Newtonian particle dynamics (Chi *et al*., 1989). This research investigated how students studied a chapter from a physics textbook, and in particular, how students studied the worked example problems provided in the text. Chi *et al.* analyzed the self-explanations that students generated when studying the worked examples and investigated how students' self-explanations related to their learning outcomes. The main finding was that good learners produced more self-explanations of the physics content than did poor learners. Specifically, good learners generated many explanations that refined or expanded upon the conditions for particular actions taken in the examples, extrapolated the consequences of actions beyond those stated in the examples, imposed goals for particular actions or sets of actions, or explained the meaning of quantitative expressions. Thus, the self-explanations produced by good learners contained information about the conditions, consequences, goals, and meanings of the steps taken in the examples.

The fact that good learners produced more self-explanations than poor learners suggested to Chi *et al*. that self-explanations might actually foster understanding of domain principles. To evaluate this possibility, Chi *et al.* assessed the number of components of Newton's laws that students possessed before studying the examples (on a pretest), and the number of components that they mentioned in the explanations they produced while studying the examples. The good learners stated components of the laws during their explanations that they had not mentioned at all on the pretest. Thus, self-explanations served to "bootstrap" students' understanding of domain principles.

Renkl (1997) also highlighted the role of self-explanation in understanding of principles, in a study of university freshmen studying worked examples of probability problems. Renkl coded the content of participants' self-explanations of worked examples along several dimensions, and then used cluster analysis to identify "styles" of self-explanation. Students' learning outcomes varied as a function of self-explanation style, with two style groups showing positive learning outcomes: (1) individuals who used anticipative reasoning, defined as working part of the problem before viewing the worked example, and (2) individuals whose self-explanations focused on principles. Renkl's findings about the beneficial effects of a principle-based self-explanation style are compatible with Chi *et al.*'s claim that self-explanations foster understanding of domain principles.

Renkl's study also replicated Chi *et al.*'s (1989) finding that better learners produced more self-explanations than less successful learners. This finding has also been replicated by other investigators in other participant populations and content domains, including adults learning about computer programming (Pirolli & Recker, 1994) and high school students learning to solve algebra word problems (Neuman *et al.*, 2000).

Early studies of the self-explanation effect were correlational—they linked pre-existing individual differences in explanation patterns to differences in learning outcomes. However, because of the obvious educational implications of self-explanations, intervention studies soon followed. In the first controlled study of whether eliciting self-explanations fosters learning, Chi *et al*. (1994) randomly assigned eighth-grade students studying a text about the circulatory system to self-explain during the study period or to simply read the text twice. Students in the self-explanation condition learned more than students in the control condition. Furthermore, students who generated many self-explanations were more likely to achieve the correct mental model of the circulatory system over the course of the experiment than were students who generated few self-explanations.

A large number of studies in a wide variety of participant populations and content domains have shown that elicited or prompted self-explanations are beneficial for learning. These studies have spanned a number of different types of material to be explained, including text (e.g., Chi *et al.*, 1994), worked examples (e.g., Didierjean & Cauzinille-Marmeche, 1997; Atkinson, Renkl, & Merrill, 2003), problem solutions (e.g., Pine & Messer, 2000), and problem-solving activities (e.g., Neuman & Schwarz, 1998; Aleven & Koedinger, 2002). These studies have also addressed many different content areas, including computer programming (Bielaczyc, Pirolli, & Brown, 1995), analogical problems (Neuman & Schwarz, 1998), block balancing problems (Pine & Messer, 2000), problems about compound and real interest (Renkl *et al.*, 1998), and several mathematical domains, including factoring (Didierjean & Cauzinille-Marmeche, 1997), geometry (Aleven & Koedinger, 2002), mathematical equivalence (Rittle-Johnson, 2004) and probability (Wong, Lawson, & Keeves, 2002; Atkinson, Renkl, & Merrill, 2003). A few studies have yielded null effects (e.g., Mwangi & Sweller's, 1998, study of third-grade students learning to solve "compare" word problems); however, on the whole, this body of literature suggests that eliciting self-explanations is a robust and effective technique for promoting learning and cognitive change.

For elicited self-explanations, the nature of the information explained also influences what and how much explainers learn. Siegler (1995) elicited explanations among 5-year-old children solving Piagetian quantity conservation tasks. Children in one group explained their own, usually incorrect, judgments about the equivalence of the quantities in the tasks, and they were given feedback about the correctness of their judgments. Children in another group were told the correct judgment (e.g., the two quantities were the same) by the experimenter and then explained that judgment. Children who explained the experimenter's reasoning performed better than children who explained their own reasoning.

In a related study, Siegler (2002) elicited explanations among third- and fourth-grade students solving mathematical equivalence problems. Some children were told that a child at another school had provided the correct answer, and they explained why they thought that answer was correct. Other children were told about two children at another school, one who had provided the correct answer, and one who had provided an incorrect answer. The children then explained both why the correct answer was correct and why the incorrect answer was incorrect. Compared to children who explained their own reasoning, children in both groups benefited from explaining other children's answers, but those who explained both the correct and the incorrect answers benefited more than did those who explained the correct answer only. Similar findings were reported by Curry (2004) in a parallel study of college students setting up equations for algebra word problems. Thus, the effects of self-explanation also depend on the information being explained.

Many investigators who have studied self-explanation have theorized about or investigated the mechanisms that underlie the self-explanation effect. In the following sections, I discuss three potential mechanisms that may be responsible for the beneficial effects of self-explanation on learning. Because my focus in this chapter is on changes in mathematical thinking, I emphasize mechanisms that have been studied in the domain of mathematics, although I also consider some studies from other domains. My intent is not to provide an exhaustive review of mechanistic accounts of self-explanation effects, but rather to illustrate some mechanisms that have relevance for mathematics learning. I focus here on three potential mechanisms: (1) self-explaining encourages inference generation, which creates new chunks of knowledge, (2) self-explaining promotes the integration of concepts and procedures, and (3) self-explaining leads to improved problem representation.

1. Generating Inferences

Chi and colleagues' initial claim about the mechanism underlying the self-explanation effect was that self-explanations promote inference generation. Inferences go beyond what was stated in the source material, sometimes drawing on commonsense world knowledge or domain knowledge, sometimes integrating across components of the source text, and sometimes being purely deductive inferences based on statements in the source text (Chi & VanLehn, 1991; Chi, 2000). Inferences may make tacit knowledge (learners' own prior tacit knowledge, or information implicit in the text) more explicit and available for use (Chi & VanLehn, 1991), or they may help solvers to identify and repair gaps in their knowledge of the domain (VanLehn, Jones, & Chi, 1992). Thus, generating inferences is a mechanism by which new knowledge chunks can be created.

Chi and colleagues' view was based on studies of self-explanations produced in the domains of physics and biology. Might a similar process of inference generation also apply in mathematical domains? It seems likely. A hypothetical example can illustrate how inference might lead to new knowledge in a mathematical domain. Imagine a seventh-grade girl studying a worked example of a solved equation, such as the one shown in Figure 6, and providing self-explanations. At Step 3 of the worked example, the girl might notice that 3 has been subtracted from the quantity $3 + x$, and she might notice that $3 + x - 3$ yields x. Based on these observations, she might draw the inference that the purpose of the step is to isolate the variable: "OK, so now they're subtracting 3 from $3 + x$. So I guess the idea is to get the x by itself." In this self-explanation, the girl inferred the goal for the action taken in the worked example. This inference may be a new knowledge chunk for the girl, and it may fill a gap in her understanding of algebraic manipulation.

Is there empirical evidence from mathematical domains to support the view that self-explanations involve generating inferences, which leads to the formation

$$
\begin{aligned}
&1.\quad 3 + 9 + 5 = 3 + x \\
&2.\quad 17 = 3 + x \\
&3.\quad 17 - 3 = 3 + x - 3 \\
&4.\quad 14 = x
\end{aligned}
$$

Fig. 6. Sample worked example of an algebraic equation.

of new knowledge chunks? At least one study of self-explanation in a mathematical domain has yielded such evidence—Wong, Lawson and Keeves' (2002) study of students learning a new theorem in geometry. These investigators provided a group of ninth-grade students with training in self-explanation. These students were then asked to study a textbook section that presented the theorem that an angle inscribed in a semi-circle is a right angle. Students in a control group studied the same textbook section using their usual study techniques. Students in both groups provided think-aloud protocols while studying the text. Finally, all students completed a posttest that included problems that required use of the target theorem for solution.

Students' think-aloud protocols were coded for knowledge-generation activities, defined as statements or actions that involved using given information or prior knowledge to create new connections or relations. Examples include reasoning or hypothesizing that involved the target theorem and relating the target theorem to other theorems. Thus, most knowledge-generation activities involved forming inferences based on given information or prior knowledge. Students in the self-explanation group produced significantly more of these knowledge-generation statements than controls.

The beneficial effects of self-explanation were clear on the posttest, where students in the self-explanation group outperformed students in the control group by a substantial margin, particularly on the more difficult items. Moreover, a path analysis indicated that the effects of self-explanation on posttest performance involved a pathway that included knowledge-generation activities. In fact, knowledge-generation activities were a stronger predictor of posttest performance than any of the other candidate predictors tested, including prior knowledge and beliefs about mathematics. Thus, training in self-explanation fostered students' use of knowledge-generation activities, and these activities promoted success in problem solving.

This study indicates that self-explanations can be used to generate knowledge via inference in at least one mathematical domain. However, the nature of the specific connections formed and how this new knowledge came into play in problem solving were not addressed in Wong *et al.*'s study. Detailed models, such as that developed by VanLehn, Jones, and Chi (1992) for the domain of physics, will be needed to specify the processes involved with more precision.

Future studies should attempt to specify the nature of the new knowledge chunks that are generated via inference. In addition, studies that focus on mathematical domains other then geometry are needed.

2. Promoting Integration of Concepts and Procedures

One type of inference that is often evident in self-explanations of worked examples is justifications for particular steps in the solution process (Chi *et al.*, 1989; Chi & VanLehn, 1991). It seems reasonable to assume that self-explanations that focus on such justifications may foster learning by promoting integration of concepts and procedures. Other types of self-explanations may contribute to this integration, as well. For example, explanations that invoke domain concepts may strengthen links between those concepts and procedures that are based on those concepts, even when those explanations do not focus on justifying steps in the procedures.

When concepts and procedures are well-integrated, knowledge is flexible and can be generalized to new tasks and problems (Baroody, 2003). Thus, self-explanations that promote integration of concepts and procedures seem likely to foster learners' acquisition of generalizable, transferable knowledge, and consequently, better problem-solving performance.

Some suggestive evidence from studies of self-explanation in mathematical domains supports this potential mechanism. Rittle-Johnson (2004) examined the effects of prompted self-explanations on third- through fifth-grade children learning to solve mathematical equivalence problems. Children solved equivalence problems and were given feedback on their solutions. Some children were then prompted to explain both an incorrect and a correct solution that had been obtained by "children from another school" (as in Siegler's 2002 study). Children also received either instruction about a correct procedure or encouragement to invent a new procedure.

Children who self-explained performed better than children who did not self-explain on the posttest and on a transfer test that included novel problems (e.g., problems that did not have a repeated addend), regardless of whether they had received direct instruction or encouragement to invent a procedure. To succeed on the transfer test, children needed to adapt their newly learned or invented procedures for the novel problems, and conceptual knowledge is presumably necessary to guide such adaptation. Although children rarely mentioned conceptual information explicitly in their self-explanations, the fact that children who self-explained performed better on the transfer test suggests that these children had constructed connections between their new procedures and their conceptual knowledge of the domain.

Aleven and Koedinger (2002) also provided evidence for the integration of concepts and procedures in their research on self-explanations in geometry learning. Their study utilized a high school geometry curriculum, Cognitive

Tutor Geometry, that involves a computer-based instructional environment. The study focused on a unit about angles in which students learned to solve problems that relied on various theorems (e.g., the alternate interior angles theorem) for solution. Two versions of the Cognitive Tutor were compared: (1) a self-explanation version, which prompted students to self-explain as they worked problems, and (2) the standard, problem-solving version, which did not. In the self-explanation tutor, students provided justifications for each of their problem-solving steps, either by typing in the name of the theorem or principle that justified the step or by selecting the theorem or principle from a glossary. The tutor also provided feedback about whether students' explanations were correct.

Students who used the self-explanation tutor performed better on a problem-solving posttest than did students who used the standard version of the tutor. Furthermore, the advantage of self-explanation was especially pronounced on items that required deeper understanding for success (e.g., items that required students to determine that there was not sufficient information for solution).

To gain a better understanding of students' patterns of success across items on the posttest, Aleven and Koedinger (2002) fit a mathematical model to their data. The model was based on the assumption that performance on the posttest items was due to a mixture of three types of knowledge: (1) shallow procedural knowledge, which included both incorrect procedures and guessing heuristics (e.g., "angles that look the same, are the same"), (2) correct procedural knowledge, defined as knowledge of correct procedures, and (3) conceptual knowledge, defined as well-integrated verbal and visual knowledge about the theorems used in solving the problems. These three types of knowledge were hypothesized to relate to performance on different types of posttest items in different ways. For example, correct performance on easy-to-guess posttest problems could be based on any of the three types of knowledge, whereas correct performance on some of the more difficult items required conceptual knowledge. Aleven and Koedinger formulated equations that captured the hypothesized relations between the three types of knowledge and performance on the different types of posttest items. They then fit these equations to the data, generating estimates of the strength of each of the three knowledge types for students in each of the experimental conditions.

The modeling results suggested that students in the standard condition acquired primarily correct procedural knowledge, whereas students in the self-explanation condition acquired both conceptual knowledge and correct procedural knowledge. These findings are compatible with the idea that self-explanation fostered students' understanding of the conceptual basis of their problem-solving procedures.

Although Aleven and Koedinger (2002) and Rittle-Johnson (2004) have provided suggestive evidence that self-explanation promotes integration of concepts and procedures, neither study has presented direct evidence for this mechanism. Future studies should include assessment items that tap integration of concepts and procedures more directly.

3. Improvements in Problem Representation

Another possible mechanism by which self-explanation may promote learning is via improvements in problem representation. In her research on self-explanations in the domain of biology, Chi (2000) discussed this basic idea in terms of *revising mental models*. To evaluate this potential mechanism, Chi diagnosed eighth-grade students' mental models of the circulatory system before they read and self-explained a text about it. With knowledge of students' initial mental models, it was possible to identify information in the to-be-explained text that would conflict with each student's model. In an intensive microgenetic analysis of one girl's self-explanations, Chi (2000) showed that at the point when the girl detected the conflict between the information in the text and her own mental model, she attempted to resolve the conflict, "resulting in [a] long and tortuous explanation," (p. 214) which appeared to help her repair her initial mental model. The repair process involved adding features that were not present in her original model and integrating links in the model. Thus, self-explaining helped the student to construct a better representation of the circulatory system.

Research on self-explanations of analytical reasoning problems also suggests that self-explanation promotes accurate problem representation. Neuman and Schwarz (1998) found that university students who were prompted to self-explain a set of analytical reasoning problems were more successful at solving a target problem than were students who simply provided think-aloud protocols. A content analysis of the protocols suggested that good solvers (most of whom were in the self-explain condition) focused more on the "deep structure" of the target problem, whereas poor solvers (most of whom were in the control condition) focused more on the problem's surface structure. Neuman and Schwarz interpreted these data as suggesting that self-explaining supports solvers in representing the problem's deep structure. However, it should be noted that their study did not include measures of problem representation independent of problem solution.

At least one study of self-explanation in a mathematical domain has also provided suggestive evidence that self-explanations can lead to improved problem representation. Didierjean and Cauzinille-Marmeche (1997) examined self-explanations produced by ninth-grade students studying worked examples of algebra problems that involved factoring. For participants who had some initial knowledge about factoring, prompts to explain led to better performance on a posttest than either of two control conditions, one in which students simply solved problems and one in which students did nothing between pretest and posttest. Didierjean and Cauzinille-Marmeche distinguished two subgroups of participants in the explanation condition based on their patterns of performance on the posttest, namely, participants who performed well on all of the posttest problems and those who performed well only on problems that were highly similar to the worked examples. For the subgroup that performed well on all of the posttest problems, evidence from participants' verbalizations suggested that they constructed an abstract problem schema that they then applied to the posttest problems.[2] Specifically, students in this group often made remarks that pointed out structural similarities between different problems. Like Neuman and Schwarz (1998), Didierjean and Cauzinille-Marmeche argued that self-explaining led students to represent the problems in terms of an abstract schema that captured the problems' deep structure.

The evidence about improvements in representation from Didierjean and Cauzinille-Marmeche's (1997) research is indirect, because there was no explicit measure of problem representation either before or after the explanation manipulation. To my knowledge, there are no studies in mathematical domains that have directly investigated whether self-explanation promotes changes in representation. However, self-explanations may well foster improvements in students' representations of mathematics problems.

To illustrate, suppose a boy were asked to study the worked example shown in Figure 6, and to provide self-explanations. Suppose further that the boy had used the *add-all* procedure to solve similar equations in a pretest prior to studying the example, and that in a pretest of problem representation, he made several errors typical of children who use the *add-all* procedure, such as reconstructing the given equation $4 + 6 + 7 = 4 + x$ as $4 + 6 + 7 + 4$. When reading step 2 of the worked example, the boy might wonder aloud why the solver in the example did not also add the second 3 in the problem, as he himself would have done when solving this problem. Attempting to explain this solution step may lead him to notice the position of the equal sign in the problem, which he may not have represented correctly before studying the worked example: "Hmm, I wonder why they didn't add the other 3. It has a plus sign after it… Oh! But it has an equal sign in front of it." In this case, being prompted to provide explanations for each solution step impelled the child to improve his representation of the problem by accurately representing the position of the equal sign.

If children improve their problem representations as a result of self-explanations (either prompted or spontaneous), this may lead to other sorts of changes in their knowledge. As discussed in the previous section, improved problem representation may lead to gains in both conceptual and procedural knowledge. With respect to conceptual knowledge, improved problem representation may provide children with a more accurate basis for inferring concepts, as in the example presented previously (p. 93). With respect to procedural knowledge, improved problem representation may foster generation

[2]The other subgroup appeared to focus on the specifics of each worked example. Participants in this subgroup appeared to have solved the posttest problems using case-based reasoning.

of new procedures (e.g., Alibali *et al.*, 1998) or may lead to better choices among existing alternative procedures (Siegler, 2002).

The idea that self-explanation promotes improved problem representation is a plausible one, with suggestive evidence to support it. However, few studies, and none yet in the domain of mathematics, have provided compelling evidence that learners' problem representations do in fact improve as a result of self-explanation. Research on this issue in mathematical domains is needed.

4. *Summary*

Existing data support the possibilities that self-explaining leads to changes in mathematical knowledge because (1) self-explaining encourages explainers to generate new knowledge by making inferences, and (2) self-explaining promotes the integration of concepts and procedures. There is less empirical evidence from mathematical domains to suggest that producing self-explanations leads to improved problem representation, which in turn may lead to gains in both conceptual and procedural knowledge. However, this pathway has received indirect support from studies in other problem-solving domains, so it remains a plausible candidate mechanism worthy of future test in mathematical domains.

B. GESTURE PRODUCTION AND KNOWLEDGE CHANGE

The research on explanation and knowledge change has focused primarily on the *verbal* component of explanations. However, in many cases, people produce gestures as well as speech when they articulate inferences or explain how they solved problems. Might these gestures also play a role in knowledge change? Mounting evidence suggests that they do.

Many previous studies have investigated the gestures speakers produce when explaining their problem solutions, both in mathematical tasks (e.g., Goldin-Meadow, Alibali, & Church, 1993), and in non-mathematical tasks (e.g., Crowder & Newman, 1993). It is now widely accepted that speakers' gestures reveal their understanding of the tasks being explained (e.g., Garber *et al.*,1998; Alibali *et al.*, 1999).

Several studies have shown that gestures index transition periods in development and learning. In particular, discrepancies between gesture and speech in children's task explanations, called "gesture–speech mismatches", are associated with readiness to learn. For example, Church and Goldin-Meadow (1986) found that children who frequently produced gestures that mismatched their speech on a pretest of Piagetian conservation were more likely to learn from a brief lesson about conservation than were children who seldom produced gesture–speech mismatches. Similar findings have been reported in several studies of children learning to solve mathematical equivalence problems (Perry, Church, & Goldin-Meadow, 1988; Goldin-Meadow & Singer, 2002). (For examples of gesture–speech mismatches in children's explanations of mathematical equivalence problems, see Table I.)

However, it is not only mismatching gestures that are associated with knowledge change—any gestures may be important. In a study of children learning the concept of mathematical equivalence, Alibali and Goldin-Meadow (1993) found that children who gestured while explaining problems on a pretest tended to learn more from a brief lesson than children who did not gesture, regardless of whether their gestures matched or mismatched speech. The posttest in this study was a paper-and-pencil test that included both addition problems (e.g., $3 + 9 + 5 = 3 + \_$) and multiplication problems (e.g., $4 \times 2 \times 3 = 4 \times \_$). As seen in Figure 7, non-gesturers tended to succeed only on the addition problems on the posttest, whereas gesturers tended to succeed on both types of problems. Thus, gesture production was associated with deeper learning.

Why is gesture production associated with knowledge change? There are at least two possibilities. First, gesture may be an epiphenomenon. It may reflect speakers' knowledge, but not be directly involved in the process of knowledge change itself. Alternatively, gesture may play a functional role in the process of knowledge change. That is, producing gestures may actually influence the path of learning. In this section, I review evidence that producing gestures plays a functional role in knowledge change.

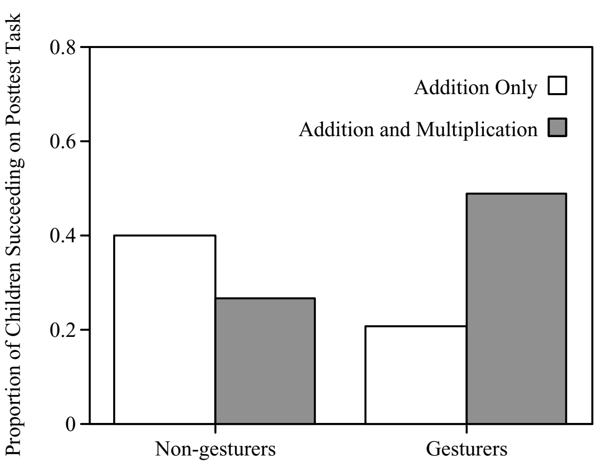

Fig. 7. Proportion of children who produced gestures (N = 43) and who did not produce gestures (N = 15) who succeeded on the paper-and-pencil posttest. Children who produced gestures tended to succeed on both addition and multiplication problems, whereas children who did not produce gestures tended to succeed on addition problems only. Adapted from Alibali, M. W, & Goldin-Meadow, S., Transitions in learning: What the hands reveal about a child's state of mind, Cognitive Psychology, *25, 468–523, copyright © 1993, with permission from Elsevier.*

Research on the function of gesture has focused on several possible mechanisms by which gesture may play a role in knowledge change. Here, I consider two such mechanisms: (1) gesture generates or activates knowledge based in perception and action, making this information available for reasoning and problem solving, and (2) gesture "off-loads" or frees working memory for other purposes, such as procedure generation.

1. Gesture Generates or Activates Knowledge Based in Perception and Action

One pathway by which gesture may play a role in knowledge change is by generating or activating knowledge based in perception and action, making this information available for other cognitive processes. A great deal of evidence suggests that speakers' gestures highlight perceptual and motoric information for *listeners* (e.g., Graham & Argyle, 1975; Kendon, 1994; Valenzeno, Alibali, & Klatzky, 2003). Here, I consider evidence that producing gestures may highlight such information for speakers themselves. If gesture production leads speakers to focus on particular types of information, then speakers who produce gestures should traverse a different path in learning than speakers who do not gesture.

Gestures often express information that is based in perception or motor actions. For example, when a speaker talks about a cup, the accompanying gestures may represent perceptually salient properties of the cup (e.g., size, shape, or height) or information about how the speaker could physically manipulate the cup (e.g., by picking it up, tilting it, or drinking from it). Producing such gestures may activate this perceptual and motor information so that it is readily available for reasoning and problem solving. As a consequence, speakers who produce gestures may tend to focus on perceptual and motor information, rather than other sorts of information, in their problem solutions and problem explanations.

Several lines of research support the claim that gesture promotes a focus on perceptual-motor information. First, gesture appears to help speakers activate images or maintain them in memory. In one study of this issue, de Ruiter (1998) presented speakers with arrays of shapes and lines and asked them to describe them to a listener. For some speakers, the images remained visible during the descriptions, and for other speakers, the images were removed during the description. Speakers produced more gestures when the images were no longer visible, suggesting that gestures helped them to retrieve the images and maintain them in mind. These findings were replicated by Wesp and colleagues, using still-life paintings instead of arrays of shapes (Wesp *et al.*, 2001).

Second, when speakers are prevented from gesturing, they focus less on perceptual or motor information. However, it should be noted that most gesture-prohibition studies have been conducted in non-mathematical domains. For example, Rimè and colleagues examined the content of spontaneous conversations held by speakers who were allowed to gesture and speakers who were prohibited from gesturing (Rimè *et al.*, 1984). The content of the speech was assessed using a "computer program of content analysis conceived to quantify the degree of speech imagery" (p. 317). When gestures were prohibited, speakers received lower imagery scores, supporting the view that gesture promotes a focus on perceptual information. However, this conclusion is necessarily tentative, because Rimè *et al.* provided little information about their content analysis program.

Stronger evidence on this issue was provided by Alibali, Spencer, and Kita (2005), who studied the procedures adults use to solve gear movement prediction problems with gesture allowed and with gesture prohibited. In gear movement prediction problems, participants imagine an array of gears described by the experimenter, and determine how a particular gear would move if another gear were moved in a certain way (e.g., "Imagine five gears are arranged in a circle. If you try to turn the gear on top clockwise, what would the gear just to its left do?"). Past research had shown that solvers frequently used gestures to model the movements of each individual gear as they attempt to solve the problems (Schwartz & Black, 1996).

Alibali *et al.* found that participants who were allowed to gesture tended to use depictive procedures, in which they described the movement of each individual gear, throughout the set of problems. In contrast, participants who were prevented from gesturing often generated rule-based procedures (most often the parity rule, which holds that if there are an odd number of gears, the last gear goes the same direction as the first, and if an even number, the last gear goes in the opposite direction). Gesture appeared to help solvers to mentally simulate the actions of each gear, and therefore promoted use of depictive procedures. Thus, gesture promoted reasoning based on motor actions rather than abstract rules. Speakers solved the problems correctly at comparable rates in both conditions, but gesture influenced their choice of problem-solving procedures.

Only one study to date has examined gesture and perceptual-motor reasoning in a mathematical domain. Alibali, Kita and colleagues (Alibali *et al.*, 2001; Alibali & Kita, 2005) studied the justifications that first- and second-grade students provided for Piagetian quantity conservation tasks when they were allowed to gesture and when they were prohibited from gesturing. All children solved and provided justifications for two sets of conservation tasks. During the first set, all children were allowed to gesture; for the second set, children were randomly assigned to either a gesture-allowed or a gesture-prohibited condition. When gesture was allowed, children focused primarily on perceptual characteristics of the task objects (e.g., the heights of the glasses of water, the lengths of rows of checkers) in their verbal explanations of the tasks. When gesture was prohibited, children were more likely to invoke information that was not perceptually present in the display, such as the initial equivalence of the quantities or the transformation that had been performed.

These findings suggest that, when gesture is allowed, perceptual information is highlighted as children attempt to formulate an explanation for the conservation task. When gesture is prohibited, perceptual information is less salient, so children are more likely to focus on non-perceptual information in formulating their responses. From this study, it is not clear whether prohibiting gestures influenced children's representations of the tasks themselves, or the aspects of those representations that children chose to verbalize. However, in either case, producing gestures appeared to play a role in children's thinking.

But *how* might gesture contribute to thinking? According to Kita (2000), speakers use gestures to explore alternative ways of organizing perceptual and motoric information in planning spoken utterances, in an effort to find a conceptualization that can readily be expressed in speech (see also Hostetter and Alibali, 2004)). I suggest that this exploration may lead to changes in the speaker's mental representation of the problem at hand. In the conservation task, children believe that the objects contain information relevant to the experimenter's question, so when gesture is allowed, children use gesture to explore the physical characteristics of the objects as they attempt to represent the situation and formulate their responses. As a consequence of this exploration, perceptual information should be central in children's representations of the tasks when gesture is allowed. When gesture is prohibited, children cannot use gesture to explore the perceptually present array, so other information about the tasks should be relatively more salient. This is exactly the pattern of data observed by Alibali and Kita in children's verbal explanations.

This view implies that children's problem representations should differ when gesture is allowed vs. when gesture is prohibited. Furthermore, if producing gestures provokes changes in children's problem representations, it may ultimately lead to changes in children's procedural and conceptual knowledge, as described in the first section of this chapter. Future studies that directly test these predictions are needed.

Are there other mathematical domains besides quantity conservation in which gesture might influence reasoning and problem solving? Gestures are especially likely to accompany speech that has spatial or imagistic content (Lavergne & Kimura, 1987; Hadar & Krauss, 1999), so mathematical domains that involve visual imagery, such as geometry, may be particularly likely to elicit gestures. Geometry problems often involve reasoning with images, and using gestures in reasoning about or explaining such problems might promote accurate representation of the problem content.

Likewise, solving story problems often involves forming a mental image of the problem situation, and producing gestures could help speakers to envision such situations more accurately. This possibility gains support from a study of adults solving algebra story problems dealing with discrete and continuous change (Alibali *et al.*, 1999). Participants were asked to read story problems silently, and then explain the gist of each problem to another participant. Speakers produced many gestures when describing the problems. After describing each problem, participants were asked to solve that problem. Speakers who produced gestures that conveyed discrete, incremental change during their problem descriptions tended to use additive procedures to solve the problems, whereas speakers who produced gestures that conveyed smooth, continuous change during their problem descriptions tended to use multiplicative procedures for solving the problems. The gestures speakers produced while describing the problems may have influenced their mental representations of the problems, and consequently, influenced their solution procedures.

In sum, there is growing evidence that gesture promotes a focus on perceptual and motor information. To date, most studies of this function of gesture have focused on non-mathematical tasks, such as the gear movement prediction task. It seems probable that gesture could also influence problem representation in mathematical domains by highlighting perceptual information, particularly for tasks that involve a perceptual component. However, research directly addressing this issue is needed.

2. Gesture Off-Loads Working Memory

Another possible mechanism by which gesture production leads to knowledge change is by off-loading working memory, so that it can be used for other purposes, such as procedure discovery.

A number of researchers have suggested that gestures provide support for working memory. In one study of this issue, Alibali and DiRusso (1999) investigated the role of gesture in preschoolers' counting. Children counted more accurately when they gestured to the counted objects themselves than when they were prohibited from gesturing or when a puppet gestured for them as they produced the number words. Alibali and DiRusso (1999) argued that gestures serve to externalize some of the contents of working memory, so they need not be held internally. They further proposed that keeping track of counted objects and tagging each object with a number word require fewer working memory resources when the objects are marked physically, with gestures, than when the objects are marked visually, by looking at each item (as when gesture was prohibited or when the puppet gestured). When children gestured to each of the counted objects themselves, they counted more successfully.

Other researchers have explored the role of gesture in managing the memory demands of problem explanation. Goldin-Meadow *et al.* (2001) asked two groups of participants to explain mathematical problems while remembering sets of words or letters. Nine- and ten-year-old participants solved and explained mathematical equivalence problems, and adult participants solved and explained factoring problems of the form $x^2 + 5x + 6 = (\ ?\ )(\ ?\ )$. Participants in both groups were prohibited from gesturing during some of their problem explanations and free to gesture during other problem explanations. Both child and adult participants remembered more of the words or letters when they gestured than when they did not. Goldin-Meadow *et al*. interpreted these findings as showing that gesture off-loads working memory, freeing up capacity for the secondary task of remembering words or letters. In essence, gesture "lightens the load" of explanation.

A possible alternative account of Goldin-Meadow *et al*.'s findings is that, instead of gesture lightening the load, being prohibited from gesturing may actually *add* to speakers' load. If this were the case, the findings would be due, not to the beneficial effects of gesture, but to the deleterious effects of being still. To address this possibility, Goldin-Meadow *et al*. reanalyzed the data from a subset of participants—those who chose not to gesture on some of the problem explanations for which they were allowed to do so. The effect of not gesturing was the same whether participants refrained from gesturing spontaneously or did so in response to the experimenter's instructions. Both child and adult participants remembered more of the words or letters when they gestured than when they did not gesture, whether by choice or by instruction. Furthermore, these findings could not be attributed to differences in the difficulty of solving the problems. The percentage of problems solved correctly did not differ as a function of whether participants gestured, did not gesture by choice, or did not gesture by instruction during the problem explanations.

These findings were replicated by Wagner, Nusbaum, and Goldin-Meadow (2004) with the factoring task as the primary task, and with two different secondary tasks: a verbal memory task (remembering lists of letters) and a visuo-spatial memory task (remembering the locations of dots on a grid). In both cases, participants performed better on the memory task when they produced gestures during their problem explanations.[3] Furthermore, as in the prior study, not gesturing spontaneously had the same effects as not gesturing in response to instruction—in both cases, participants remembered less when they failed to gesture.

The findings of Goldin-Meadow *et al*. and Wagner *et al*. suggest that gestures off-load working memory, freeing capacity for other tasks. But how might

[3]In light of research showing that gesture promotes a focus on perceptual information (Alibali & Kita, 2005; Alibali, Spencer, & Kita, 2005), it seems surprising that prohibiting gesture had a comparable impact on verbal and visuo-spatial memory. One possible resolution of this puzzle has to do with differences in the types of gestures produced across the studies. The factoring task studied by Wagner *et al*. elicited primarily pointing gestures that referred to the numbers on the blackboard. In contrast, the conservation and gears tasks studied by Alibali and colleagues elicited primarily iconic gestures, which depict information via handshape or motion. It is possible that different types of representations underlie pointing and iconic gestures. Moreover, it seems likely that different types of gestures serve different cognitive functions. However, testing these possibilities remains a task for future research.

off-loading working memory be involved in knowledge change? One possibility is that, if resource demands of a primary task (e.g., explanation) are minimized, resources can be used for activities that lead to knowledge change, such as forming more accurate representations of the problems to be solved, or discovering better, more efficient procedures (e.g., Shrager & Siegler, 1998).

If this is the case, then gesture prohibition should make procedure discovery less likely. However, as described previously, in a non-mathematical task, Alibali, Spencer, and Kita (2005) found just the opposite—when speakers were prohibited from gesturing, they were *more* likely to discover a rule-based procedure for solving gears problems. In this study, speakers who were free to gesture perseverated in their use of a depictive procedure. These findings highlight that many factors contribute to processes of procedure selection and procedure change. Producing gestures may decrease the resource demands of explanation, but the availability of gesture for physical modeling may make participants unlikely to shift away from the depictive procedure.

3. Summary

Empirical evidence provides tentative support for at least two mechanisms by which gesture production is involved in knowledge change. Gesture promotes a focus on perceptual and motor information, and may thereby influence problem representation. There is evidence for this mechanism in the literature, although not primarily in mathematical domains. Alternatively, or in addition, gesture may serve to off-load working memory. There is evidence from multiple mathematical domains that gesture serves this function; however, the role of this function in knowledge change needs to be better specified.

IV. Concluding Remarks

In this chapter, I have considered two broad classes of change mechanisms that have been proposed to apply in the domain of mathematical reasoning: (1) mechanisms that involve reciprocal relations between knowledge of problem-solving procedures and knowledge of concepts, and (2) mechanisms that involve expressing knowledge in speech and gestures. Within each class, several candidate mechanisms were described, the empirical evidence for each was reviewed, and gaps in the research base addressing each mechanism were identified.

In this final section, I address two issues. First, are there any core similarities among the mechanisms discussed in this chapter? I consider two characteristics that are shared by many of the mechanisms: (1) the role of the learner's own activity, and (2) a focus on problem representation. Second, is it possible to predict which mechanisms will apply in particular situations? I discuss the possibilities that some mechanisms are better suited for certain types of content or for certain types of learners. I also consider whether there are developmental differences in which mechanisms apply.

A. SIMILARITIES AMONG MECHANISMS

The present review of change mechanisms provokes an obvious question: Are there any similarities that are shared by many or all of them?

1. Emphasis on Learner's Activity

One notable similarity is that many of the mechanisms discussed herein depend on the learner's own activity. Some of the mechanisms require overt actions on the part of the child, such as producing self-explanations or producing gestures. Many of the mechanisms involve constructive mental processes, such as generating inferences (either based on improved problem representations, or as a result of self-explanation), reflecting about why procedures work, forming links between concepts and procedures, and mentally simulating actions or images with the support of gestures. Knowledge change via these mechanisms is not a passive process, but instead requires active processing on the part of the learner.

However, active processes are ones that learners can choose to engage in or not to engage in. Indeed, even when prompted to engage in active processes, individuals differ in how much they do so. For example, Chi *et al.* (1994) found that, among a group of learners who were prompted to provide self-explanations of a text about the circulatory system, some provided more extensive self-explanations than others, and consequently learned more.

Given the central role of active processing in many mechanisms of knowledge change, one key issue for future theories of knowledge change is explaining why some learners are more likely than others to engage in active processing. Better understanding of this issue should shed light on individual differences in learning outcomes. In particular, variations in learners' own activities may be associated with variations in learning from instruction.

2. Emphasis on Problem Representation

Another notable similarity across many of the mechanisms discussed herein is an emphasis on problem representation. Many of the mechanisms involve changes in representation: gains in conceptual knowledge lead to improved problem representation, which may instigate procedure generation; gains in procedural knowledge lead to improved problem representation, which may provide a basis for inferring concepts; self-explanations may lead to improved representation of the problem domain, which may lead to gains in both conceptual and procedural knowledge; and producing gestures promotes a focus on certain types of information in learners' problem representations.

In addition to its role in mechanisms discussed in this chapter, problem representation may also play a role in other change mechanisms, not discussed herein, that involve problem-solving performance. Problem representation is a key determinant of problem-solving performance (e.g., Siegler, 1976; McNeil & Alibali, 2004), and performance plays a crucial role in altering knowledge in some models of developmental change (e.g., McClelland, 1995; Shrager & Siegler, 1998). According to such models, each time a problem is solved, the results feed back to influence the learner's knowledge about that class of problems. Thus, because problem representation influences how problems are solved, changes in problem representation are a potentially powerful engine of knowledge change.

B. WHICH MECHANISMS APPLY WHEN?

Given that there are many possible mechanisms of knowledge change, is it possible to predict which mechanisms may apply in particular situations or for particular individuals?

1. Differences across Content Domains

It seems probable that certain mechanisms are more likely to be engaged in some content domains than others. Much of mathematical learning involves solving problems, so mechanisms that involve procedures may be particularly relevant in mathematical domains. Some of the mechanisms discussed in this chapter are triggered by learning new procedures (e.g., learning new procedures leads to improvements in representation, and improved representations are a basis for inferring concepts). Other mechanisms involve procedures in other ways (e.g., self-explanation of mathematical content promotes the formation of links between concepts and procedures). More generally, knowledge change in mathematics may be particularly likely to engage mechanisms that involve procedures, either as a starting point for change or as an object of change.

It may also be the case that certain sub-domains within mathematics are particularly well suited for certain mechanisms and not others. For example, geometry is a highly visual domain, so mechanisms that involve gesture production may be especially likely to apply in learning about geometry.

2. Individual Differences

It also seems highly probable that individuals differ in the particular mechanisms that they engage. Unfortunately, little is known about how individual differences in cognitive skills, learning styles, personality traits, and other characteristics influence the mechanisms that individuals use. However, it is easy to generate hypotheses about what mechanisms might be likely or unlikely to apply depending on individual characteristics. It seems likely, for example, that individual differences in verbal skill might influence frequency of self-explanation, which has consequences for several mechanisms of knowledge change. Similarly, it is easy to imagine that individual differences in learners' propensities to produce gestures might influence whether or not learners engage mechanisms that involve gesture production. However, to date, neither of these hypotheses has been tested.

The literature does include some hints about dimensions of individual differences that may be associated with tendencies to apply particular change mechanisms. McNeil and Alibali (2000) addressed one of these dimensions—namely, variations in the goals that children bring to the learning situation. In their study, children who were provided learning goals (e.g., understand the problems) were more likely to display conceptual gains in response to procedural instruction than children who were provided with performance goals (e.g., do well on the problems).[4] These findings suggest that children who enter a learning situation with learning goals (regardless of whether those goals are intrinsic to the child or imposed from the outside) may be particularly likely to reflect about why procedures work. Studies such as this one represent a tentative first step toward addressing how mechanisms of change depend on individual differences. This issue is an important arena for future work.

3. Developmental Differences

It is also worth considering whether there are developmental differences in which mechanisms apply and when. It is possible that different mechanisms are most commonly engaged at different ages. However, little research has addressed this issue directly. No studies to my knowledge have compared whether children of different ages are differentially likely to acquire conceptual knowledge after learning new procedures, or differentially likely to benefit from self-explaining.

Most research on mechanisms that involve reciprocal relations between concepts and procedures has been conducted with elementary school children. In their early school years, children learn many new concepts and procedures, and the domains children are learning about (e.g., mathematical equivalence, decimal fractions) are tractable for study. However, learning of concepts and procedures is not limited to the elementary school years. For example, children learn procedures and concepts related to counting in the preschool years. Research on early counting has also focused on relations between concepts and procedures, specifically on the possibilities that preschoolers infer counting principles from implementing the counting procedure (e.g., Briars & Siegler, 1984), and that preschoolers' knowledge of counting principles informs their implementation

[4]This difference between the learning and performance goal groups was observed at the immediate posttest. Two weeks later, at a follow-up assessment, there were no differences between the groups in the proportion of participants who made conceptual gains.

of the counting procedure (e.g., Gelman & Meck, 1983). It seems likely that the reciprocal relations between concepts and procedures that have been observed in various domains are a consequence of how knowledge is organized, rather than due to some characteristic of child learners *per se*. If this is the case, then the same mechanisms should apply in learners of any age.

Research on mechanisms that involve expressing knowledge in language and gesture has been conducted with both children and adults. The benefits of self-explanation have been documented in elementary school children (e.g., Pine & Messer, 2000; Rittle-Johnson, 2004), high school students (e.g., Aleven & Koedinger, 2002; Wong, Lawson, & Keeves, 2002) and adults (e.g., Chi *et al*., 1989; Renkl, 1997), and the effects of producing gestures on procedure use have also been documented both in children (Alibali & Kita, 2005) and in adults (Alibali, Spencer, & Kita, 2005). Of course, in order for explanations or gestures to influence learners' knowledge, those learners must be old enough to produce explanations or gestures in the first place. However, once learners are able to do so, it seems likely that the mechanisms should apply regardless of age.

The possibility that certain change mechanisms apply regardless of age has not been directly tested for any of the mechanisms discussed in this chapter. Furthermore, because individuals of different ages are typically learning different things, age and content domain are likely to be confounded, so it would be difficult to conduct a strong test of the invariance of change mechanisms across development. However, despite the challenges inherent in testing such claims, research on this issue would be valuable.

4. Multiple Mechanisms

Finally, it is important to acknowledge that in real-world learning situations, multiple mechanisms of change are certainly in play at any given time. For example, consider a girl who learns a new problem-solving procedure during a lesson in her mathematics class. The girl may begin to represent features of the problem that she had not previously noticed. In addition, learning the new procedure may provoke her to self-explain as she solves practice problems, and her self-explanations may incorporate gestures as well as speech. Thus, she may engage multiple change mechanisms simultaneously, and the mechanisms may feed into and reinforce one another. At the same time, the boy sitting at the desk beside her, who experiences the same procedural lesson, may engage a different set of mechanisms, and he may learn something entirely different from the lesson.

Understanding knowledge change in real-world settings will require an appreciation of the simultaneous and interactive nature of mechanisms of change, as well as the considerable variability that is manifested across content domains and across individuals. Although these issues present substantial challenges, a better understanding of underlying mechanisms also promises substantial payoffs, both for advancing theory and for informing practice.

ACKNOWLEDGEMENTS

Preparation of this chapter was supported by grant #BCS-0096129 from the National Science Foundation. Any opinions, findings, conclusions or recommendations expressed herein are those of the author and do not necessarily reflect the views of the National Science Foundation.

I thank Chuck Kalish, Nicole McNeil, Bethany Rittle-Johnson, and Susan Wagner for thoughtful comments on previous versions of the chapter, and Julia Evans, Susan Goldin-Meadow, Sotaro Kita, Eric Knuth, Ken Koedinger, and Mitchell Nathan for helpful discussions of the issues addressed herein.

REFERENCES

Aleven, V. A. W. M. M., & Koedinger, K. R. (2002). An effective metacognitive strategy: Learning by doing and explaining with a computer-based Cognitive Tutor. *Cognitive Science*, *26*, 147–179.

Alibali, M. W. (1999). How children change their minds: Strategy change can be gradual or abrupt. *Developmental Psychology*, *35*, 127–145.

Alibali, M. W., Bassok, M., Solomon, K. O., Syc, S. E., & Goldin-Meadow, S. (1999). Illuminating mental representations through speech and gesture. *Psychological Science*, *10*, 327–333.

Alibali, M. W., & DiRusso, A. A. (1999). The function of gesture in learning to count: More than keeping track. *Cognitive Development*, *14*, 37–56.

Alibali, M. W., & Goldin-Meadow, S. (1993). Transitions in learning: What the hands reveal about a child's state of mind. *Cognitive Psychology*, *25*, 468–523.

Alibali, M.W., & Kita, S. (2005). *On the role of gesture in thinking and speaking: Prohibiting gesture alters children's explanations of conservation tasks*. Manuscript submitted for publication.

Alibali, M. W., Kita, S., Bigelow, L. J., Wolfman, C. M., & Klein, S. M. (2001). Gestures play a role in thinking for speaking. In C. Cavé, I. Guaïtella, & S. Santi (Eds.), *Oralité et gestualité: Interactions et comportements multimodaux dans la communication [Orality and gestuality: Multimodal interaction and behavior in communication]. Actes du colloque [Proceedings of the meeting of] ORAGE* (pp. 407–410). Paris, France: L'Harmattan.

Alibali, M. W., & McNeil, N. M. (in preparation). *Changes in problem representation as a mechanism of knowledge change*.

Alibali, M. W., McNeil, N. M., & Perrott, M. A. (1998). What makes children change their minds? Changes in problem encoding lead to changes in strategy selection. In M. A. Gernsbacher & S. Derry (Eds.), *Proceedings of the 20th Annual Conference of the Cognitive Science Society* (pp. 36–41). Mahwah, NJ: Erlbaum.

Alibali, M. W., Ockuly, K. M., & Fischer, A. D. (2005). *Sources and consequences of changes in problem representation*. Paper presented at the Annual Meeting of the Jean Piaget Society, Vancouver, British Columbia.

Alibali, M. W., Spencer, R. C., & Kita, S. (2005). *Spontaneous gestures influence strategy choice in problem solving*. Paper presented at the Biennial Meeting of the Society for Research in Child Development, Atlanta, Georgia.

Atkinson, R. K., Renkl, A., & Merrill, M. M. (2003). Transitioning from studying examples to solving problems: Effects of self-explanation prompts and fading worked-out steps. *Journal of Educational Psychology*, *95*, 774–783.

Baroody, A. J. (2003). The development of adaptive expertise and flexibility: The integration of conceptual and procedural knowledge. In A. J. Baroody & A. Dowker (Eds.), *The development of arithmetic concepts and skills* (pp. 1–33). Mahwah, NJ: Erlbaum.

Bielaczyc, K., Pirolli, P., & Brown, A. L. (1995). Training in self-explanation and self-regulation strategies: Investigating the effects of knowledge acquisition activities on problem solving. *Cognition and Instruction*, *13*, 221–252.

Bisanz, J., & LeFevre, J. (1990). Strategic and nonstrategic processing in the development of mathematical cognition. In D. F. Bjorklund (Ed.), *Children's strategies: Contemporary views of cognitive development* (pp. 213–244). Hillsdale, NJ: Erlbaum.

Blöte, A. W., Van der Burg, E., & Klein, A. S. (2001). Students' flexibility in solving two-digit addition and subtraction problems: Instruction effects. *Journal of Educational Psychology*, *93*, 627–638.

Briars, D., & Siegler, R. S. (1984). A featural analysis of preschoolers' counting knowledge. *Developmental Psychology*, *20*, 607–618.

Byrnes, J. P., & Wasik, B. A. (1991). Role of conceptual knowledge in mathematical procedural learning. *Developmental Psychology*, *27*, 777–786.

Canobi, K., Reeve, R. A., & Pattison, P. E. (2003). Patterns of knowledge in children's addition. *Developmental Psychology*, *39*, 521–534.

Carpenter, T. P. (1986). Conceptual knowledge as a foundation for procedural knowledge: Implications from research on the initial learning of arithmetic. In J. Hiebert (Ed.), *Conceptual and procedural knowledge: The case of mathematics* (pp. 113–132). Hillsdale, NJ: Erlbaum.

Chi, M. T. H. (2000). Self-explaining expository texts: The dual processes of generating inferences and repairing mental models. In R. Glaser (Ed.), *Advances in instructional psychology* (Vol. 5, pp. 161–237). Mahwah, NJ: Erlbaum.

Chi, M. T. H., Bassok, M., Lewis, R., Reimann, P., & Glaser, R. (1989). Self-explanations: How students study and use examples in learning to solve problems. *Cognitive Science*, *13*, 145–182.

Chi, M. T. H., de Leeuw, N., Chiu, M. H., & LaVancher, C. (1994). Eliciting self-explanations improves understanding. *Cognitive Science*, *18*, 439–477.

Chi, M. T. H., Feltovich, P. J., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive Science*, *5*, 121–152.

Chi, M. T. H., & VanLehn, K. (1991). The content of physics self-explanations. *Journal of the Learning Sciences*, *1*, 69–106.

Church, R. B., & Goldin-Meadow, S. (1986). The mismatch between gesture and speech as an index of transitional knowledge. *Cognition*, *23*, 43–71.

Crowder, E. M., & Newman, D. (1993). Telling what they know: The role of gesture and language in children's science explanations. *Pragmatics and Cognition*, *1*, 341–376.

Curry, L. A. (2004). The effects of self-explanations of correct and incorrect solutions on algebra problem solving performance. In K. Forbus, D. Gentner, & T. Regier (Eds.), *Proceedings of the 26th Annual Meeting of the Cognitive Science Society*. Mahwah, NJ: Erlbaum.

de Ruiter, J.-P. (1998). *Gesture and speech production*. Doctoral dissertation, Katholieke Universiteit Nijmegen, Nijmegen, The Netherlands.

Didierjean, A., & Cauzinille-Marmeche, E. (1997). Eliciting self-explanations improves problem solving: What processes are involved? *Cahiers de Psychologie Cognitive/Current Psychology of Cognition*, *16*, 325–351.

Dweck, C. S., & Elliot, E. S. (1983). Achievement motivation. In P. H. Mussen & E. M. Hetherington (Eds.), *Handbook of child psychology: Vol. IV. Social and personality development* (pp. 643–691). New York: Wiley.

Dweck, C. S., & Leggett, E. L. (1988). A social-cognitive approach to motivation and personality. *Psychological Review*, *95*, 256–273.

Elliot, E. S., & Dweck, C. S. (1988). Goals: An approach to motivation and achievement. *Journal of Personality and Social Psychology*, *54*, 5–12.

Fuson, K. C. (1990). Conceptual structures for multiunit numbers: Implications for learning and teaching multidigit addition, subtraction, and place value. *Cognition and Instruction*, *7*, 343–403.

Fuson, K. C., & Briars, D. (1990). Using a base-ten blocks learning/teaching approach for first- and second-grade place-value and multidigit addition and subtraction. *Journal for Research in Mathematics Education*, *21*, 180–206.

Garber, P., Alibali, M. W., & Goldin-Meadow, S. (1998). Knowledge conveyed in gesture is not tied to the hands. *Child Development*, *69*, 75–84.

Gelman, R., & Meck, E. (1983). Preschoolers' counting: Principles before skill. *Cognition*, *13*, 343–359.

Goldin-Meadow, S., Alibali, M. W., & Church, R. B. (1993). Transitions in concept acquisition: Using the hand to read the mind. *Psychological Review*, *100*, 279–297.

Goldin-Meadow, S., Nusbaum, H., Kelly, S. D., & Wagner, S. M. (2001). Explaining math: Gesturing lightens the load. *Psychological Science*, *12*, 516–522.

Goldin-Meadow, S., & Singer, M. A. (2002). From children's hands to adults' ears: Gesture's role in the learning process. *Developmental Psychology*, *39*, 509–520.

Graham, J. A., & Argyle, M. (1975). A cross-cultural study of the communication of extra-verbal meaning by gestures. *International Journal of Psychology*, *10*, 57–67.

Hadar, U., & Krauss, R. M. (1999). Iconic gestures: The grammatical categories of lexical affiliates. *Journal of Neurolinguistics*, *12*, 1–12.

Hiebert, J., & LeFevre, P. (1986). Conceptual and procedural knowledge in mathematics: An introductory analysis. In J. Hiebert (Ed.), *Conceptual and procedural knowledge: The case of mathematics* (pp. 1–27). Hillsdale, NJ: Erlbaum.

Hiebert, J., & Wearne, D. (1989). Cognitive changes during conceptually based instruction on decimal fractions. *Journal of Educational Psychology*, *81*, 507–513.

Hiebert, J., & Wearne, D. (1996). Instruction, understanding and skill in multidigit addition and subtraction. *Cognition and Instruction*, *14*, 251–283.

Hostetter, A. B., & Alibali, M. W. (2004). On the tip of the mind: Gesture as a key to conceptualization. In K. Forbus, D. Gentner, & T. Regier (Eds.), *Proceedings of the 26th Annual Meeting of the Cognitive Science Society*. Mahwah, NJ: Erlbaum.

Kendon, A. (1994). Do gestures communicate? A review. *Research on Language and Social Interaction*, *27*, 175–200.

Kieran, C. (1981). Concepts associated with the equality symbol. *Educational Studies in Mathematics*, *12*, 317–326.

Kita, S. (2000). How representational gestures help speaking. In D. McNeill (Ed.), *Language and gesture* (pp. 162–185). Cambridge, UK: Cambridge University Press.

Lavergne, J., & Kimura, D. (1987). Hand movement asymmetry during speech: No effect of speaking topic. *Neuropsychologia*, *25*, 689–693.

Licht, B. G., & Dweck, C. S. (1984). Determinants of academic achievement: The interaction of children's achievement orientations with skill area. *Developmental Psychology*, *20*, 628–636.

Ma, L. (1999). *Knowing and teaching elementary mathematics: Teachers' understanding of fundamental mathematics in China and the United States*. Mahwah, NJ: Erlbaum.

McClelland, J. L. (1995). A connectionist perspective on knowledge and development. In T. J. Simon & G. S. Halford (Eds.), *Developing cognitive competence: New approaches to process modeling* (pp. 157–204). Hillsdale, NJ: Erlbaum.

McNeil, N. M., & Alibali, M. W. (2000). Learning mathematics from procedural instruction: Externally imposed goals influence what is learned. *Journal of Educational Psychology*, *92*, 734–744.

McNeil, N. M., & Alibali, M. W. (2004). You'll see what you mean: Students encode equations based on their knowledge of arithmetic. *Cognitive Science*, *28*, 451–466.

McNeil, N. M., & Alibali, M. W. (in press-a). Knowledge change as a function of mathematics experience: All contexts are not created equal. *Journal of Cognition and Development*.

McNeil, N. M., & Alibali, M. W. (in press-b). Why won't you change your mind? Knowledge of operational patterns hinders learning and performance on equations. *Child Development*.

Mwangi, W., & Sweller, J. (1998). Learning to solve compare word problems: The effect of example format and generating self-explanations. *Cognition and Instruction*, *16*, 173–199.

Neuman, Y., Leibowitz, L., & Schwarz, B. (2000). Patterns of verbal mediation during problem solving: A sequential analysis of self-explanation. *Journal of Experimental Education*, *68*, 197–213.

Neuman, Y., & Schwarz, B. (1998). Is self-explanation while solving problems helpful? The case of analogical problem solving. *British Journal of Educational Psychology*, *68*, 15–24.

Perry, M. (1991). Learning and transfer: Instructional conditions and conceptual change. *Cognitive Development*, *6*, 449–468.

Perry, M., Church, R. B., & Goldin-Meadow, S. (1988). Transitional knowledge in the acquisition of concepts. *Cognitive Development*, *3*, 359–400.

Pine, K. J., & Messer, D. J. (2000). The effect of explaining another's actions on children's implicit theories of balance. *Cognition and Instruction*, *18*, 35–51.

Pirolli, P., & Recker, M. (1994). Learning strategies and transfer in the domain of programming. *Cognition and Instruction*, *12*, 235–275.

Renkl, A. (1997). Learning from worked-out examples: A study in individual differences. *Cognitive Science*, *21*, 1–29.

Renkl, A., Stark, R., Gruber, H., & Mandl, H. (1998). Learning from worked-out examples: The effects of example variability and elicited self-explanations. *Contemporary Educational Psychology*, *23*, 90–108.

Rimè, B., Shiaratura, L., Hupet, M., & Ghysselinckx, A. (1984). Effects of relative immobilization on the speaker's nonverbal behavior and on the dialogue imagery level. *Motivation and Emotion*, *8*, 311–325.

Rittle-Johnson, B. (2004). Promoting flexible problem solving: The effects of direct instruction and self-explaining. In K. Forbus, D. Gentner, & T. Regier (Eds.), *Proceedings of the 26th Annual Meeting of the Cognitive Science Society*. Mahwah, NJ: Erlbaum.

Rittle-Johnson, B., & Alibali, M. W. (1999). Conceptual and procedural knowledge of mathematics: Does one lead to the other? *Journal of Educational Psychology*, *91*, 175–189.

Rittle-Johnson, B., Siegler, R. S., & Alibali, M. W. (2001). Developing conceptual understanding and procedural skill in mathematics: An iterative process. *Journal of Educational Psychology*, *93*, 346–362.

Schwartz, D., & Black, J. B. (1996). Shuttling between depictive models and abstract rules: Induction and fallback. *Cognitive Science*, *20*, 457–497.

Shrager, J., & Siegler, R. S. (1998). SCADS: A model of children's strategy choices and strategy discoveries. *Psychological Science*, *9*, 405–410.

Siegler, R. S. (1976). Three aspects of cognitive development. *Cognitive Psychology*, *8*, 481–520.

Siegler, R. S. (1995). How does change occur? A microgenetic study of number conservation. *Cognitive Psychology*, *28*, 225–273.

Siegler, R. S. (2002). Microgenetic studies of self-explanation. In N. Grannott & J. Parziale (Eds.), *Microdevelopment: Transition processes in development and learning* (pp. 31–58). Cambridge, UK: Cambridge University Press.

Siegler, R. S., & Crowley, K. (1994). Constraints on learning in nonprivileged domains. *Cognitive Psychology*, *27*, 194–226.

Siegler, R. S., & Shipley, C. (1995). Variation, selection, and cognitive change. In T. Simon & G. Halford (Eds.), *Developing cognitive competence: New approaches to process modeling* (pp. 31–76). Hillsdale, NJ: Erlbaum.

Valenzeno, L., Alibali, M. W., & Klatzky, R. L. (2003). Teachers' gestures facilitate students' learning: A lesson in symmetry. *Contemporary Educational Psychology*, *28*, 187–204.

VanLehn, K., Jones, R. M., & Chi, M. T. H. (1992). A model of the self-explanation effect. *Journal of the Learning Sciences*, *2*, 1–59.

Wagner, S. M., Nusbaum, H., & Goldin-Meadow, S. (2004). Probing the mental representation of gesture: Is handwaving spatial? *Journal of Memory and Language*, *50*, 395–407.

Wesp, R., Hess, J., Keutmann, D., & Wheaton, K. (2001). Gestures maintain spatial imagery. *The American Journal of Psychology*, *114*, 591–600.

Wong, R. M. F., Lawson, M., & Keeves, J. (2002). The effects of self-explanation training on students' problem solving in high-school mathematics. *Learning and Instruction*, *12*, 233–262.

This Page Intentionally Left Blank

A SOCIAL IDENTITY APPROACH TO ETHNIC DIFFERENCES IN FAMILY RELATIONSHIPS DURING ADOLESCENCE

Andrew J. Fuligni and Lisa Flook

DEPARTMENT OF PSYCHOLOGY, UNIVERSITY OF CALIFORNIA,
LOS ANGELES, CA 90095, USA

In the past 25 years, systematic research into the development of adolescents within the family has surged dramatically. *PsycInfo*, the premier database of psychological and related journals, documents only 414 articles, chapters, and dissertations on adolescence and family relationships from 1970 to 1979.[1] In contrast, 1758 papers appeared from 1980 to 1989 and 2912 papers were completed from 1990 to 1999. The total for the 1990s represents a sevenfold increase in published research on family relationships during the teenage years since the 1970s, as compared to a fourfold increase in publications on

[1]Search conducted on May 14, 2002, simultaneously using "adolescence" as a population term and "family relations" as a descriptor term.

Advances in Child Development and Behavior
R Kail (Editor)

adolescence in general. Along with the increase in the sheer number of studies of family relationships during adolescence has come a recent expansion of the ethnic backgrounds of the families being investigated, such that a field that had been dominated by studies of white, middle-class adolescents increasingly includes investigations of more ethnic minority, immigrant, and culturally diverse teenagers and their families. The increase in diverse samples has occurred much later and at a slower rate than research on family relationships in general, but there has been a noticeable rise in work on ethnically diverse families in the United States that now allows for reviews of the literature in order to identify general trends and offer tentative conclusions.

In this chapter, we argue that taking a social identity approach to adolescent development in the family can help in the understanding of ethnic differences in family relationships within American society. Current theories and explanations for ethnic differences in family relationships focus on the importance of factors such as cultural traditions, socioeconomic status and stress, and discrimination and other aspects of social threat. We believe that viewing family membership as a social identity for adolescents, particularly among those from ethnic minority families, can help to explain why these different factors are important for ethnic variability in family relationships. Families are not only collections of dyadic relationships, they also are social groups functioning in a society that distributes resources, opportunities, and challenges unequally and according to social categories that often are defined by family membership and origin. The family, therefore, is one of the primary contexts by which adolescents interpret and make sense of larger social categories such as ethnicity, having implications for the manner in which adolescents identify with social groups.

If family membership indeed can serve as a social identity for adolescents, then the principles of social identity that have been applied to other social groups such as workplaces, peer groups, and political parties also may be applied to the family. Several theories in different disciplines have addressed issues of social and collective identity. Within psychology the tradition most frequently associated with a social identity approach is social identity theory, which was first proposed by Tajfel to describe the dynamics of intra- and inter-group relations among different ethnic and cultural groups (Tajfel, 1972; Tajfel & Turner, 1979). Borrowing from social identity theory and related work, the basic premises of a social identity approach that are relevant for understanding family relationships during adolescence are fourfold: like other members of the human species, adolescents have a powerful tendency to seek out and identify with social groups; adolescents will identify with social groups when they believe that they are valued members of those groups; adolescents will identify more strongly with their social groups when functional use is made of those groups and when those groups perceive external threat; and group identification will result in a tendency to act in support of the well-being of the group, as well as greater internalization of the values of that group.

There is likely to be little disagreement over these basic principles, and we are not the first observers to bring one or more of them to bear upon specific aspects of adolescent development within the family (e.g., Tyler, Degoey, & Smith, 1996). Rather, we are suggesting that as one of the primary social groups to which children belong, the family can function as a significant social identity for adolescents and that some of the dynamics that govern the intra- and inter-group relations of other social groups also apply to families. A social identity approach, therefore, offers a useful lens with which to view the growing research based on ethnic variations in family relationships during adolescence. In addition, a social identity approach can help to stimulate and enhance new and continuing areas of inquiry.

In this chapter, we first describe a social identity approach to group identification and argue that family membership can serve as an important social identity for adolescents. Second, we discuss the reasons why family identification may be particularly salient for teenagers from ethnic minority backgrounds in the United States. Third, we outline the ways in which a social identity approach can help to explain observed patterns of ethnic differences in family relationships during adolescence. Finally, we suggest how a social identity approach can generate new and significant questions to be pursued in the next era of research on ethnic differences in family relationships.

I. A Social Identity Approach

As defined by Tajfel (1972, p. 292), a social identity is "the individual's knowledge that he belongs to certain social groups together with some emotional or value significance to him of his group membership." Although a social identity is clearly an aspect of the self-system of the individual, it usually distinguished from a personal identity that rests on idiosyncratic attributes and relationships. A social or collective identity is considered to be a part of the self that derives from group membership, processes, and inter-group behavior (Ashmore, Deaux, & McLaughlin-Volpe, 2004). Examples of social identities include those based upon gender, ethnicity, political party, and occupation. Thousands of studies have tested the principles of social identity theory among groups defined along real and salient social categories (e.g., gender and ethnicity), as well as among novel and artificial groups created within a laboratory setting. Several general patterns of inter- and intra-group dynamics have emerged from these studies, and the following are particularly relevant for family relationships during adolescence: individuals have a powerful tendency to identify with social groups, even when the groups are defined in minimal ways; individuals are more likely to identify with groups when they believe that they are respected and valued members of the group; individuals will identify more strongly with their groups when functional use is made of the groups and when the groups perceive external threat; and group identification is associated with a tendency to act in support of the well-being of the group, and a greater internalization of the values of the group (Hogg, 2003; Tyler, 1999).

A social identity approach has been applied to the study of virtually any social group, both real and artificial, outside of the family. We believe, however, that family membership can be an important social identity for children and adolescents. Many of the features used to define other social groups within the study of social identity also can be applied to the family. The family is the first and perhaps primary social group to which children belong. The family is usually the social group that is ultimately responsible for the care and feeding of children, and children depend on the family for other resources and opportunities for their development. The family is one of the primary referents by which children define and organize their experiences. To the outside world, children's family membership usually is made socially obvious through salient cues such as shared surname, co-residence, and physical appearance. Family membership is experienced by children through shared values, norms, and beliefs that often differentiate families from one another. Family membership, therefore, is a salient and significant way in which the social world is differentiated and defined for children and adolescents, creating the conditions favorable for the establishment of family as a social identity.

Family membership can function as a social identity for children as early as middle childhood, when children begin to develop the cognitive abilities necessary in order to identify with social groups (Ruble *et al.*, 2004). The years of adolescence, however, should be a period when figuring out the importance of family as a social identity becomes a particularly salient developmental task for children. Children engage in a great deal of both personal and collective identity work during the teenage years because of increased cognitive skills and an expansion of their social worlds (Harter, 1990). The social groups to which they belong expand from the years of middle childhood, and become intertwined with issues of social acceptance and conformity (Brown, 1990). Adolescents increasingly become aware of the broader social categories in which they are placed by the larger society, such as those defined by gender and ethnicity, and the opportunities and constraints that these categories place upon their values, behaviors, and aspirations (Phinney, 1990). Social identities, therefore, have great salience during the teenage years as adolescents attempt to negotiate the social groups and categories into which they are placed either willingly or unwillingly. Negotiating the place that family membership holds within the larger set of social groups of children's lives is one of the most critical developmental tasks facing children during the adolescent years.

Developmentalists generally have not taken such a social identity approach to family relationships during adolescence, instead focusing on dyadic relationships between individual family members. Approaches that have focused on whole-family functioning, such as family-systems theories and family cohesion studies, move beyond single dyads but still focus on aspects of intimate relationships such as conflict and emotional closeness. A social identity approach to the family is related to, but nevertheless distinct from analyses of dyadic and intimate relationships within the family. Relationships with individual parents and siblings have implications for whether adolescents identify with the family as a social group, and it is unlikely that adolescents with extremely poor family relationships will hold family as an important or valued social identity. But a social identity approach to the family focuses on the group level, and, to paraphrase Tajfel's original definition, a family identity refers to adolescents' knowledge that they belong to the family as a social group together with the emotional or value significance to them of their family membership. Family identity refers to a sense of "we-ness" that goes beyond simply the collection of individual dyadic relationships with primary and extended family members. It is a collective identity that if salient and strong, results in the internalization of values and group-oriented behaviors that stem from being a group member rather than from just the dyadic relationships with individual members. For example, an adolescent may attend family events or assist siblings even if they are not currently getting along with family members, simply because they are "part of the family". Similarly, a young adult may assist aging parents regardless of the quality of their relationships because "they are my parents".

Our point is not that family membership must be a social identity for children and adolescents, or that it is of the same importance throughout development and across different individuals and ethnic groups. Rather, we believe that the potential exists for family membership to serve as a social identity for children during the adolescent years, and it is the variability in the importance of family identity that helps to account for some of the ethnic differences that have been observed in family relationships during adolescence. In the following sections, we first argue that family serves as an especially important social identity for ethnic minorities in the United States, and then interpret major findings regarding ethnic differences in family relationships during adolescence through the lens of a social identity approach to the family.

II. Ethnicity and Family Identity

Many reasons exist for believing that family membership serves as a particularly important social identity for adolescents from ethnic minority families in the United States. First, traditions of family support, assistance, and identification exist within the cultural backgrounds of most ethnic minority groups. Many Asian traditions, such as Confucianism, emphasize family respect and devotion (Ho, 1996). Similarly, a loyalty and commitment to family are often expected from individuals in Latin American societies (Sabogal *et al.*, 1987). Family assistance and support also has long been a cultural tradition for African American and Native American families (Harrison *et al.*, 1990; Joe & Malach, 1998).

Interestingly, studies of acculturation and generational change within families suggest that these traditions of family importance and assistance remain strong even after ethnic minority families have lived in the United States for several generations. In our research, we have observed that ethnic minority status is a far stronger predictor than generational status of adolescents' attitudes toward supporting, assisting, and respecting the authority of the family. For example, adolescents from Filipino, Mexican, Chinese, and Central American backgrounds all believe in the importance of assisting family members more strongly than adolescents from European backgrounds (Fuligni, Tseng, & Lam, 1999). Few differences exist among those from Asian and Latin American backgrounds according to whether they or their parents were born in the United States, and when they do exist, the differences are small and the adolescents' attitudes remain much different than those of adolescents from European backgrounds (the majority group). In addition, even among children as young as 7 and 9 years of age, the beliefs about family support and assistance among those with African American backgrounds are more similar to the beliefs of children from immigrant minority groups (i.e., Chinese and Dominican) than children from European American backgrounds (Fuligni *et al.*, in press).

The continued importance of family membership and obligation across several generations of ethnic minority families in an American society that so strongly emphasizes individual initiative and achievement points to the strength of the cultural traditions of these groups as well as the existence of contemporaneous social factors that heighten the importance of family membership as a social identity. Social identity theory and numerous supportive studies suggest two particularly relevant contemporaneous factors: group identification increases when group membership is made functionally salient and when groups perceive external threat (Hogg, 2003). Ethnicity and ethnic group membership are social categories that are made functionally salient in American society. Ethnicity is one of the primary social categories by which individuals are grouped, and resources, opportunities, and challenges are distributed according to ethnic background. The functional use of ethnicity becomes particularly salient for children during the years of adolescence. Educational resources and success increasingly fall along ethnic lines during early and middle adolescence, with Asian and white students being disproportionately placed in higher academic tracks with higher quality teachers than African American and Latino students (Dornbusch, Glasgow, & Lin, 1996; Oakes *et al.*, 1991). Organized activities, sports, and popular culture and leisure activities become ethnically defined (e.g., Schreiber & Chambers, 2002). Peer groups cleave along ethnic lines upon entry into middle school, leaving even adolescents who had mixed-race friends in elementary schools with more ethnically homogenous peer groups (Shrum, Cheek, & Hunter, 1988).

In addition to living in a society in which ethnicity is made functionally salient, ethnic minority families perceive and experience more external threat than those in the majority (García Coll *et al.*, 1996). The threat may be as minimal as feeling different and out of place and as great as experiencing direct hostility, prejudice, and discrimination. The perception and experience of external threat increase as children enter adolescence and increasingly become aware of the larger society. Instances of discrimination and being the target of ethnic slurs or harassment happen more often to teenagers than to younger children (Fisher, Wallace, & Fenton, 2000; Szalacha *et al.*, 2003). One study of a group of ethnically diverse adolescents found that about half of the youths reported an incident of being the target of a racial slur and approximately one-third said that they were threatened by their peers because of their race or ethnicity (Fisher, Wallace, & Fenton, 2000). Youths from African American and Latino backgrounds generally report higher levels of adult and institutional discrimination than Asian American youths, but those from Asian American backgrounds tend to report among the highest levels of ethnic discrimination and harassment from peers (Phinney & Chavira, 1995; Way, 1998). For example, Asian American teenagers are more likely to report being teased and bothered by peers because of their ethnicity whereas African American and Latino youth report more mistrust from adults such as store owners and authority figures. In general, adolescents from European American backgrounds report levels of discrimination far lower than those reported by those from ethnic minority backgrounds.

It is important to note that an early hypothesis of Tajfel's social identity theory was that when group members perceive that their group is of low social status, which is one of the primary messages of social threat and discrimination, they would try to protect their self-image by actually disidentifying with their group and seeking to join other groups. According to this line of reasoning, members of ethnic minority groups actually would show a tendency to disidentify with their ethnic background because of their generally lower social standing in American society. Yet this early hypothesis of social identity theory has not been supported, particularly when the larger society ascribes group membership to individuals and makes the boundaries between groups relatively impermeable. Instead of trying to deny their group membership, members in low status groups attempt to improve their standing by challenging prevailing characteristics of their group, redefining the essential features of their group, and engaging in social comparisons on other dimensions that are advantageous to them (Hogg, 2003). Members of ethnic minority groups often engage in such actions, which serve to strengthen their group identification under conditions when the functional use and social threat associated with ethnic group membership work to their disadvantage.

The functional use of ethnicity and the perception of external threat among ethnic minority groups have implications for identification with the family because family of origin is one of the primary ways in which individual's ethnic group membership is established in American society. Adolescents may attempt to choose their ethnicity, and often engage in very creative and sophisticated negotiations of their ethnic identity across time and place, but the range of identity options available to them is constrained by the larger society according to their family membership (Waters, 1999). Schools and other governmental institutions place adolescents into official ethnic categories such as Black, White, Latino, Asian, and Native American on the basis of their family of origin. Even peers from similar ethnic backgrounds place limits on the kinds of ethnic identities that adolescents can adopt, often by pointing to adolescents' family membership. Ethnographies of African American and Latino teenagers have highlighted how American-born members of these groups sometimes distinguish themselves from foreign-born adolescents from the Caribbean and Latin America and form separate peer groups and cliques on the basis on nativity (Matute-Bianchi, 1991; Waters, 1999). For example, Matute Bianchi observed that in a high school dominated by Mexican-descent students, American-born students were more likely to consider themselves to be "Mexican Americans" and "Chicanos" whereas the foreign-born students considered themselves "Mexicanos," and peer groups among the students generally consisted of members of the same social category.

We believe that because ethnic group membership is so closely tied to family membership, the social identity dynamics related to ethnicity should be relevant to family identification. Specifically, family membership should be a particularly important social identity for adolescents from ethnic minority backgrounds because of the functional use of ethnicity in American society and the perception of external threat among ethnic minority groups. Taking a social identity approach to the family, therefore, may help us to better understand the nature of ethnic differences in family relationships during the teenage years.

III. Ethnic Differences in Family Relationships

If the dynamics of social identity that apply to other social groups also impact family relationships, then a social identity approach to the family would predict four specific patterns of ethnic differences in family relationships during adolescence. First, adolescents from ethnic minority groups would have a stronger identification with the family than other youths. Second, ethnic minority teenagers would feel a greater sense of obligation to support and assist the family than other youths. Third, because the social identity dynamics that are relevant for ethnic differences in family relationships operate at the between-group level, the family identifications and obligations of ethnic minority adolescents would be greater than what would be predicted from their relationships within the family. Similarly and finally, adolescents from ethnic minority families would show greater internalization of the values of the family than would be predicted from relationships within the family. Although more research needs to be done in order to fully test these four predictions, several sets of findings already exist to suggest that the predictions will be born out.

A. FAMILY IDENTIFICATION AND OBLIGATION

Many studies have highlighted the great importance of family membership and family obligation to adolescents from ethnic minority backgrounds (Azmitia *et al.*, 1996; Cooper *et al.*, 1993; Freeberg & Stein, 1996; Gaines *et al.*, 1997; Janoff-Bulman & Leggatt, 2002; Phinney, Ong, & Madden, 2000). We have observed similar ethnic differences in our studies of adolescents' sense of obligation to the family. In one study of approximately 1000 high school students from a variety of ethnic backgrounds, we asked adolescents to complete measures that assessed their sense of obligation to support, assist, and respect family members (Fuligni, Tseng, & Lam, 1999). The first measure, which we called *current assistance*, referred to the degree to which adolescents believed that they should assist with household tasks and help the family. Youths indicated how often they felt they should engage in activities such as run errands for the family, help out around the house, take care of brothers and sisters, and spend time with other family members. The second type of obligation was called *family respect* and students were asked to evaluate the importance of respecting parents and older family members, doing well for the sake of the family, and making sacrifices for the family. The final aspect of family obligations that we assessed involved the value the students placed upon *supporting the family in the future*. Youths indicated how important they believed it was to help their parents financially in the future, live or go to college near their parents, and help take care of their parents and other family members in the future.

As shown in Figure 1, adolescents from Asian and Latin American families endorsed all three aspects of family obligations more strongly than the youths from European backgrounds. Chinese, Filipino, Mexican, and Central/South American youths believed that they should assist and spend time with their family, respect their parent's wishes and make sacrifices for their family, and support their family in the future more than did their European American peers. These differences tended to be large, sometimes reaching more than a full

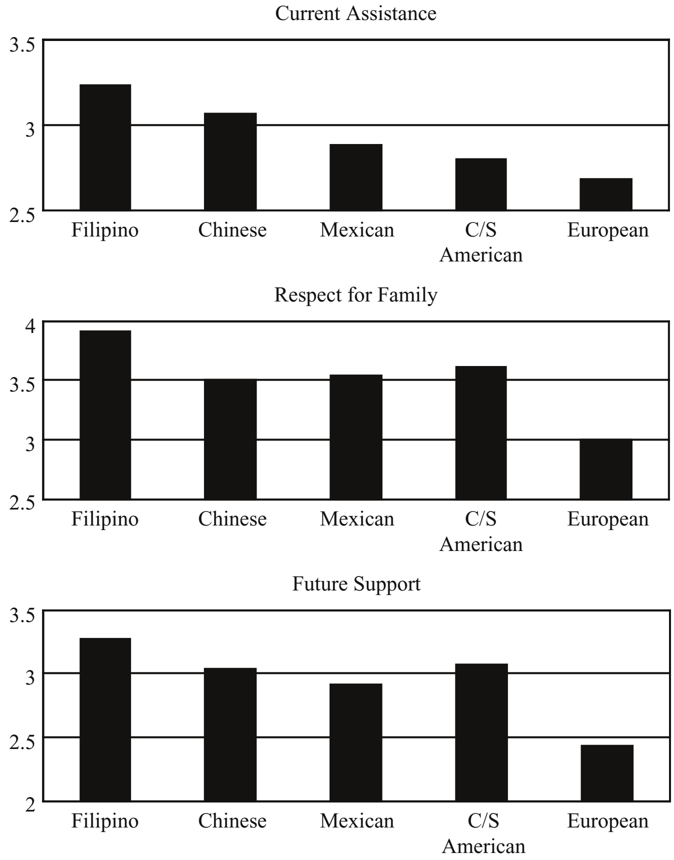

Fig. 1. Ethnic differences in adolescents' sense of obligation to the family. *Figure adapted from results presented in Fuligni, Tseng, and Lam (1999). All attitudes were measured on a 5-point scale, where a higher score indicates a stronger attitude. "C/S" refers to Central/South American.*

standard deviation in magnitude, and could not be accounted for by ethnic variations in socioeconomic status and family composition (e.g., family size and the presence of grandparents). Consistent with other research, therefore, ethnic background appears to be the strongest predictor of adolescents' sense of obligation to the family. Additionally, most studies indicate that teenagers from ethnic minority backgrounds, regardless of the number of generations their families have lived in American society, generally have a greater inclination to support and assist the family than adolescents from European American backgrounds, who represent the ethnic majority in the United States.

An important gap in existing research, however, is the lack of studies that examine generic identification with the family among adolescents from different ethnic and immigrant groups. The evidence just cited focuses on teenagers' sense of obligation and duty to support, assist, and respect the authority of the family. Although a motivation to voluntarily support the group is a consistent correlate of group identification in many studies (e.g., Tyler, 1999), it is important to assess adolescents' generic identification with the family in a manner that does not necessarily imply obligation and assistance. Given the basic significance of the family for all children and adolescents, regardless of ethnic and cultural background, there may well be very few ethnic differences in a generic identification with the family which would suggest that the family is an important group in adolescents' lives. Rather, ethnic differences may be most evident in the aspect of identification that implies an obligation to support and assist the group. If so, it remains to be debated whether family identification without a sense of obligation is a weaker identification, or whether it simply implies a difference in the nature of family membership as a social identity.

B. WITHIN-FAMILY RELATIONSHIPS

The third and fourth predictions of a social identity approach to the family state that family obligation and internalization would be greater among ethnic minority adolescents than would be predicted by their relationships within the family. At this point we need to distinguish between social identity dynamics that operate at the between-group and the within-group levels. Issues of functional use and external threat operate at the between-group level. In other words, the focus of these issues is on how resources, opportunities, and challenges are distributed among different ethnic groups in American society. To the extent that these dynamics operate, they should produce ethnic group differences in identification with the family that are independent of the interactions and relationships among members within each of the groups. The nature of the relationships within families does have an important impact upon identification with the family, but it is more important in explaining variations in identification among different members of the same family and the same ethnic group. For example, the quality of adolescents' relationships with their parents may account for the difference in family identification between two Mexican American adolescents, but it will be less relevant for understanding the difference between a Mexican American and a European American adolescent. Another way to conceptualize the dynamics that operate at the two levels is to consider them to be additive: an individual adolescent's identification with the family depends on both the adolescent's relationships within the family as well as the adolescent's ethnic group membership.

In our research, a consistent finding has been that ethnic differences are larger and more evident in adolescents' sense of obligation to the larger family as compared to their relationships with individual family members. In the study that was described previously, we asked teenagers to report on their levels of closeness and conflict with their parents (Fuligni, 1998). All adolescents, regardless of their ethnic background, indicated similarly moderate levels of closeness with their mothers and slightly, but significantly lower levels of closeness with their fathers. The teenagers from different ethnic backgrounds also reported similarly low levels of conflict with mothers and fathers, although the frequency of conflict with mothers was greater than the rate of conflict with fathers. In no case did the adolescents from Asian or Latin American backgrounds indicate closer, more intimate, or less conflictual relationships with their mothers and fathers than those from European American backgrounds. The ethnic minority adolescents even reported a decline in closeness with mothers and fathers with increasing age that was similar to the decline reported by the European American adolescents.

What makes the divergent patterns of ethnic similarities in dyadic relationships and ethnic differences in family identification and obligation so intriguing is the fact that adolescents' dyadic relationships with their parents and their sense of family obligation are significantly correlated with one another at the individual level. As shown in Table I, our studies have shown that adolescents who feel closer and more intimate with their mothers and fathers feel a greater sense of obligation to the support, assist, and respect the family. These results are consistent with other work that has taken a relational approach to social identity in other social groups. For example, Tyler (1999) has observed that employees who have good relationships and feel trusted and respected by supervisors feel a greater sense of identification and pride in the company and are more willingly to voluntarily assist other employees. Indeed, a central tenet of a social identity approach is that individuals will more strongly identify with a group if they feel that they are a valued and respected member of the group. Adolescents with close and intimate relationships with their parents likely feel a greater sense of value and respect than other teenagers, leading to a stronger sense of obligation to the family. Yet the within-group dynamics of social identity and the family do not

TABLE I

Correlations between Family Obligation and Parent–Adolescent Cohesion

	Current assistance (r)	Respect for family (r)	Future support (r)
Cohesion with mother	0.35	0.38	0.33
Cohesion with father	0.36	0.31	0.26

Note: Table adapted from figures presented in Fuligni, Tseng, and Lam (1999). All correlations significant at $p < 0.001$.

seem to be as relevant for ethnic differences in family relationships, as demonstrated by the fact that ethnic minority teenagers have a stronger sense of family obligation than majority teenagers who report the same levels of closeness and intimacy with their parents.

The pattern of findings just described suggests that a social identity approach to the family also offers a way to resolve an apparent contradiction in some cultural and developmental theories about adolescent development in the family. Some theorists have argued that non-Western and ethnic minority families deemphasize personal autonomy and press for children to concentrate on the collective needs and wishes of the group. Yet at the same time, ethnic minority American adolescents and those in collectivist societies still believe in a personal domain that allows children and adolescents to participate in decision making about certain aspects of their lives (Smetana, 2002). The apparent contradiction exists in part because some theorists have conceptualized adolescent autonomy as a unidimensional construct, rather than taking the contemporary developmental perspective which holds that adolescents' relationships within the family involve the dual dimensions of autonomy and relatedness (Smetana, 2002). Therefore, even European American adolescent development involves consideration of the needs and wishes of group members, and development among teenagers in ethnic minority families and collectivistic societies involves concern over independence and personal agency.

A social identity approach to the family offers an additional way to resolve the apparent contradiction between contemporary cultural and developmental approaches to adolescent development within the family, at least in terms of ethnic differences in American society. Social identity theory holds that group identification is enhanced when individuals believe that they are valued and respected members of the group. An important way in which a sense of value and respect is engendered is by allowing group members to make contributions to the group and to respect group members' ideas and opinions (Hogg, 2003; Tyler, 1999). Personal agency, therefore, is not antithetical to a concern for the needs and wishes of the larger group. In fact, allowing for the development of personal agency helps to maintain group identification and enhance the ability of group members to become capable individuals who can make valuable contributions to the larger group. It would seem important for the group's long-term interests, therefore, to provide members with certain domains in which they can develop a sense of personal agency. The same is true for families from ethnic minority and cultural backgrounds that emphasize family identity and obligation, particularly in American society. Personal agency, and the associated sense of value and respect such agency provides for developing adolescents, is likely to be as critical to the development of a sense of family identity as is the emphasis upon group solidarity and togetherness. The dynamics of personal agency and group solidarity are essential issues in human relations in general and in adolescent development specifically, and they likely represent a fundamental aspect of ethnic minority family relationships in American society. A social identity approach could be a fruitful way to study these dynamics, as will be discussed in Section IV.C of this chapter.

C. INTERNALIZATION OF VALUES

The fourth and final social identity prediction about ethnic differences in the family is that adolescents from ethnic minority families would show a greater internalization of values than would be predicted on the basis of their relationships within the family. Truly testing this prediction requires studies that independently measure the values of different family members and examine whether the concordance between these values are greater in ethnic minority families than would be predicted from their dyadic relationships within the family. Few studies that include all of these elements exist. One study assessed both parents' and adolescents' values about familism among an ethnically diverse sample and found that whereas parents reported higher values than adolescents, few ethnic differences in the discrepancy between parents' and adolescents values emerged (Phinney, Ong, & Madden, 2000). The nature of the dyadic relationships between adolescents and parents, however, were not reported. It is also unclear whether the correlations between parents' and adolescents' values varied across the different ethnic groups, which is a more direct assessment of adolescents' internalization of family values. Nevertheless, the use of both parents and adolescents in the Phinney *et al.* study represents the methodological advance the needs to be included more often in studies of ethnic minority families.

Although more research that assesses ethnic variations in parent–adolescent value similarity needs to be done, evidence from other work suggests that adolescents from ethnic minority families report stronger values in some areas than would be predicted from their relationships within the family. For example, adolescents from ethnic minority backgrounds generally have stronger values of education as compared to those from European American backgrounds, even though their parents use higher levels of controlling and authoritarian parenting that is generally associated with weaker values of education and lower levels of internalization in general (Fuligni, 2001). Studies have observed that Asian American parents, in particular, use more controlling parenting practices than European American parents (Chao, 2001). Differences between African American and Latino parents and their European American counterparts are somewhat less consistent, but when differences are observed, they are generally in the direction of more controlling and power assertive styles among ethnic minority than majority parents (Steinberg, 2001).

Despite these differences in parenting practices, adolescents from ethnic minority backgrounds often report values of education that are equal to or even greater than the values of their European American peers, even though power assertive parenting styles are usually associated with a lower level of value internalization. One explanation for these findings is the idea that controlling parenting styles have different meanings within the cultural backgrounds of many ethnic minority families. For example, Chao (1994) has suggested that controlling parenting within Chinese families does not have the same negative connotation as it does within European American families because a controlling parent is more culturally acceptable in Chinese traditions that more strongly emphasize filial piety and hierarchical relations within families. Alternatively, parents may differentially exert control across domains. For example, the finding that Asian American adolescents are granted more personal autonomy in academic endeavors is consistent with their achievement in this realm (Asakawa & Csikszentmihalyi, 1998). Another explanation focuses on the fact that ethnic minority families are more likely to live in neighborhoods that present many dangers to their children, thereby making greater parental control over adolescents' lives a sensible adaptive strategy (McLoyd *et al.*, 2000). These explanations share an emphasis upon understanding the specific traditions and life circumstances of ethnic minority families.

A social identity approach to the findings regarding family relationships and internalization in ethnic minority families is complementary to these explanations, and suggests that the cultural and contemporary contexts surrounding these controlling parenting practices do not compromise and perhaps even enhance adolescents' identification with the family. If controlling parenting practices are indeed more culturally acceptable within Chinese and other ethnic minority families, then they may not compromise the adolescents' sense of value and respect in their families. Additionally, although controlling practices are evident in Asian American families, a domain of personal agency remains available to adolescents within these families that could provide the sense of value and respect that is necessary for group identification (Fuligni, 1998; Smetana, 2002). The use of control in response to perceived threat and danger, which has been observed in African American and other ethnic minority families, likely sends adolescents a different message than seemingly capricious and arbitrary assertions of parental authority. The message may be one that emphasizes families' response to a shared threat, which could actually enhance adolescents' sense of identification. Both of these examples of the context in which parental control occurs within ethnic minority families suggest that the internalization of values within ethnic minority families are greater than would be predicted by family relationships because of a greater social identification with the family.

Other evidence in support of the idea that family identification plays an important role in the internalization of values in ethnic minority families comes from the links between family obligation and academic motivation that we have observed in our studies of ethnically diverse teenagers (Fuligni, 2001; Fuligni & Tseng, 1999). In addition to assessing adolescents' sense of obligation to the family, we have measured many different values of adolescents regarding academic achievement and educational attainment. To tap the students' *value of academic success*, adolescents were asked to rate the importance of outcomes such as "Doing well in school", "Being one of the best students in your class", and "Going to college after high school." Adolescents' perceptions of the *future utility of education* was measured by having students respond to statements such as "Going to college is necessary for what I want to do in the future", "I need to get good grades in school to get a good job as an adult", and "Doing well in school is the best way for me to succeed as an adult." We also assessed adolescents' subject-specific values of English and mathematics. Correlational analyses were conducted in order to determine whether a sense of family obligation was related to youth's academic motivation. As shown in Table II, the correlations demonstrate a notable and consistent link between an emphasis upon assistance to the family and a value of academic achievement. Adolescents who believed that they should assist, support, and respect their family placed a stronger value upon achieving a measure of academic success and going on to college. These youths also had a stronger belief in the utility and importance of schooling, English, and mathematics for their future lives and occupations as adults.

We also examined whether the adolescents' educational values varied across ethnic backgrounds (Fuligni, 2001). As shown in Table III, Asian adolescents—those with Chinese and Filipino backgrounds—consistently report the highest level of academic motivation. These youths place more importance upon succeeding in school, going on to college, and have stronger faith in the importance and utility of education for their adult lives as compared to those from European backgrounds. The differences between the Latin American and European American students are less consistent, but the Mexican and Central/

TABLE II

Correlations between Family Obligation and Academic Attitudes

	Current assistance (*r*)	Respect for family (*r*)	Future support (*r*)
Math value	0.25	0.34	0.25
English value	0.20	0.17	0.14
Value of academic success	0.33	0.40	0.28
Future utility of education	0.26	0.33	0.22

Note: Table adapted from figures presented in Fuligni and Tseng (1999). All correlations significant at $p < 0.001$.

TABLE III

Ethnic Differences in Adolescents' Academic Attitudes

	Chinese (*M*)	Filipino (*M*)	Mexican (*M*)	C/S American (*M*)	European (*M*)
Math value	4.07*	4.08*	3.91*	4.02*	3.57
English value	4.49*	4.30*	3.89	4.13	3.96
Value of academic success	4.17*	4.09*	3.56	3.76*	3.45
Utility of education	4.59*	4.53*	4.17*	4.26*	3.99

Note: Table adapted from results presented in Fuligni (2001). All values were measured on a 5-point scale, where a higher score indicates a stronger value.

*indicates whether a group mean is significantly different from the European group mean at $p < 0.05$ or less.

South American students do indicate a value of academic success and a belief in the utility of education that is either equal to or greater than their peers from European backgrounds. The latter differences are especially notable because Latino adolescents receive significantly lower grades in school, indicating these students place a higher value on academic success than their equally achieving European American peers. Indeed, when we control for students' school performance, all of the ethnic minority students—both Asian American and Latino—have stronger values of education. Such a finding suggests that Latino adolescents, perhaps because of the greater economic and social challenges that they experience, require a higher level of motivation in order to achieve at the same level as their European American peers.

Finally, mediation analyses indicated that the tendency for adolescents from Asian and Latin American backgrounds to place more emphasis upon the importance and utility of education as compared to their equally achieving European American peers is associated with a sense of duty to support and assist the family. When we control for the youths' sense of family obligation, the ethnic differences in motivation become significantly reduced (Fuligni, 2001). This significant statistical mediation suggests that the desire for Asian and Latin American students to support, assist, and respect their families leads these youths to place more value upon the importance and usefulness of education than their European American peers who are achieving a similar level of performance in school.

These results are suggestive of the idea that the importance of family as a social identity among adolescents from ethnic minority families results in a greater internalization of values than would be predicted from their relationships within the family. The findings are by no means definitive, however, and continued research into the impact of family identity on adolescents' internalization of values represents one of the new and continuing directions for research offered by a social identity approach to the family.

IV. New and Continuing Areas of Inquiry

In addition to enhancing the understanding of existing research, a social identity approach to ethnic differences in families suggests areas of work that would be fruitful to begin or continue. Specifically, it would be important to examine how children and adolescents develop a sense of family identity, the various meanings associated with family identity across ethnic and cultural groups, and the role of family identity as a socializing agent in adolescent development.

A. EMERGING UNDERSTANDING OF FAMILY AS A SOCIAL IDENTITY

Because family membership serves as an important social identity for adolescents from ethnic minority families, we need to better understand how such an identity develops and is acquired. Answering this question requires going beyond examining developmental changes in the dyadic relationships between adolescents and their parents, and beginning to focus on adolescents' ideas about their membership in and identification with the larger family. It also requires assessing the extent to which adolescents feel a sense of "we-ness" with the family that goes beyond their relationships with individual family members. Numerous measures and scales that assess identification with other social groups and categories such as gender and ethnicity currently exist, and attempts could be made to adapt such measures for assessing family identification. In addition, behavioral measures could be used to concentrate on the potential behavioral manifestations of family identification, such as family assistance and support. Finally, several ethnographies have been conducted on family identity, solidarity, and obligation among ethnic minority families (e.g., Burton & Lawson Clark, in press; Gibson & Bhachu, 1991; Suárez-Orozco & Suárez-Orozco, 1995; Zhou & Bankston, 1998), and it would be enriching to use such qualitative and mixed-method techniques to focus specifically on the question of how children come to understand and make sense of their family membership during the years of adolescence.

Research on the development of adolescents' sense of family identity would profit from focusing on three particular issues: the conditions that heighten and enhance family as a social identity, the manner in which families negotiate what it takes to be a valued family member, and the implications of family identity for other important social identities during adolescence. In terms of the conditions that enhance family identification among youths from ethnic minority families, we already have highlighted the importance of the functional use of ethnicity and the social threat experienced by ethnic minorities. New research could focus on the links between children's emerging understanding of these social phenomena and the extent to which they identify with their families. When and how do children begin to understand the associations between the distribution of resources, social threats, and ethnicity and how does that understanding relate to how they identify about their families? What are the most salient aspects of functional use and social threat for developing children and adolescents? In addition to these factors, research could focus on issues such as socioeconomic resources, family emergencies and distress, and the neighborhood ethnic and economic contexts. Studies of the role of these social factors in family identification should be sure to separately assess the actual existence of these factors, adolescents' understanding and interpretation of the factors, and the extent to which adolescents link the social conditions facing their family with their identification and sense of "we-ness" with the larger family.

Because adolescence is a developmental period that involves the transformation of family relationships to accommodate the increasingly mature and competent child, an assessment of the development of family identity requires a focus on how families and adolescents negotiate what it takes to be a good member of the family. According to social identity theory and many supportive studies, group identification is enhanced when members feel like they are valued members of the group (Tyler, 1999). Therefore, it is important to understand the criteria by which family members determine whether adolescents are valued members of the family. Parents likely initially dictate the prevailing norms and expectations for children in the family, such as doing well in school, staying out of trouble, and helping other members of the family. Adolescents tend to generally agree with the prevailing norms of the family, but they and their parents often must negotiate the specific details of these norms as they apply to the adolescents' daily lives. For example, what qualifies as doing well in school? Does staying out late on weekends count as getting into trouble, or is that acceptable as long as the adolescent does not get involved with the police? A particularly salient issue for ethnic minority families, given the importance of family obligation and assistance, is the question of what counts as helping family members. Do adolescents fulfill their family obligations by simply doing well in school and staying out of trouble, with the understanding that following a straight path helps and brings honor to the family? If family obligation requires direct assistance to family members, how much is enough? Does helping siblings with homework fulfill one's obligations, or does being a good family member require more substantial household chores or even helping parents at work? The specific criteria by which being a good family member is determined likely involves a degree of give and take between adolescents and their parents, particularly for teenagers with immigrant parents who have traditional cultural expectations that may conflict with the norms of being a teenager in American society (Phinney, Ong, & Madden, 2000). The manner in which adolescents try to resolve such conflicts and more generally attempt to have input into what it takes to be a good member of their family would be important to study, as it should have significant bearing on the extent to which family membership will function as an important social identity for the adolescent.

Finally, research should address how the development of family identity takes place alongside the development of other social identities, such as gender and ethnic identity. One assumption behind the social identity approach to ethnic differences in family relationships that we outlined in this chapter is that adolescents' understanding of their ethnic group membership is linked to their sense of identification with the family. Family and ethnic identity, therefore, should be associated with one another, particularly among adolescents from ethnic minority backgrounds. At the same time, family identity should be linked to ethnic identity and gender identity if, for example, being a "good Chinese child" or a "good Mexican daughter" is an important expectation within the family. To the extent that being a valued family member is linked to being "good" representatives of other social categories such as ethnic group or gender, then family identity should be linked to these other social identities. In contrast, if other important social identities for the adolescent conflict with being a good family member, then either those social identities or the family identity may be diminished. Or, the adolescent may engage in creative ways to keep both identities, either by separating the two as much as possible or by trying to redefine the identities so that they do not conflict with one another. For example, an adolescent who values his identity as an athlete may try to avoid conflicting with the high academic expectations of his family by studying late into the night after practice or by casting his athletic pursuits as a way to make him more attractive to competitive colleges. Adolescents often creatively try to balance the multiple worlds in their lives, often with success but sometimes with difficulty (Cooper *et al.*, 2002; Phelan, Davidson, & Yu, 1998). On the one hand, some adolescents may balance an obligation to spend time with the family with their desire to socialize with peers by inviting their friends over to their house for meals or special occasions. On the other hand, practical family demands may also shape adolescents' choices such that the necessity for childcare may impede on adolescents' time with friends or attention to academics. In one of our studies, the potential conflict between family and school demands was described by a ninth grade student from a Mexican immigrant family:

> Yeah, sometimes I get irritated and frustrated about the fact that I have to sit late at night. Sometimes during the weekday, they [her parents] would go late at night to Wal-Mart or something, or to the market because they wouldn't have time during the day. So, she [her mother] leaves it up to me to watch my little brother or sister. Sometimes I have a lot of homework so I tell her I have homework and she says, "Oh, you have to watch your brother and sisters." I wind up staying up really late or sometimes I wind up finishing it in class.

Whether successful or unsuccessful, these negotiations should have important implications for the development of both family identity and other critical social identities in adolescents' lives.

B. MEANING OF FAMILY IDENTITY

As discussed earlier, it remains to be determined whether adolescents from ethnic minority groups have a greater generic identification with the family than those from European American backgrounds, or whether the nature of adolescents' family identify differs across ethnic groups. In other words, family membership may be important for all youths, regardless of ethnic background, given the primary importance of families for child and adolescent development. The meaning, values, and norms associated with family identity may be what distinguish families of different ethnic backgrounds from one another. Social identity theory and associated approaches to social identity suggest that it is just as important to determine the defining characteristics of social groups as it is to analyze individuals' degree of identification with those groups (Ashmore, Deaux, & McLaughlin-Volpe, 2004). The same should be true for families from different ethnic backgrounds. A family identity could have very different implications for adolescents' development depending upon how that identity is defined in different ethnic groups.

An example of the potentially different meaning of family identity across different ethnic groups was discussed earlier in terms of the importance placed upon family obligation and assistance among families from ethnic minority and immigrant backgrounds. Within these families, the meaning of being a good family member is associated with a collection of values and norms that emphasize the role of children and adolescents in assisting and supporting other family members. Family obligation also often implies the necessity of taking into account the needs and wishes of the family when making decisions and choices about activities, leisure time, and future educational and occupational plans. In other families, being a member of the family may not imply these types of obligations. Identification with the family may still be strong, but it just does not necessitate the kinds of responsibilities that exist in many ethnic minority and immigrant families.

Research on the role of family identity, therefore, should assess the meaning along with the level of identification among families from different ethnic and cultural backgrounds. Research on the meaning of family identification may be done in several ways. Investigators could use open-ended and qualitative interview techniques to tap adolescents' ideas about what it means to be good and valued members of their families. Adolescents could be asked what values and norms are most important to their families, and which tend to distinguish their families from families with different backgrounds. These values and norms could be probed more deeply by developing scenarios in which such values and norms are violated, and asking adolescents how such transgressions would be viewed within the family. Are the adolescents punished? If so, how severely would they be punished? Could the transgressions result in shaming or ostracism from the family? Collectively, these types of probes could yield defining characteristics of the family with valuable information about the family's "bottom line"—that is, the value or norm that simply cannot be violated in order to be considered a good member of the family.

An additional way to examine the meaning of family identity in different ethnic groups is to examine how identification is associated with other aspects of adolescent development. Investigators could study the implications of family identification for things such as family assistance, educational values and behaviors, religious beliefs, anti-social and pro-social behavior, and political and social attitudes and behaviors. An association between family identity and certain beliefs, such as those towards sexuality or morality, would suggest that these beliefs help to define the meaning of family identity. Likewise, a link between greater educational values and family identity would indicate that believing in the importance of educational effort and success is part of being a good member of the family. One interesting approach would be to examine whether the association between family identity and other aspects of adolescent development varies across different ethnic groups. For example, an association between family identification and religiosity that exists in some groups but not others would be the evidence that the meaning of family identity varies across groups, implying certain religious beliefs in some groups but not others. The potential for such a pattern of findings highlights the importance of assessing both the level and the meaning of family identification among adolescents from different ethnic groups.

C. FAMILY IDENTITY AS A SOCIALIZING AGENT

The tendency for social identification to be associated with greater internalization of group values suggests that family identity can be an important socializing force in adolescent development. In particular, it can be a potent yet seemingly indirect socializing agent of adolescents' values and behaviors because heightening an adolescent's sense of identification with the family should lead adolescents to seek out the values of the family and endeavor to become a valued and respected member of the family (Tyler, 1999). It would be important, therefore, to not only examine how specific values and behaviors are socialized within the family through direct messages, supervision, and regulations. A social identity approach would suggest that investigators also should pay attention to how families emphasize and enhance identification with the family as a social group, and the extent to which that identification is associated with the acquisition of desired values and behaviors.

Among ethnic minority families, the practices associated with what has been called ethnic or racial socialization involve several techniques known to heighten group identification. As children become adolescents, ethnic minority parents increasingly offer their children a collection of messages involving awareness of racism, knowledge of cultural history and practices, and ethnic pride (Hughes & Chen, 1999). These messages are sometimes delivered in response to an adolescent's negative racial experience or in a pro-active manner in order to prepare the adolescents for such experiences in the future. Interestingly, the aspect of racial socialization that deals with making children aware of racism essentially involves teaching children about the ways in which race and ethnicity are associated with the distribution of resources and social threat in society, social factors that are related to stronger group identification. Teaching children about the family's cultural history, which often involves family-related themes, serves to heighten the distinctiveness of the adolescent's family and ethnic group. Finally, messages about ethnic pride serve to maintain and even heighten adolescents' identification with the group while the adolescents are beginning to learn that their social group may not be particularly valued or respected in the larger society. The relevance of ethnic and racial socialization for family identification is that parents likely invoke themes of family solidarity and "we-ness" when delivering these messages, resulting in adolescents making the connection between their memberships in their ethnic group and their family.

Given the potential implications of ethnic and racial socialization for family identification, it would be interesting to see how the different aspects of ethnic and racial socialization identified by Hughes (2003), such as preparation for bias and ethnic pride, interact with one another to produce family identification. For example, preparation for bias may produce a stronger identification with the family only in the context of messages about ethnic pride. Developmentally, it would also be profitable to examine these associations dynamically over time. Some aspects of racial socialization create ambivalence for ethnic minority parents because parents want their children to be optimistic about their futures, but also to be aware of the challenges that society will present to them because of their ethnic or racial background (Hughes & Chen, 1999). Examining the role of family identification in this aspect of socialization in ethnic minority families could offer insights into how these complex messages are received and understood by the adolescents themselves.

It would be informative to examine the implications of other aspects of parenting and family relationships for adolescents' identification with the family. The extent to which family decisions are made jointly, taking into account either the wishes or the needs of the entire family, could have implications for adolescents' family identity. Discourse in the family could be analyzed for the frequency with which language involves references to "we" as a family group, as opposed to individual family members. The level at which children's contributions to the family, either through their opinions or instrumental contributions such as chores or achievements, are welcomed and valued by the larger family should also have implications for family identification. In general, research should pay attention to family practices that tend to highlight family distinctiveness and pride, however subtle the practices may be, as well as actions that allow adolescents to believe that they are valued and respected members of the group.

Finally, the manner in which these family practices socialize children in seemingly unrelated family values and beliefs would be worthwhile to examine. Are adolescents who experience these types of parenting techniques and family relationships more likely to seek out and internalize family values? Are they more likely to actively develop their beliefs and behaviors so that they are in accord with the goals of the parents and the family? What is particularly promising about examining the implications of family identity as a socializing agent is that such an approach considers adolescents active self-socializers. Children are more likely to adopt values and beliefs that they feel they actively chose (Grolnick, Deci, & Ryan, 1997), and identity-based self-socialization involves active choice on the part of the adolescents. In other words, adolescents who more strongly identify with their families are more likely to want to be valued and respected members of that group, and they will actively engage in a process by which they try to act in accord with the goals and values of the group. It would be interesting to see whether values and beliefs adopted through this socialization route would be more stable and richly developed as compared to values and beliefs that were directly taught by parents and more passively received by the adolescents themselves.

V. Conclusion

The surge in research on family relationships during adolescence in the past 25 years recently has begun to extend to the study of ethnic minority families. The focus on families from different ethnic, generational, and cultural backgrounds needs to continue given the demographic changes taking place in the United States in which ethnic minority children increasingly comprise larger proportions of the American population. As the number of studies of adolescent development within ethnic minority families rises, the need for different theoretical and methodological approaches also increases. We believe that understanding ethnic differences in family relationships requires going beyond the traditional emphasis on dyadic relationships and focusing on the extent to which families serve as particularly important and salient social identities for adolescents from ethnic minority backgrounds. By no means do we mean that dyadic family relationships are unimportant for ethnic minority adolescents during the years of adolescence. Rather, in addition to focusing on dyadic relationships, researchers need to consider ways in which the social status of ethnic minority groups may shape the roles families play in children's lives. We believe that a social identity approach to the family offers a promising lens through which to view the importance of families for ethnic minority adolescents. A social identity approach cannot account for all of the ways in which ethnicity plays a role in family relationships, but such a focus could help to integrate findings that have emerged in recent years as well as generate new and potentially fruitful questions as we move into the next quarter century of research on adolescent development within the family.

ACKNOWLEDGEMENTS

The authors would like to thank members of the Russell Sage Foundation Working Group on Social Identity and the Foundation for Psychocultural Research-UCLA CBD Interdisciplinary Research Group on Adolescence for their feedback on the ideas expressed in this chapter. Jacqueline Goodnow, Sandra Graham, and Diane Ruble offered helpful suggestions that were integrated into our discussion. Support during the preparation of this chapter was provided to the first author by the Russell Sage Foundation and the William T. Grant Foundation and to the second author by the National Science Foundation.

REFERENCES

Asakawa, K., & Csikszentmihalyi, M. (1998). The quality of experience of Asian American adolescents in activities related to future goals. *Journal of Youth and Adolescence*, *27*, 141–163.

Ashmore, R. D., Deaux, K., & McLaughlin-Volpe, T. (2004). An organizing framework for collective identity: Articulation and significance of multidimensionality. *Psychological Bulletin*, *130*, 80–114.

Azmitia, M., Cooper, C. R., Garcia, E. E., & Dunbar, N. D. (1996). The ecology of family guidance in low-income Mexican–American and European–American families. *Social Development*, *5*, 1–23.

Brown, B. B. (1990). Peer groups and peer cultures. In S. S. Feldman & G. R. Elliot (Eds.), *At the threshold: The developing adolescent* (pp. 171–196). Cambridge, MA: Harvard University Press.

Burton, L. M. & Lawson Clark, S. (in press). *House and homeplace in the lives of low-income African American families*. New York: The Guilford Press.

Chao, R. K. (1994). Beyond parental control and authoritarian parenting style: Understanding Chinese parenting through the cultural notion of training. *Child Development*, *65*, 1111–1120.

Chao, R. K. (2001). Extending research on the consequences of parenting style for Chinese Americans and European Americans. *Child Development*, *72*, 1832–1843.

Cooper, C. R., Baker, H., Polichar, D., & Welsh, M. (1993). Values and communication of Chinese, Filipino, European, Mexican, and Vietnamese American adolescents with their families and friends. In S. Shulman & A. W. Collins (Eds.), *Father–adolescent relationships* (pp. 73–89). San Francisco, CA: Jossey-Bass.

Cooper, C. R., Cooper, R. G. Jr., Azmitia, M., Chavira, G., & Gullatt, Y. (2002). Bridging multiple worlds: How African American and Latino youth in academic outreach programs navigate math pathways to college. *Applied Developmental Science*, *6*, 73–87.

Dornbusch, S. M., Glasgow, K. L., & Lin, I. C. (1996). The social structure of schooling. *Annual Review of Psychology*, *47*, 401–429.

Fisher, C. B., Wallace, S. A., & Fenton, R. E. (2000). Discrimination distress during adolescence. *Journal of Youth and Adolescence*, *29*, 679–695.

Freeberg, A. L., & Stein, C. H. (1996). Felt obligations towards parents in Mexican-American and Anglo-American young adults. *Journal of Social and Personal Relationships*, *13*, 457–471.

Fuligni, A. J. (1998). Parental authority, adolescent autonomy, and parent-adolescent relationships: A study of adolescents from Mexican, Chinese, Filipino, and European backgrounds. *Developmental Psychology*, *34*, 782–792.

Fuligni, A. J. (2001). Family obligation and the academic motivation of adolescents from Asian and Latin American, and European backgrounds. In A. Fuligni (Ed.), *Family obligation and assistance during adolescence: Contextual variations and developmental implication* (pp. 61–76). San Francisco, CA: Jossey-Bass.

Fuligni, A. J., Alvarez, J. M., Bachman, M., & Ruble, D. N. (in press). Family obligation and the academic motivation of young children from immigrant families. In C. R. Cooper, C. García Coll, T. Bartko, H. Davis, & C. Chatman (Eds.), *Hills of gold: Rethinking diversity and contexts as resources for children's developmental pathways*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Fuligni, A. J., & Tseng, V. (1999). Family obligations and the achievement motivation of children from immigrant and American-born families. In T. Urdan (Ed.), *Advances in motivation and achievement* (pp. 159–184). Stamford, CT: JAI Press.

Fuligni, A. J., Tseng, V., & Lam, M. (1999). Attitudes toward family obligations among American adolescents from Asian, Latin American, and European backgrounds. *Child Development*, *70*, 1030–1044.

García-Coll, C., Crnic, K., Lambert, G., Wasik, B. H., Jenkins, R., Vázquez García, H., & McAdoo, H.P. (1996). An integrative model for the study of developmental competencies in minority children. *Child Development*, *67*, 1891–1914.

Gaines, S. O. Jr., Marelich, W. D., Bledsoe, K. L., Steers, W. N., Henderson, M. C., Granrose, C. S., Barájas, L., Hichs, D., Lyde, M., Takahashi, Y., Yum, N., Ríos, D. I., García, B. F., Farris, K. R., & Page, M. S. (1997). Links between race/ethnicity and cultural values as mediated by racial/ethnic identity and moderated by gender. *Journal of Personality and Social Psychology*, *72*, 1460–1476.

Gibson, M. A., & Bhachu, P. K. (1991). The dynamics of educational decision making: A comparative study of Sikhs in Britain and the United States. In M. A. Gibson & J. U. Ogbu (Eds.), *Minority status and schooling: A comparative study of immigrant and involuntary minorities* (pp. 63–96). New York: Garland Publishing.

Grolnick, W. S., Deci, E. L., & Ryan, R. M. (1997). Internalization within the family: The self-determination theory perspective. In J. E. Grusec & L. Kuczynski (Eds.), *Parenting and children's internalization of values* (pp. 135–162). New York, NY: Wiley.

Harrison, A. O., Wilson, M. N., Pine, C. J., Chan, S. Q., & Buriel, R. (1990). Family ecologies of ethnic minority children. *Child Development*, *61*, 347–362.

Harter, S. (1990). Self and identity development. In S. S. Feldman & G. R. Elliot (Eds.), *At the threshold: The developing adolescent* (pp. 352–387). Cambridge, MA: Harvard University Press.

Ho, D. (1996). Filial piety and its psychological consequences. In M. H. Bond (Ed.), *The handbook of Chinese psychology* (pp. 155–165). Hong Kong: Oxford University Press.

Hogg, M. A. (2003). Social identity. In M. R. Leary & J. P. Tangney (Eds.), *Handbook of self and identity* (pp. 462–479). New York, NY: Guilford Press.

Hughes, D. (2003). Correlates of African American and Latino parents messages to children about ethnicity and race: A comparative study of racial socialization. *American Journal of Community Psychology*, *31*, 15–33.

Hughes, D., & Chen, L. (1999). The nature of parents' race-related communications to children: A developmental perspective. In L. Balter & C. S. Tamis-LeMonda (Eds.), *Child psychology: A handbook of contemporary issues* (pp. 467–490). Philadelphia, PA: Psychology Press.

Janoff-Bulman, R., & Leggatt, H. K. (2002). Culture and social obligation: When "shoulds" are perceived as "wants". *Journal of Research in Personality*, *36*, 260–270.

Joe, J. R., & Malach, R. S. (1998). Families with native American roots. In E. W. Lynch & M. J. Hanson (Eds.), *Developing cross-cultural competence* (2nd ed., pp. 127–158). Baltimore, MD: Brookes Publishing.

Matute-Bianchi, M. E. (1991). Situational ethnicity and patterns of school performance among immigrant and non-immigrant Mexican-descent students. In M. A. Gibson & J. U. Ogbu (Eds.), *Minority status and schooling: A comparative study of immigrant and involuntary minorities* (pp. 205–248). New York: Garland Publishing.

McLoyd, V. C., Cauce, A. M., Takeuchi, D., & Wilson, L. (2000). Marital processes and parental socialization in families of color: A decade review of research. *Journal of Marriage and the Family*, *62*, 1070–1093.

Oakes, J., Selvin, M. J., Karoly, L., & Guiton, G. (1991). *Educational matchmaking: Toward a better understanding of curriculum and tracking decisions*. Santa Monica, CA: Rand.

Phelan, P., Davidson, A. L., & Yu, H. C. (1998). *Adolescents' worlds: Negotiating family, peers, and school*. New York, NY: Teachers College Press.

Phinney, J. S. (1990). Ethnic identity in adolescents and adults: Review of research. *Psychological Bulletin*, *108*, 499–514.

Phinney, J. S., & Chavira, V. (1995). Parental ethnic socialization and adolescent coping with problems related to ethnicity. *Journal of Research on Adolescence*, *5*, 31–53.

Phinney, J. S., Ong, A., & Madden, T. (2000). Cultural values and intergenerational value discrepancies in immigrant and non-immigrant families. *Child Development*, *71*, 528–539.

Ruble, D. N., Alvarez, J., Bachman, M., Cameron, J., Fuligni, A. J., García-Coll, C., & Rhee, E. (2004). The development of a sense of "We": The emergence and implications of children's collective identity. In M. Bennett & F. Sani (Eds.), *The development of the social self*. East Sussex, England: Psychology Press.

Sabogal, F., Marin, G., Otero-Sabogal, R., VanOss Marin, B., & Perez-Stable, E. J. (1987). Hispanic familism and acculturation: What changes and what doesn't? *Hispanic Journal of Behavioral Sciences*, *9*, 397–412.

Schreiber, J. B., & Chambers, E. A. (2002). After-school pursuits, ethnicity, and achievement, for 8th- and 10th-grade students. *Journal of Educational Research*, *96*, 90–100.

Shrum, W., Cheek, N. Jr., & Hunter, S. (1988). Friendship in school: Gender and racial homophily. *Sociology of Education*, *61*, 227–239.

Smetana, J. G. (2002). Culture, autonomy, and personal jurisdiction in adolescent–parent relationships. In R. V. Kail & H. W. Reese (Eds.), *Advances in child development and behavior* (Vol. 29, pp. 51–87). San Diego, CA: Academic Press.

Steinberg, L. (2001). We know some things: Parent–adolescent relationships in retrospect and prospect. *Journal of Research on Adolescence*, *11*, 1–19.

Szalacha, L. A., Erkut, S., Garcia Coll, C., Alarcon, O., Fields, J. P., & Ceder, I. (2003). Discrimination and Puerto Rican children's and adolescent's mental health. *Cultural diversity and ethnic minority psychology*, *9*, 141–155.

Suárez-Orozco, C., & Suárez-Orozco, M. M. (1995). *Transformations: Immigration, family life, and achievement motivation among Latino adolescents*. Stanford, CA: Stanford University Press.

Tajfel, H. (1972). Social categorization. In S. Moscovici (Ed.), *Introduction a la psychologie sociale* (Vol. 1, pp. 272–302). Paris: Larousse.

Tajfel, H., & Turner, J. C. (1979). An integrative theory of intergroup conflict. In W. G. Austin & S. Worchel (Eds.), *The social psychology of intergroup relations* (pp. 33–47). Monterey, CA: Brooks/Cole.

Tyler, T. (1999). Why people cooperate with organizations: An identity-based perspective. *Research in Organizational Behavior*, *21*, 201–246.

Tyler, T., Degoey, P., & Smith, H. J. (1996). Understanding why fairness of group procedures matters: A test of the psychological dynamics of the group-value model. *Journal of Personality and Social Pscyhology*, *70*, 913–930.

Waters, M. C. (1999). *Black identities: West Indian immigrant dreams and American realities*. New York: Russell Sage.

Way, N. (1998). *Everyday courage: The lives and stories of urban teenagers*. New York: New York University Press.

Zhou, M., & Bankston, C. L. (1998). *Growing up American: How Vietnamese children adapt to life in the United States*. New York: Russell Sage Foundation.

WHAT DEVELOPS IN LANGUAGE DEVELOPMENT?

LouAnn Gerken

DEPARTMENTS OF PSYCHOLOGY AND LINGUISTICS, UNIVERSITY OF ARIZONA,
TUCSON, AZ 85721, USA

I. Introduction

Human language exhibits a fascinating duality. On the one hand, nearly every member of our species develops language to a greater or lesser extent, and ours

Advances in Child Development and Behavior
R Kail (Editor)

is the only species that has a communication system with the informational power of human language. On the other hand, no member of our species is born with an overt language. Rather, we come to master the language or languages of our community over the first few years of life. Thus, the dual nature of language is that it is both a species characteristic and a skill acquired by individual humans. One measure by which theories of human language can be differentiated is their emphasis on species vs. individual aspects of the phenomenon. In this chapter, I consider mainly the question of how language might develop in an individual learner. However, any adequate theory of language development must also take a position on how language development in individuals is related to linguistic ability in our species. Therefore, I also highlight a subset of species characteristics that we need to bear in mind as we attempt to answer what in fact does development in language development.

Only in the second half of the 20th century was the process by which individual humans develop a language acknowledged as a mystery worthy of scientific investigation. Because children gradually come to produce and comprehend utterances "like" the ones to which they are exposed, the potential complexity of the acquisition process is obscured. The mystery hinges on the word "like", which lets slip the fact that children produce utterances that they have never heard before. The creativity of children's utterances is masked to most observers, because these utterances contain words and structures that are part of the language of the community. The novelty of children's combinations only becomes apparent when these deviate in some way from the adult language, as when a child produces *goed* instead of *went*. Thus, a simple answer to the question, what develops in language development, is the ability to generalize from the utterances encountered in the input to novel utterances that are consistent with those of the language community. The nature of the generalization process is the central focus of this chapter.

The data that might bear on the question of what develops in language development can be organized in a variety of ways. One would be by language domain—phonology, syntax, the lexicon (e.g., Gerken, 2002). However, children do not necessarily divide the subject of learning a language into the same domains as language researchers do. An organization by language domain can, therefore, easily misses aspects of development that are common across domains. Another possible organization might be in terms of theories of language development. However, as noted previously, theories differ on how much emphasis they place on the development of language in individuals vs. language as a species characteristic. These differences of emphasis often make comparing entire theories extremely frustrating, because they seldom make sufficiently precise predictions about the same types of data.

The organizational device I have chosen for this chapter is a series of four questions that any adequate theory of language development must eventually answer.

Linguistic generativity—how creative are we?
What is the computational endowment of the language learner?
What constitutes evidence of generalization?
Are there constraints on what is learnable?

These four questions reflect the intersection of research on how language is acquired by individual humans with studies of language as a species characteristic. The first question concerns our human capacity to create language, in the absence of a model from which to learn. The second concerns what mental abilities are possessed by our species that allow individuals to acquire language and how these abilities are related to those found in non-humans. The third question concerns how we evaluate human language ability and what types of representational abilities we need to impute to human language learners. The fourth question concerns whether our species is constrained or biased to perceive or represent certain types of information and how such constraints or biases influence our capacity for language learning. For each question, I provide some background information about why it is important for ultimately answering the question of what develops in language development. I also provide a sample of data that bear on the question, and I end each section with an assessment of how close I believe the field is to answering it.

II. Linguistic Generativity—How Creative are We?

When we consider how learners might generalize from utterances they have encountered to new ones, we cannot escape noticing that language use in individuals is both creative and habitual. We can and do generate new utterances every day. You can hear a new count noun such as *fenisole* and be willing to bet your life savings that the plural is *fenisoles*, with the plural marker sounding like /z/ and not /s/. We also re-use utterances that we or someone else generated. But even beyond habitual social expressions like *how are you?* and *have a nice day*, we are discovering the habitual nature of language structure. Bever (1970) pointed out the sheer frequency in English of sentences composed of a noun phrase, a transitive verb, and a noun phrase (e.g., *the dog chased the cat*). Furthermore, not only are these structures frequent, but their interpretation is consistent: treat the first noun as the agent and the second noun and the patient/theme. Analyses of large corpora have revealed a wealth of such local patterns that have been shown to affect adult processing (e.g., Ellis, 2002; Ferreira, Bailey, & Ferraro, 2002; MacDonald, Pearlmutter, & Seidenberg, 1994).

Similarly, you may have never encountered *fenisole*, but you have encountered *sole* and know how it is pluralized. Such recurring patterns in language tempt many, especially non-linguists, to assume that what develops in human language is simply the ability to recognize and interpret a set of phonological and syntactic patterns. On such a view, language can be represented in probabilistic, distributed networks, without the need for discreet symbols or rules, such as S → NP VP (i.e., a sentence is composed of a noun phrase and a verb phrase) (Elman *et al.*, 1996; Lewis & Elman, 2001). Proponents of this view make the important point that much of our ability to comprehend and produce language does not require linguistic machinery like rules. Creativity, on this view, can be considered the result of substitution of one element for another element "like" it.

A. LINGUISTIC UNITS ACROSS LANGUAGES

However, the word "like" should again give us pause for thought. On what grounds do we determine that two linguistic elements (sounds, words, phrases) are sufficiently like each other that we should expect them to enter into the same linguistic pattern? Thus far, analyses that reveal frequent patterns in corpora, and computational models that generate new utterances consistent with a particular language, employ units of analysis supplied to them by linguistically savvy humans (e.g., phonological features, thematic roles, lexical classes). Although the units of analysis themselves might be discoverable through a large-scale statistical analysis over other, still more primitive units, no workable proposals of this kind have been made.

Furthermore, most linguists claim that the same units of analysis are useable across a variety of languages (e.g., Baker, 2001; Greenberg, 1963). If this claim is correct, an individual human learning a single language is either finding in that language, or bringing to it, core elements shared across languages learned by other humans. We must remember when evaluating the claim about universality of linguistic components that the goal of much of linguistic theory over the past 50 years has been to find linguistic universals (e.g., Chomsky, 1955). This goal, and the tools of analysis at our disposal, may have caused linguists to overestimate similarities across languages. For example, sounds transcribed with the same phonetic symbols across languages may have quite different articulatory/acoustic manifestations. Nevertheless, the degree of success linguists have achieved using basic linguistic units of analysis is remarkable. Any theory of language development must take a position on the nature of these putatively shared units. Either convincing arguments must be provided that these shared units are an illusion, or, if they are not, they must be explained in terms of a biological endowment for language, or in terms of a cognitive and/or pragmatic solution to the problem of information transmittal that is independently adopted in each language.

B. THE CREATION OF NEW LANGUAGES

Another source of data that bears on the linguistic creativity and biological endowment of humans comes from situations in which human learners receive little or no language input. In one study, researchers examined a property of the grammar of Nicaraguan Sign Language (NSL), as the language was being formed by two cohorts of deaf adults and children (Senghas, 2003; Senghas & Coppola, 2001). The members of the first cohort provided the input to the members of the second cohort. The marker in question was a spatial modulation of a basic verb (e.g., *pay*) to indicate the object of the verb. For example, a referent *man* might have been indicated as an arbitrary location in space (e.g., to the left of the signer), and to indicate that someone was paying the man, the verb *pay* would be made in the direction of the previously established referent. Members of the first cohort showed little use of this spatial verb marker. Adult members of the second cohort also showed little use of the marker. However, children of the second cohort substantially increased the use of the marker, but only in a set of contexts that were linguistically appropriate. These results are consistent with American deaf children, who regularize the inconsistent use of grammatical forms produced by parents who learn American Sign Language as adults after having a deaf infant (Ross & Newport, 1996; Singleton & Newport, in press).

Hudson and Newport (1999) explored the conditions under which language learners are most likely to make changes to their input. They exposed adults to an artificial language system in which nonsense nouns and verbs referred to objects and actions in an artificial world presented on a video display. The language contained four types of sentences: intransitive, transitive, negative intransitive, and negative transitive, presented with a optional negative marker followed by verb, subject, and an optional object. There were two determiners, and nouns were randomly assigned to occur with just one of them. Participants were assigned to one of four conditions, with consistency of determiner use in the input varying across conditions. Determiners were used either 45, 60, 75, or 100% of the time. In a sentence completion task, adults matched their production of determiners to the condition they were in. This finding suggests that adults do not readily systematize unsystematic input. In contrast, work by Hudson Kam and Newport (2005) shows that 5- to 7-year-old children, when faced with sentences in which a determiner was presented 60% of the time, regularized determiner use. Different children regularized usage differently, with some producing a determiner all of the time and others none of the time. These results suggest that children are more likely than adults to create a language system that is more systematic than their input.

The studies by Senghas, Newport and their colleagues all have in common the fact that learners take elements from their input that are not used systematically and make them systematic in ways consistent with existing languages. Senghas (2003) describes the situation in this way: "… this new version of the language is not unrelated to its model; it is derived from it. Forms that exist in free variation or with some other function in the language of the first cohort were available for the second cohort to use as raw materials for creating new form–function mappings. If the first stage were not necessary, all of NSL would have appeared in a single sweep, instead of being built cohort by cohort."

Other researchers have examined the situation when a deaf child has no language input at all and creates a gestural communication system called "home sign" (e.g., Feldman, Goldin-Meadow, & Gleitman, 1978; Goldin-Meadow & Mylander, 1998). Goldin-Meadow and Mylander (1998) report that four American and four Chinese deaf children, who were never exposed to a signed language, each showed a set of linguistic properties in their utterances that the authors convincingly argue cannot easily be attributed to the gestural input of their parents. First, they produced sequences expressing sentence-like content, as well as gestures for single words. Second, they produced complex sentences. Third, each child showed a pattern of production consistent with ergative languages, such as Inuit.[1] In sentences whose meaning involved only a single obligatory noun (e.g., *the <u>mouse</u> went in the hole*), children expressed the verb followed by the noun (*went mouse*). In sentences whose meaning involved two obligatory nouns (e.g., *the mouse ate the <u>cheese</u>*), children sometimes omitted one of the nouns, and they were consistent in retaining the patient (*cheese*), placing it after the verb. Agents, when they appeared, were place before the verb. Why every child adopted the same pattern is not clear; however, their systematic use of the same devices for mapping meaning to form is remarkable (e.g., Gleitman, 1990).

C. ASSESSMENT OF CREATIVITY

The similarity among existing languages, at least in terms of basic units of analysis, the human ability to not only learn, but create language, and the advantage of children over adults in language generalization suggest three related points. First, the machinery that we hypothesize for the task of language learning must, at the very least, be compatible with the creative function of language. Second, theories of language development must take a stand on the origin

[1]An ergative language is one that marks the subject of transitive verbs distinctly from the subject of intransitive verbs and the object of transitive verbs. In contrast, languages like English mark the subjects of transitive verbs and intransitive verbs the same way (e.g., the subject of *He ate fish* and *He ate* both appear in nominative case).

of linguistic units like phonetic features and syntactic categories, either taking them as innate primitives or demonstrating how they can be derived without tacitly assuming other linguistic knowledge on the part of the child. Third, the mechanism proposed to account for language development needs to also account for differences in linguistic creativity between children and adults. Some theories, mostly from the field of generative linguistics, have taken these phenomena as the central data requiring explanation. However, as I point out in subsequent sections, other important aspects of language development are better explained within other frameworks.

III. What is the Computational Endowment of the Language Learner?

The debate about whether linguistic generalizations are largely learned from exposure to input or innately given and simply triggered by specific input utterances has many facets. One of these concerns the assumptions that we can make about the computational abilities of the young language learner. In particular, we need to determine how much information about their input infants store, and what types of analytic tools they are able to apply to the input. A learner who stores only a few temporally contiguous utterances has a much less fruitful database to analyze than one who fully or partially stores a large number of utterances. The former learner is likely to need more "built in" generalizations than the latter. Similarly, a learner who notes only dependencies between adjacent units (e.g., syllables) has an impoverished tool-kit compared with a linguist, who looks for relations between non-adjacent units in single utterances, as well as patterns across utterances. The analytically impoverished learner needs more help from biology than the better equipped one.

A. STORING EXEMPLARS

Beginning with how much linguistic information human learners store, a variety of data bear on this question. In the domain of children's language production, work by Munson (2001) has demonstrated that children are more accurate and more consistent at producing nonsense words that contain phoneme sequences that are more frequent in the target language. Similarly, Zamuner and colleagues (Zamuner, 2003; Zamuner, Gerken, & Hammond, 2004) have demonstrated that children at the early stages of word production are better able to produce the same word-final consonant when it is part of a phoneme sequence that is more frequent in English. Such studies demonstrate that learners are able to keep track of the frequency of phoneme sequences in the target language. However, the frequency effect in these studies might be one of motor practice instead of frequency calculations over stored utterances. That is, children have more frequently attempted to produce words with the more frequent patterns.

The motor interpretation of frequency effects is ruled out by studies with prelinguistic infants, who also distinguish between phoneme sequences that are frequent vs. infrequent in their language (Jusczyk, Luce, & Charles-Luce, 1994; Gerken & Zamuner, in press). Because most studies examining segment frequency effects have calculated frequency over words in the lexicon (lexical types, e.g., Beckman & Edwards, 2000; Munson, 2001), we might infer that infants calculate the frequency of segment sequences across their growing lexicons. However, Zamuner and Gerken found the effect in 7.5-month-olds, who have very small lexicons at best. These data suggest that infants perform segment sequence frequency calculations over word or utterance tokens, implying that they store a large number of these tokens.

Further evidence that infants store words as individual exemplars/tokens and not lexical types can be found in the work of Houston and Jusczyk (2000, 2003). They found that 7.5-month-olds, who were familiarized with a passage produced by one female talker, were able to recognize particular words from the passage if they were produced by a different female talker, but not a male talker. In contrast, 10.5-month-olds recognized the words regardless of talker differences. Taken together, these data suggest that 7.5-month-old infants are storing utterance tokens at least the size of a word. Because words are ultimately stored (as lexical types), the storage of word-sized units is impressive, but perhaps not surprising.

Is there evidence that infants also store tokens of larger utterances, such as entire phrases or sentences? This question is an important one, because adults exposed to an artificial grammar are best able to learn aspects of the language when exposed to the same set of words participating in different sentence structures (Morgan, Meier, & Newport, 1989). In one developmental study that bears on this question, researchers exposed 9-month-olds to sentences exhibiting either alternating structures (Subject–Verb–NP and Subject–Verb–Subordinate Clause) or the same structure (all S–V–NP or S–V–S). Infants were then tested using the headturn preference procedure (Kemler Nelson *et al.*, 1995) on sentences with the same structure(s) they had heard during training and containing either the same words or different words than in training. Listening times to the condition with alternating structures and the same words heard in training were longest (Jusczyk & Kemler Nelson, 1996). If infants had stored only words during training, they should have found test items with the same words as the training items equally interesting, regardless of whether these words occurred in the same or different sentence structures. Therefore, infants' preference for test utterances with the same words but different structures from training suggests that they stored entire utterances during training and were able to compare test utterances to the stored training utterances.

This study, although intriguing, was not originally designed to ask whether learners store utterance tokens. A more straightforward design to address that question would be to familiarize infants with one set of sentences and test them on the same sentences or sentences with small changes in word order or in sublexical details.

B. COMPUTING DESCRIPTIVE STATISTICS

Turning to the question of the analytic tools that language learners have at their disposal, there is a growing body of data on this topic as well. From the data already presented, we know that learners can compute the frequency with which certain forms appear in their input. Additional data make it clear that they can do considerably more than that. One study suggests that infants are able to compute the frequency distribution of a set of input (Maye, Werker, & Gerken, 2002). Six- and 8-month-old infants were exposed for about 2 min to syllables that varied along the acoustic dimension represented by the endpoints of [d] as in *day* and the unaspirated [t] in *stay* along with filler stimuli. (Adult English-speakers perceive both endpoints as /d/ when presented in syllable-initial position.) All infants heard all of the stimuli from an eight-token continuum. However, half of the infants heard a stimulus set in which most tokens came from the middle of the continuum (tokens 4 and 5, unimodal group, dashed line in Figure 1), whereas the other half heard a set in which most tokens came from near the endpoints (tokens 2 and 7, bimodal group, solid line in Figure 1). During test, infants' listening times were measured as they were exposed to trials comprising either an ongoing alternation between the two endpoints (tokens 1 and 8, alternating trials) or a single stimulus from the continuum repeated (non-alternating trials). Each trial

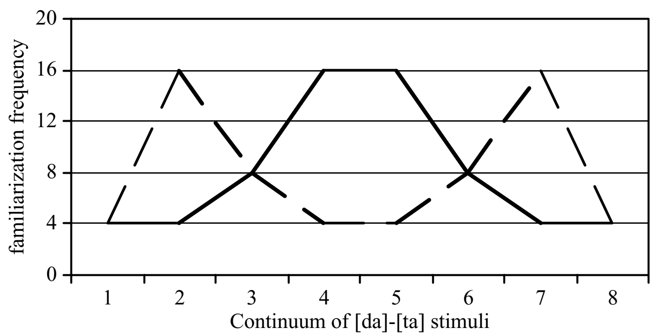

Fig. 1. Schematic of familiarization conditions in Maye, Werker, and Gerken (2002). The dashed line represents the frequency with which stimulus tokens were presented in the bimodal condition and the solid line represents the unimodal condition. Stimulus tokens 1 and 8 were presented for discrimination at test.

ended when the infant stopped fixating the visual target for a predetermined time. Only infants from the bimodal group responded differentially to the alternating vs. non-alternating trials. One interpretation of these findings is that exposure to a bimodal distribution helped infants determine that the acoustic dimension in question was potentially relevant. In contrast, exposure to a unimodal distribution made it more likely that infants would ignore the same acoustic difference. These results suggest that infants are able to perform some sort of tacit descriptive statistics on acoustic input.

C. COMPUTING TRANSITIONAL PROBABILITIES

Other research suggests that infants can track transitional probabilities between adjacent elements in strings. In their landmark study, Saffran, Aslin, and Newport (1996) familiarized 7.5-month-olds for about 2 min with four trisyllabic nonsense words (e.g., *bidaku*) strung together in random order (with the proviso that the same word not occur twice in sequence) and with no breaks between words (e.g., *bidakupadotigolabubidakutupiro…*). Infants were then tested to see whether they would discriminate two of the familiarized words (e.g., *bidaku*) from two "part words" made up of one syllable of one familiarized word and two syllables from another (e.g., *kupado*). They listened longer to part word stimuli than to familiarized words, indicating that they could discriminate the two word types.

Saffran and colleagues proposed that infants were able to use the statistical likelihood of one syllable following another to extract the actual trisyllabic words from the training stimuli. For example, in the training word *bidaku* the syllable *bi* is followed by *da* and *ku* is followed by *bi* with a probability of 100%. The syllable *pa* follows *ku* only 33% of the time. A subsequent study ruled out the possibility that effect was due only to the fact that words occurred more frequently in the training stimuli than part words (Aslin, Saffran, & Newport, 1998). Therefore, infants were apparently able to compute transitional probabilities to determine what syllables behaved as words. Similar studies using either tone or visual sequences as stimuli revealed that infants' ability to track transitional probabilities is not limited to linguistic stimuli (Kirkham, Slemmer, & Johnson, 2002; Saffran *et al.*, 1999).

Interestingly, although infants appear to be able to use transitional probabilities on a variety of input types, they may treat statistically extracted syllable strings as potential words. Saffran (2001) presented infants with the continuous syllable strings described previously. She then embedded the words and part words in either English sentence frames (*I like my _*), nonsense frames (*zy fike ny*) or tone frames (*AC#G_*). Only infants in the English frame condition listened longer to words than part words, the opposite pattern they showed when words and part words are presented with no frame at all. Although it is difficult in infant artificial language studies to interpret the direction of listening time preference, these results are suggestive. If supported by future research, they might indicate that language learners use a set of general purpose computational devices, but that different learning domains become specialized because of differences in the types of sensory and computational devices best suited for processing them (Saffran, 2001).

D. COMPUTING NON-ADJACENT DEPENDENCIES

Infants not only have the ability to track sequential dependencies, but non-adjacent dependencies as well. Santelmann and Jusczyk (1998) showed that infants are able to detect violations in dependencies between English morphemes, such as auxiliary *is* and progressive suffix-*ing*. Infants in their studies listened to sets of sentences like *Grandma is singing* vs. *Grandma can singing*. Eighteen-month-olds, but not 15-month-olds, showed a preference for grammatical sentences, but only when the distance between the two morphemes was between one and three syllables.

One potential problem with this study is that infants' preference for grammatical sentences containing *can* was not assessed. That is, infants were not tested on their ability to discriminate sentences like *Grandma can sing* vs. *Grandma can singing*. Therefore, the preference for the *is/-ing* relation could simply reflect a preference for the more frequently occurring auxiliary *is*. However, infants' ability to keep track of non-adjacent dependencies is further supported by work using a grammar with 18-month-olds (Gómez, 2002). Infants were familiarized with an artificial grammar of the form AXB and CXD, in which there is a dependency between the A and B elements and between the C and D elements. Importantly, Gómez found that, only when the second (X) element was selected from a large pool of possible syllables (24), could infants detect the relation between the first and third (A–B or C–D) elements in the grammar. Of course, in natural language, long distance dependencies encompass linguistic constituents (e.g., *Arielle called her soccer-playing friend Sara up*). Thus, to fully link current developmental research on non-adjacent dependencies to natural language, infants' sensitivity to constituent structure must ultimately be assessed.

E. COMPUTING THE RELATIONS OF IDENTICAL ELEMENTS

Infants also appear to have the ability to detect patterns of repeating or alternating elements. In one study, Gómez and Gerken (1999) presented 12-month-olds with a subset of strings produced by one of two finite state grammars. In one experiment using the headturn preference procedure, half of the infants were trained for about 2 min on strings from Grammar 1 and half on strings from Grammar 2. The two grammars allowed for a variety of strings, but importantly only Grammar 1 contained immediate repetitions of a syllable (e.g., *vot-pel-pel*).

During test, infants heard strings from the two grammars, but now instantiated in a new set of syllables. This was done by pairing each word from one vocabulary with one from the other vocabulary. Thus, an infant who heard a string like *jed-fim-fim-tup* in training might hear a string like *vot-pel-pel-jic* in test. Infants were able to transfer to the new strings, showing longer listening times to test items that matched their familiarization grammar.

Additional studies by Gómez and Gerken (1998) revealed that infants' performance was based on the pattern of repeating or alternating syllables. Infants who were exposed to strings with no repeating or alternating syllables were able to generalize to new test strings in the same vocabulary, but they were not able to transfer to test strings with new vocabulary. The finding that infants are able to generalize based on the pattern of repeating elements in a string was supported by Marcus and colleagues (Marcus *et al.*, 1999). They familiarized 7-month-olds with strings containing repetitions either at the beginning or end (e.g., *wi-wi-di* vs. *wi-di-di*) and tested them on new strings that matched or failed to match their familiarization language (e.g., *ko-ko-ba* vs. *ko-ba-ba*). Infants showed a preference for the repetition pattern that was different from the one with which they were familiarized.

F. FINDING PATTERNS OVER MULTIPLE UTTERANCE TYPES

Finally, infants appear to be able to make inferences about possible language structure from information contained over different types of utterances. Based on a study with adults by Guest, Dell, and Cole (2002), Guest, Dell, and Cole (2000) familiarized 9-month-olds with five types of three- to five-syllable words from one of two artificial languages that differed in their stress assignment principles. No single familiarization word type exhibited all of the stress assignment principles for the language. During test, infants heard new words with different stress patterns than the ones heard during familiarization, although the test words of Language 1 were consistent with the stress assignment principles of Language 1, and the test words of Language 2 were consistent with the stress assignment principles of Language 2. Importantly, Language 1 and Language 2 test words had the same stress patterns, and differed only in the placement of a heavy (CVC) syllable. For example, *do-TON-re-MI-fa* was a test word from Language 1, and *do-RE-mi-TON-fa* was a test word from Language 2; capital letters indicate stressed syllables. Infants discriminated the test words, suggesting that they were able to generalize to new words by combining information from the different types words encountered during familiarization. Six-month-olds did not discriminate the test words.

Another study demonstrating cross-utterance computations concerns infants' ability to infer proto-categories from the distribution of case inflections for Russian gender (Gerken, Wilson, & Lewis, 2005). Seventeen-month-old infants

TABLE I

Stimuli Used with 17-Month-Olds by Gerken, Wilson, and Lewis (2005)

		Feminine words			
Polkoj	rubashkoj	ruchkoj	**vannoj**	knigoj	korovoj
Polku	rubashku	ruchku	vannu	knigu	**korovu**
		Masculine words			
uchitel'ya	stroitel'ya	zhitel'ya	**medved'ya**	korn'ya	pisar'ya
uchitel'yem	stroitel'yem	zhitel'yem	medved'yem	korn'yem	**pisar'yem**

Note: Bolded words were withheld during familiarization and comprised the grammatical test items. An apostrophe after a consonant indicates that the consonant is palatalized in Russian. Ungrammatical words were *vannya*, *korovyem*, *medevedoj*, *pisaru*.

with no prior exposure to any language but English were presented with words from a Russian gender paradigm, shown in Table I. Feminine words appeared with the case endings *oj*, *u*, and masculine words with *ya* and *yem*. The 16 non-bolded words in Table I were presented in random order during familiarization. At test, the four bolded words were presented on alternate trials from ungrammatical words created by putting the wrong gender case ending on the bolded stems. Infants discriminated the grammatical from ungrammatical test items, suggesting that they were able to detect that a stem occurring with *oj* should also occur with *u*, and a stem that occurs with *ya* should also occur with *yem*. Interestingly, they made this inference under certain conditions and not others, a point to which I return when discussing possible constraints on generalization (Section V.B.3). Twelve-month-olds familiarized with the same Russian materials showed no evidence of learning. However, research using a similar paradigm with 12-month-olds showed evidence for a precursor to this type of generalization (Gómez & LaKusta, 2004).

G. HYPOTHESIS EVALUATION

One of the most vexing problems facing language learners, and therefore theories about language development, is how to know when a generalization one is making is correct. Much has been made in the field of language development about the lack of negative evidence or feedback provided by parents about the correctness of children's utterance forms (e.g., Brown & Hanlon, 1970). The need for feedback arises from situations when there is more than one possible basis of generalization for a given set of input. Work in using Bayesian models of generalization (e.g., Tenenbaum & Griffiths, 2001), coupled with infant artificial language learning research may cause us to reconsider how much of a problem the lack of negative evidence really is. To illustrate, consider the stimuli from the study by Marcus *et al.* (1999), which are shown in Table II.

TABLE II

AAB Familiarization Stimuli Used by Marcus *et al.* (1999)

	B			
A	di	je	li	we
le	leledi	leleje	leleli	lelewe
wi	wiwidi	wiwije	wiwili	wiwiwe
ji	jijidi	jijije	jijili	jijiwe
de	dededi	dedeje	dedeli	dedewe

Note: The first column and diagonal were used as familiarization stimuli in the studies by Gerken (2004a).

If one considers all of the information in the table, a succinct generalization is that all strings have an AAB form. However, if one considers only the data in the first column, all of the strings not only have an AAB form but also end in the syllable *di*. Which generalization is correct? If the infant makes the AAB generalization, and the target language really has the "ends in *di*" form, and if no adult corrects the child for creating the AAB-conforming utterance *leleje*, the child could go along indefinitely with the wrong grammar. However, if the child assumes, like a young Bayesian, that the input that she is receiving represents a random subset of the possible data, she will realize that, if all of the strings in the input end in *di*, something must be up. That is, the grammar of the language must be one in which the strings end in *di*, because the statistical chances of getting the strings in the first column if the language is a random subset of an AAB grammar are very tiny indeed. In contrast, infants presented with the data from the diagonal are indeed getting a representative subset of the possible strings generated by an AAB grammar. Therefore, infants familiarized with strings from the first column should not generalize to new AAB strings that do not end in *di*, whereas infants familiarized with strings from the diagonal should. These predictions were borne out in a set of experiments with 9-month-olds (Gerken, 2004a).

H. ASSESSMENT OF COMPUTATIONAL ENDOWMENT

It is now well-accepted that infants store a substantial amount of information about the input to which they are exposed, and that they possess a surprising array of computational abilities that may be relevant both for analyzing language structure and for evaluating potential generalizations. However, infants' computational abilities raise a set of questions: first, with the ability to store so many tokens, what causes infants to combine tokens into types or otherwise engage in abstraction? Perhaps they find storing so many tokens a barrier to efficient language processing (Stager & Werker, 1997), or perhaps some other developmental change is responsible for abstraction over tokens. Second, with the potential to perform so many analyses of their input, how do infants know which computational devices to use on what data? If infants attempted to apply every computational analysis of which they are apparently capable to a particular learning problem, and the problem is only amenable to one or two solutions, we might expect infants to generalize more slowly than they do. Third, how much overlap is there in the computational skills of human infants and non-humans (Hauser, Newport, & Aslin, 2001; Hauser, Weiss, & Marcus, 2002)? If there is considerable overlap, we need to look outside of the domain of computational ability alone to explain apparent differences between human infants and our near biological relatives with respect to their language ability. Finally, in several of the studies in this section, infants of one age, but not a younger age, showed a particular effect (e.g., 17-month-olds vs. 12-month-olds in the Russian gender paradigm). However, it is important to note that in most cases, these age-related differences are not statistically reliable, if they are tested at all. Therefore, it appears that something changes in infants' computational ability, but whether this change reflects added computational power, more efficient use of existing computational power, or some other developmental change, requires further study.

IV. What Constitutes Evidence of Generalization?

Throughout this chapter, I take the ability to generalize from input to new utterances to be the gold standard of language development. Ideally, two pieces of evidence are needed to establish that children are able to make the same types of linguistic generalizations as adults. First, the utterances under consideration must not be ones that the learner has encountered before. Second, lower level, less abstract, bases of generalization need to be reasonably ruled out, if these bases can be ruled out for adults. Although the first piece of evidence can often be found, the second is almost always open for debate. Furthermore, the nature of the debate depends partially on the type of data under consideration (e.g., discrimination, comprehension, production) and the linguistic level at issue (e.g., phonology, syntax).

A. EVIDENCE OF PHONOLOGICAL GENERALIZATION

At least two domains of phonological generalization have been explored. The first concerns generalization of phonetic features, such as place of articulation and voicing. In a classic study, Berko (1958) demonstrated that children as young as 4 years could supply the correct allomorph of the English plural (/s, z, ɛz/ when presented with a new noun (*This is a wug. Here are two _*). Because the generalization is thought to involve phonetic features in adult English, we can take children's production of new nouns to be evidence of feature-based generalization. Note, however, that a child, could answer correctly based simply on stored diphones or triphones (e.g., /gz, ts, iz, ɛz/). No feature-based generalization is necessary. A similar explanation can be applied to almost any case of morpho-phonological generalization, including the much-discussed past tense generalization (e.g., Marcus *et al.*, 1992; Rumelhart & McClelland, 1987). Indeed, the same might be argued for adult English speakers, illustrating the point that, if we apply to adults the same criteria for generalization that we apply to infants and children, we will sometimes be unable to support traditional accounts of adult behavior. One indirect way to determine whether children's responses are based on abstract features or stored diphones would be to ask whether children's production of novel plurals is uniform in its development, or whether plurals ending in more frequent diphones are mastered sooner than plurals ending in less frequent diphones. However, diphone frequency might correlate with other factors, such as ease of production, making this diagnostic weak at best.

A more direct way to demonstrate feature-based generalization would be to withhold a subset of phonemes from the learner's input and subsequently test generalization to the withheld items. No one has attempted such a study with young children to look at their productions during test. However, several studies have begun to use such methods with infants. I describe one of these later. Staying in the domain of phonological production for the moment, though, another approach to establishing feature-based generalization is to ask whether children's first mastered phonemes cluster by feature. For example, are coronals likely to be mastered sooner than velars? Although a number of claims have been made that this is the case, statistical analyses, which asked whether early mastered consonants clustered by feature more frequently than would be predicted by chance, revealed no evidence of feature classes (Zamuner, Gerken, & Hammond, in press). In sum, the data from children's productions do not make a particularly compelling case for feature-based generalization.

In contrast, a study of infant discrimination demonstrated feature-based generalization. The study is a follow-up to the work of Maye, Werker, and Gerken (2002), reported in Section III.B. In it, Maye and Weiss (2003) familiarized 9-month-olds with a unimodal or bimodal distribution of an eight-token continuum of prevoiced and short-lag stops with a coronal or velar place of articulation, synthesized from tokens produced by a Hindi speaker ([ga], [ka], [da], [ta]). During test, infants heard tokens three and six from a continuum with the unfamiliarized place of articulation (see Figure 1 for a schematic). Infants from the bimodal group were able to generalize from the familiarization continuum and discriminate the new tokens, suggesting that their basis of generalization was a phonetic feature that was in common to familiarization and test stimuli. Infants from the unimodal and no familiarization groups were not able to discriminate. Because phonetic features have acoustic correlates, infants may simply be storing acoustic events and generalizing based on acoustic similarity. However, even if an acoustic explanation is possible, only infants from the bimodal group engaged in category-based generalization.

The second domain of phonological generalization that has been explored concerns lexical stress assignment. One study of this type was presented in Section III.F (Gerken, 2004b). In it, infants were able to discriminate two types of words that had different stress patterns from the ones with which they were familiarized, based on the location of a heavy (CVC) syllable. Thus, infants discriminated words like *do-TON-re-MI-fa*, which was generated by Language 1, from *do-RE-mi-TON-fa*, which was generated by Language 2. Clearly infants were able to generalize across the utterances encountered during familiarization. However, the only heavy syllable used in this study was *TON*, and a follow-up study suggests that infants generalize only to new stress patterns containing this syllable. That is, infants appear to generalize partially based on the location of *TON*, not on the location of heavy syllables, as would be suggested by linguistic theories of stress (e.g., Dresher, 1999). What might cause infants to generalize based on units more akin to heavy syllables is not year clear.

B. EVIDENCE OF SYNTACTIC GENERALIZATION

Turning to generalization in syntax, one of the earliest characterizations of children's multiword utterances was proposed by Martin Braine (1963). Braine analyzed the words produced by three boys from about 18 months of age. He defined a word as a part of an utterance that appeared either alone or in combination with more than one other word. He noted that the majority of word combinations produced by all three boys seemed to exhibit a structure he labeled "pivot–open." Pivots in Braine's system comprise a small set of frequently used words such as *want* or *do*, which occur with words from a larger set, called open words. Legal combinations in this proto-grammar were open, pivot–open, open–pivot, and open–open. Crucially, Braine claimed that pivots could not occur alone and that two pivots could not occur together. Examples of pivot–open and open–pivot combinations appear in Table III. As is typical in studies of

TABLE III

Examples of Pivot–Open and Open–Pivot Constructions from Braine (1963)

Pivot = *want*	Pivot = *do*
want baby	bunny do
want car	daddy do
want jeep	momma do

children's spontaneous speech, no evidence is provided that any of the utterances is new. It seems unlikely that the child whose data appear in Table III ever heard an adult produce *bunny do* as the sole content of an utterance, but he might have heard *Should the bunny do it* or other similar utterances that contains the two-word sequence. But even in the extreme case that all of the child utterances are simply snippets of longer adult utterances, children do not snip randomly. Rather, there is apparent systematicity in their two-word selections, and we can treat this systematicity as a type of generalization.

The basis on which the children in Braine's study generalized is a subject of debate (e.g., Bloom, 1971). Thinking of the system Braine described as a formal grammar leads to several problems. First, it is not possible to predict which pivots precede opens and which follow. Second, some pivots occur only with some opens and not others. That is, the system is more restricted than the pivot–open notion predicts. Third, some words that were classified as pivots based on their frequency and their ability to combine with a large set of opens also combined with other words classified as pivots. For example, the same child who produced the utterances in Table III also produced *want do*. Braine claimed that, in this case, *do* was functioning as an open, not a pivot. Without an independent way of assigning words to categories, such an explanation appears circular. More generally, identifying the basis of generalization in speakers who are almost certainly constrained in the length of their utterances and the topics they find worthy of comment is a difficult, and perhaps pointless, undertaking. Nevertheless, Braine provided some of the first evidence that children's utterances reflect generalization based on word order and on some proto-category structure.

Rather than attempting to discover a child syntactic system that differed from that of adults, Valian (1986) applied to children's utterances four tests for grammatical categories used for adult language (Table IV). For example, one test for knowledge of noun phrases is the substitution of *it* for an entire phrase, as in *A wagon go boom. It go zoom zoom zoom.* A test for knowledge of adjectives is that they appear between a determiner and noun in a noun phrase, as in *a big bear*. Based on such tests, Valian concluded that there was evidence for most syntactic categories in most of the six 24- to 29-month-olds whose utterances she collected.

TABLE IV

Tests Used by Valian (1986) to Examine Young Children's Syntactic Categories

Test	Example
Word order	Determiners precede nouns
Substitutability	*it* substitutes for Noun Phrases
Multiple appearance	Adjectives should occur before nouns and as predicate adjectives
Subcategories	Count nouns used differently than mass nouns

An advantage of Valian's approach is that it applies to children the same criteria that are used to establish the basis of generalization in adult language.

A possible disadvantage of Valian's approach was pointed out a decade later by Pine and Martindale (1996). These researchers noted that Valian was primarily looking for errors and did not fully explore the contexts in which different categories were used. They observed that, in a group of seven children of similar ages to those studied by Valian, determiners were used in determiner + noun constructions and in verb + determiner + noun constructions, but almost never in determiner + noun + verb constructions or preposition + determiner + noun constructions. That is, determiner use appears to be restricted in child utterances in a way that it is not in adult utterances. This approach of testing whether children use particular syntactic classes in most or all of the possible adult contexts has been growing in popularity (e.g., Tomasello, 2000), and researchers who take this approach have a valid reason for doing so. If children used the full range of forms used by adults, we would have much clearer evidence that they were generalizing in a manner like adults.

However, in the case of the Pine and Martindale data just discussed, there is a potential explanation in the domain of constraints on language production. In particular, children are clearly better able to produce unstressed syllables in certain prosodic positions than others (Demuth, 2001; Gennari & Demuth, 1997; Gerken, 1994, 1996b). This is true both of morphemic unstressed syllables (e.g., *the*) and non-morphemic syllables (e.g., the first syllable of *giraffe*). Gerken (1994, 1996a,b) demonstrated that children are more likely to omit determiners and other unstressed syllables in sentence-initial position (*the bear chased Jane*) than in sentence-internal position (*Jane chased the bear*). Other data suggest that syllables and segments that are transcribed as omitted in fact have acoustic manifestations in instrumental analyses of children's speech (Carter & Gerken, 2004; Sadrzadeh, 2002). Furthermore, data on the motor patterns involved in producing different stress patterns reveals that strong–weak and weak–strong patterns differ in their demands on the immature motor system (Goffman, in press; Goffman & Malin, 1999; Goffman & Smith, 1999). A high priority for research on language development is to propose and test models of language production for children of the sort that have been proposed for adults (e.g., Levelt, 1989; Wijnen, 1990). Until the field has a better understanding of how mastering the task of producing speech affects patterns in the output, we run a considerable risk of misinterpreting children's production errors and production gaps. Therefore, the production gaps in the determiner system noted by Pine and Martindale (1996) may not reflect an incomplete understanding of the contexts in which English determiners can appear, but rather an incomplete mastery of the production of unstressed syllables in prosodically difficult contexts.

Other researchers have focused on gaps specifically in children's verb usage (e.g., Akhtar & Tomasello, 1997; Clark, 1996; Pine, Lieven, & Rowland, 1998; Tomasello, 1992, 2000, 2001). In brief, the claim is that children under the age of three fail to use verbs productively, but rather use them in single construction types. This claim has been supported in studies of both spontaneous speech and elicited productions of novel verbs. One study of children's ability to generalize the use of newly learned verbs employed a task in which children were asked to act out with props sentences produced by an experimenter (Akhtar & Tomasello, 1997). Children were given commands like *Make Cookie Monster dack Big Bird*. In one study, the command was preceded by the experimenter performing a novel action and labeling it; for example *This is called dacking*. In another study, the novel verb was only presented in the act out command. The question was whether children would treat the subject of the verb as the agent in whatever action they performed. The results suggest that children under the age of 3 years are very limited in their ability to generalize new verbs heard in a neutral context to a transitive sentence structure. In contrast, studies using preferential looking or pointing methodologies provide evidence that children are able to use the sentence structure (e.g., transitive) in which a novel verb appears as a cue to its interpretation (Fisher, 2002; Naigles, 1990; Naigles & Kako, 1993). For example, a 2-year-old child hearing *Big Bird is gorping Cookie Monster* is more likely to point to a video in which Big Bird is pushing Cookie Monster up and down than one in which both characters are turning arm circles.

How can we explain the discrepancy among these studies? One possibility is that using a verb in a new frame in language production is difficult. Indeed, if one selects at random any verb from an adult's vocabulary, for example *eat*, and tracks how often one adult speaker uses that verb in a week, there will probably be an enormous asymmetry in construction types used. For example, *eat* might not be used in a passive construction in the entire adult sample. This example suggests that certain verbs lend themselves to particular constructions, given the exigencies of on-line sentence production. The example could also be taken as evidence that children master individual verbs one by one, partly depending on their usage in their input. But, as Fisher (2002) points out, children could still use some basic properties of verb syntax in sentence interpretation without being able to show full productivity in the use of particular verbs. Fisher's account does not explain the results of the act out studies by Akhtar and Tomasello (1997). However, in these studies, presenting the novel verb in a two-clause imperative (e.g., *Make X verb Y*) may have taxed young children's comprehension abilities in a way that they were not taxed in the single clause stimuli used in preferential looking or pointing studies (X is verbing Y). Therefore, Fisher's (2002) notion that children have some basic understanding of how verbs function in sentence structure can be upheld. How 3-year-olds come to combine this early developing understanding of basic sentence structure with particular verbs to result in more productive verb use is not well understood.

One question that has not been addressed in the debate on early verb generalization is whether learners are able to anticipate the contexts in which a verb can occur, based on the verb's structural similarity to other verbs. For example, imagine a hypothetical experiment that uses the headturn preference procedure with 17-month-olds. Infants are exposed to a verb paradigm with some possible constructions withheld, as in the Russian gender study described in the previous section. Half of the infants would be familiarized with the sentences in Table V, and half would be familiarized with the *surprise* and *whisper* type verbs swapped (e.g., *The bunny fegged the duck*). During test, all infants would hear on alternate trials the withheld sentences from the paradigm (grammatical sentences) or sentences in which the verbs from the withheld sentences appear in the wrong verb frame (ungrammatical sentences). Imagine that infants from the two groups showed the opposite pattern of listening times. For example, suppose infants who were familiarized with the sentences in Table V showed longer listening times to *the bunny was bived* than to *the bunny was pilked*, and infants familiarized with the opposite set of sentences showed the opposite pattern of listening times.

Given infants' performance in the Russian gender paradigm study, such a finding would not be completely surprising. It would certainly indicate that infants can generalize from verbs that they have heard to new structures for those verbs, based on verb structure paradigms. This generalization would seem to be exactly the type that Tomasello and colleagues say that children under three years do not make in their meaningful verb use. Such a result would suggest that

TABLE V

Materials from a Hypothetical Infant Study of Verb Frame Generalization

Verb types	Familiarization sentences				
Surprise type verbs	The bunny gorped the duck The duck was gorped	The duck shomed the bunny The bunny was shomed	The bunny tealed the duck The duck was tealed	The duck vushed the bunny The bunny was voshed	The duck bived the bunny
Whisper type verbs	The duck fegged to the bunny The duck fegged	The bunny tammed to the duck The bunny tammed	The duck seebed to the bunny The duck seebed	The bunny pilked to the duck The bunny pilked	The bunny kaymed to the duck

Note: Grammatical test items for one group of infants: *the bunny was bived*, *the bunny kaymed*; ungrammatical test items: *the bunny was kaymed*, *the bunny bived*.

children learn about sentence structure paradigms before they attach meaning to particular verbs used in those paradigms. Combining early structural knowledge with meaning, which is required for mature performance in the studies by Tomasello and colleagues, poses a considerable challenge (Gerken, 2002; Gómez and Gerken, 2000; Naigles, 2002).

C. ASSESSMENT OF GENERALIZATION

The studies in this section point out that, although it is often possible to find evidence that infants and children generalize—in the sense that they produce, comprehend, or discriminate new utterances—it is often difficult, if not impossible, to determine the basis of generalization. This is especially true in studies of language production, because the act of production itself introduces constraints and patterns on children's utterances. Furthermore, learners appear to be very conservative in their generalizations, regardless of how they are tested (also see Section V). That is, they need multiple pieces of evidence before making a particular generalization. They appear to require the most evidence before producing a new structure, and more generally, in tasks requiring the association of form and meaning (Gerken, 2002; Gómez & Gerken, 2000; Naigles, 2002). Although they still require more than a single datum before demonstrating generalization in tasks that require them to discriminate two novel forms, they appear to require less evidence in this context. The conditions under which children are more or less conservative in their generalizations need to be explored through a larger set of parallel experiments using production vs. discrimination measures with stimuli of different types.

V. Are There Constraints on What is Learnable?

A central focus of theoretical approaches to language development is the induction problem—the logical point that any set of language input supports an infinite number of possible generalizations (e.g., Goodman, 1954). The magnitude of the induction problem depends partly on the computational endowment of the learner, discussed in Section III. Regardless of the learner's computational prowess, however, the induction problem prompts consideration of possible constraints on the types of generalizations that learners make. We can divide work on constraints on language development into two general categories internal to the language learner: constraints on the possible forms of human languages and perceptual, cognitive, and pragmatic constraints on the learner. At least two variants of each type of constraint are discussed here.

A. CONSTRAINTS ON THE POSSIBLE FORMS OF HUMAN LANGUAGES

In this section, I discuss two general approaches to constraints on possible forms of human languages that have consequences for theories of language development. One is the Principle and Parameters theory first proposed by Chomsky (1981). On this view, children are born with a set of innate parameters that reflect the possible forms that human languages can take. Some properties are common to all languages, and all children should show evidence of these absolute constraints from the earliest stages of language development. Other linguistic properties vary parametrically across languages, and learners must determine which value of each parameter is correct for their target language. The other linguistic theory that has consequences for theories of development is Optimality Theory (Prince & Smolenksy, 1993; Prince & Smolensky, 1997). On this view, linguistic patterns that are explained by rules in parameter setting accounts are explained instead in terms of the interaction of a number of ranked constraints on the forms of utterances. In Sections V.A.1 and V.A.2, I discuss these constraints in more detail and evaluate some sample data that have been used to support them.

1. Absolute Constraints on the Forms of Human Language

Perhaps the most famous example of an absolute constraint was offered by Chomsky (1980) (also see Crain, 1991). In it, a child hearing the subset of her input consisting of the statement–question pair *The man is tall, Is the man tall?* potentially faces a dilemma. She has no basis on which to determine whether the first instance of *is* is moved to the front of the sentence to make a question, or whether the more abstract unit *main verb* is moved. If a child made the former generalization, she would believe that *Is the man who tall is sad?* is a possible English sentence. In contrast, a child who believed the basis of generalization to concern the main verb would reject that sentence in favor of the grammatical *Is the man who is tall sad?*

Note that the nature of the problem here is that, after the child hears a two-clause question, the hypothesis concerning the first *is* would be ruled out. The existence of new data that would rule out some incorrect hypotheses distinguishes this type of example from another type that I discuss later. What makes children's mastery of such examples a puzzle, however, are claims that (a) children do not produce sentences like *Is the man who tall is sad*, and (b) they are not exposed early in development to two-clause questions of the sort that would rule out the incorrect hypothesis (e.g., Chomsky, 1980; Crain, 1991). If such claims are correct, it is argued children must be strongly constrained by human biology in such a way that, no matter what subset of the input they encounter, they will always make the same (correct) generalization. A great deal of the current research in linguistic theory is devoted to analyzing human languages in such a way that they can be learned without generalization errors (Baker, 2001).

I have raised several reasons why we should be cautious about interpreting children's errors and lack of errors in language production. Nevertheless, it is worthwhile to ask whether there are indeed examples of production errors that appear to reflect the wrong basis of generalization. One such example comes from Peters and Menn (1993). Menn's son Daniel apparently encountered a large number of plurals and possessives in which the stem ended in /r/ or a vowel. Rather than interpreting the sibilant plural or possessive marker as a morpheme, he appears to have drawn the conclusion that no words can end in /r/ or a vowel without a following sibilant. For example, *water* was produced as [ɔɛ*r*s].

Can it be that errors reflecting the incorrect basis of generalization do not exist in syntax, but do in morpho-phonology, as in the case with Daniel? Possibly. But another possibility is that, by the time children are producing sufficiently long utterances to reveal errors of the sort raised in Chomsky's famous example, they have already corrected any incorrect generalizations they might have made as they have been exposed to additional input. Studies in which infants are presented with artificial languages and tested on what they have learned are beginning to support this interpretation (e.g., Gerken, 2004a). Recall, for example, that in a study presented in Section III.G, 9-month-olds who were familiarized with strings ending in the syllable *di* (e.g., *le le di and wi wi di*, etc.) did not generalize to new strings with an AAB structure, whereas infants who were familiarized with strings that did not all end in the syllable *di* did generalize. Presumably, infants familiarized with strings ending in *di*, as in the previous study, but with a few strings with different final syllables added to the end of the list, *would* generalize to new AAB strings. Such a result would support the notion that infants are able to update their basis of generalization as new data come in.

With respect to the question of whether children hear the relevant input early in language development, it would be surprising if young learners never heard such utterances. A variety of data indicate that parents speak to their young children in short, but not structurally simple, utterances (e.g., Morgan, 1986; Newport, Gleitman, & Gleitman, 1977). Furthermore, computational models of syntactic generalization suggest that two clause constructions, such as *The boy who likes you is nice*, which are not questions, might be sufficient to make the appropriate generalization (see Section III.F, Lewis & Elman, 2001). Such utterances provide information about basic sentence structure that can be used in by learners who are able to make cross-utterance comparisons.

But what of linguistic generalizations for which it is claimed that no amount of input could lead to the correct generalization? Crain (1991) presented several examples of generalizations of this type. One of these concerns possible referents for pronouns in sentences like *When he ate the hamburger, the Smurf was in the box.* Most readers will agree that the pronoun *he* in this sentence can refer either to the Smurf or to another character. In contrast, *He was in the box when the Smurf ate the hamburger* cannot be interpreted with *he* referring to the Smurf. The linguistic analysis of coreference is stated to bar coreference between the pronoun and noun in the second sentence, thereby allowing all other coreference, including the two possible coreferences discussed in the first sentence. The specific statement barring coreference in the second sentence is that a pronoun cannot refer to a noun that it c-commands. In brief, main clause subjects c-command any noun phrase elsewhere in the sentence. Thus, in the second sentence, *he* c-commands *the Smurf* and therefore cannot refer to that noun phrase. This example, like all of the examples given by Crain (1991) concerns a statement about what *cannot* occur. It is the negative nature of the linguistic rule that poses the greatest challenge to all but a highly constrained learner.

However, a child might instead generalize based on the positive examples of coreference encountered. For example, when *he* is the subject of the main clause (as in the second sentence), the pronoun always refers to a referent outside the sentence, but when *he* is the subject of a subordinate clause (as in the first sentence), the pronoun can refer to a referent outside the sentence or one inside the sentence. Although it might be more formally elegant to state the constraint in the negative, there is no reason why it cannot be stated to reflect what is allowed to occur (Harris, 1991; Lasnik, 1991; McCawley, 1991; Reinhart, 1983). Let me end this example by noting that a child making a generalization of the sort just proposed still needs to distinguish between main and subordinate clauses. Therefore, to understand the developmental process, we cannot jettison all forms of linguistic analysis. This example simply demonstrates that if some formulations of a linguistic generalization are learnable and others are not, we might adopt a bias in favor of those that are learnable.

Another example of a generalization not learnable from any amount of input can be found in Lidz, Waxman, and Freedman (2003). They note that in the sentence *I'll play with this red ball and you can play with that one*, the word *one* must refer to the constituent *red ball* and not just *ball*. How, they question, might a child learn about this referential relation from the input, because in referring to *red ball*, *one* also refers to the more general *ball*. However, if we imagine the input situation of the child, she would never encounter sentences like the one here in which *one* refers to a ball that is not red. Although she may initially hypothesize that *one* refers just to the noun and not the adjective-noun constituent, the total lack of examples of *one* referring to balls other than red ones would be sufficient for certain types of learners to rule out the faulty hypothesis. As noted in Section III.G, several researchers are exploring the type of learner that could use differences between expected and actual utterances encountered (e.g., Gerken, 2004a; Tenenbaum & Griffiths, 2001). If children are such learners, this particular example of an unlearnable generalization must be discounted.

In summary, it would seem that the examples of absolute constraints on the forms of human language, and the evidence used to support these constraints, must be viewed very skeptically. This is not to say that such constraints might not exist. However, as we move from considering language production data to studies of younger and younger infants, their pattern of ability and inability suggest that they may briefly consider surprising types of generalizations.

2. *Parameterized Constraints*

Parameter setting accounts of language development are based on typological studies of human language, which have revealed that languages differ from each other in ways that can often be described in terms of two- or three-valued parameters (Baker, 2001; Chomsky, 1981; Roeper & Williams, 1987). Interestingly, each parameter value predicts a range of data in a particular language. For example, in many languages, the subject appears first in a transitive sentence, with the verb and object following in either a VO or OV order. These two orders predict a variety of other order relations in the language, shown in Table VI. The whole complex of ordering relations is captured under the Head Directionality Parameter (Baker, 2001).

The advantage of such a parameter is that a child could generalize broadly, having encountered only one of the predicted patterns. That is, a child who realized that determiners precede nouns would be able to determine the basic word order of her language, plus all of the other data shown in Table VI. On many accounts of parameter setting, each parameter has a default value, with which all learners are born. Incorrect parameter settings (i.e., default setting is A, but the target language requires setting B) can be reset based on a single input datum or trigger. On this view, the learner's job is to reset incorrectly set parameters, and errors arise when a learner has not yet reset a particular parameter. Thus, errors reflect a possible generalization for a human language, but not the child's own target language.

TABLE VI

Implications of the Proposed Head Directionality Parameter

Verb + Object order predicts	Object + Verb order predicts
Prepositions	Postpositions
Auxiliaries before main verbs	Auxiliaries after main verbs
Subordinating conjunction precedes embedded clause	Subordinating conjunction follows embedded clause
Articles precede nouns	Articles follow nouns

This is a fascinating and powerful theory that ties language development with patterns observed in languages of the world. It has the potential to account not only for children learning language through normally rich input, but also the types of situations discussed in Section II.B, in which children create language structure with little or no input (Feldman, Goldin-Meadow, & Gleitman, 1978; Goldin-Meadow & Mylander, 1998; Ross & Newport, 1996; Senghas, 2003; Senghas & Coppola, 2001; Singleton & Newport, in press).

A problem with the parameter setting account is that the evidence used to support it thus far has been quite weak. For example, evidence that children set their Head Direction Parameter very early is that they almost always produce word order consistent with their input (Baker, 2005). Thus, English learning children say *read book* not *book read*. But what theory would predict otherwise? The incorrect order is not in the input.

Other parameters are supported, not by correct performance, but by errors. For example, languages of the world differ as to whether they allow the omission of pronoun subjects of declarative sentences. English does not allow subjectless sentences, but English-learning children produce them, which has been taken as evidence that the default parameter setting of the relevant parameter (i.e., Pro-Drop) is set to allow subjectless sentences (Hyams, 1986, 1992). Hyams hypothesized that the existence of sentences like *It is raining*, which contains a semantically empty subject, allows English learners to reset the parameter to the correct value.

There are several problems with this account. One is true of all parameter setting accounts of children's errors. If the errors are corrected when the child sets her parameter based on a piece of triggering data, it is incumbent on the theory to explain why the child ignored the relevant data to that point. That is, the theory must be embedded in a psychologically plausible model of the language-learning child. A problem specific to the Pro-Drop parameter is that children omit any number of grammatical morphemes, including pronouns, determiners, and prepositions. Furthermore, as noted in Section IV.B, these omissions seem to be at least partially driven by prosodic factors (Demuth, 2001; Gennari & Demuth, 1997; Gerken, 1994, 1996b), although other factors may play a role (Becker, 2004). Evidence that young English-speakers' subjectless sentences can be attributed to production constraints can be seen in work by Bloom (1990), who demonstrated that children omit sentential subjects more frequently when their utterances are longer than when they are shorter. In a similar vein, Valian (1991) demonstrated that children learning English omit sentential subjects much less often than their Italian learning counterparts, whose target language does allow subjectless sentences.

In summary, the parameter setting account is elegant and intriguing, but proposed parameters are currently not well supported by child language data. The method of testing this account has largely been to look at the production errors that children make and associate these with possible incorrect parameter settings. To pursue such accounts in a convincing way, theorists must make *a priori* predictions about what is learnable and what is not, and what data are relevant to testing the predictions.

3. Ranked Constraints

As noted at the beginning of this section, the regularities in language that are explained by rules in a parameter setting framework are instead explained in terms of violable, ranked constraints in Optimality Theory (McCarthy & Prince, 1993; Prince & Smolenksy, 1993). On this view, an infant is born with a set of ranked constraints on the forms of utterances. The influence of highly ranked constraints is more likely to be seen in surface forms (produced by the child) than the influence of lowly ranked constraints. Constraints can be classified into two types: markedness constraints, which reflect what is typical or preferred among the world's languages, and faithfulness constraints, which dictate conformity with the ambient language. For example, a markedness constraint indicating that syllables should not have a final consonant comports with the fact that all of the worlds' languages allow such syllables, and some languages allow only syllables without a final consonant. A faithfulness constraint for English might indicate that *dog* is produced with three segments. The learner's job is to determine the correct ranking of all constraints for her language, and many errors are thought to arise because markedness constraints are initially ranked above faithfulness constraints. Thus, as in parameter setting accounts, errors reflect a possible generalization for a human language, but not the child's own target language. A difference between Optimality Theory and parameter setting accounts is that, under Optimality Theory, it is predicted that children's language will not only reflect what is possible among the world's languages, but also what is typical.

This approach to language development has been extremely productive in the accounting for children's phonological errors of production (Barlow & Gierut, 1999; Demuth, 1996; Gnanadesikan, 1996; Ohala, 1999). One example of Optimality Theory being applied to child language concerns syllable shapes. In particular, all languages of the world have consonant–vowel syllables (CV), even though some languages may have other syllable types. Therefore, we might argue that CV is the unmarked, or preferred syllable shape. In Optimality Theory, this syllable shape is reflected by a set of constraints, one of which is NO CODA, which means that words should not have final consonants. In adult English, NO CODA would be ranked below faithfulness constraints that would allow us to produce words like *dog*. However, in children, the NO CODA markedness constraint should outrank faithfulness constraints, and children should produce words without codas. This notion is consistent with the frequent observation that English-learning children produce utterances like [dɔ] for *dog*. Furthermore, children omit voiced codas more frequently than voiceless codas and velar codas more frequently than coronal codas (for a complete review, see Zamuner, 2003). Both of these statistical tendencies in children's omissions are also found across adult languages.

However, a problem with such apparent parallels in preferred forms between children and cross-linguistic data is that what is frequent cross-linguistically is often most frequent in an individual language (Greenberg, 1974). For example, CV syllables are not the only syllable allowed in many languages of the world, they are also more frequent than the other allowed syllable shapes in English. Therefore, children's coda omissions may reflect the frequency of forms encountered in the target language, and not cross-linguistic markedness, *per se*. In one study that attempted to tease apart the role of markedness and input frequency (Zamuner, 2003; Zamuner, Gerken, & Hammond, 2004), children from 21 to 29 months of age were asked to produce pairs of CVC nonsense words in which the same coda occurred in a high- or low-frequency context (e.g., /gɛl/ vs. /pʌl/). High-frequency contexts were ones in which the diphones forming the CV and VC were frequent in English, whereas these diphones were low frequency in English for the low-frequency member of the pair. Because Optimality Theory makes references to codas as a class, it has no reason to predict differences among codas depending on segmental context. Nevertheless, children were more accurate in their coda productions in the high-frequency contexts, thereby supporting input frequency over markedness as an explanation of coda omissions.

In summary, Optimality Theory, like the theory of Principles and Parameters, offers the possibility of linking language development with patterns observed in human languages. However, it suffers from several of the same problems, as well as some of its own. First, there is no statement about what causes a child to re-rank faithfulness and markedness constraints. That is, why does a child who has heard the word *dog* for 2 years begin to produce the final consonant? Second, the data that are used to support the theory nearly all come from children's errors in production, which I have noted several times are likely to require a theory of developing language production. Third, as Zamuner and colleagues have demonstrated, input frequency appears to provide a better account of children's errors, at least coda omissions, than does Optimality Theory. Finally, in Optimality Theory, possible utterance forms are selected from a set of generated candidates. The psychological mechanism by which these candidates might be generated has not been spelled out.

B. PERCEPTUAL AND COGNITIVE CONSTRAINTS

An increasingly popular approach to understanding how language development might be constrained by properties internal to language learners is that general perceptual, cognitive, and pragmatic constraints may be sufficient.

I discuss four such constraints that have been proposed: the less-is-more hypothesis, prosodic bootstrapping, the necessity of multiple cues to categories, and constraints on word learning.

1. *The Less-is-More Hypothesis*

Newport's (1991) less-is-more hypothesis concerns the problem that learners face mapping a linear sequence of morphemes that compose an utterance onto a unitary scene that the utterance describes. If the entire utterance is mapped as a unit to the scene, then the learner can never discover the compositional nature of language. At the other extreme, how is a learner to determine which elements of a 10-morpheme-utterance map onto which elements of a scene? A solution is that, if infants have a filter on how much linguistic information they can encode and remember, they have fewer language units to map to the scene. Computational models suggest that better learning is indeed achieved under circumstances of either strategically or randomly filtered input (Elman, 1993; Goldowsky & Newport, 1993). For example, Goldowky and Newport discussed a model in which forms containing three pieces of information were filtered to contain one or two pieces of information before presentation to the learning algorithm, which resulted in better learning than non-filtered input. However, other researchers argue that filtering input in these ways may be unnecessary for learning a natural language (Rohde & Plaut, 1999). Whether or not this specific hypothesis is correct, the notion that the limited processing capacity of young children might give them an advantage in learning language is one very much in the forefront of modern thinking about language development.

2. *Prosodic Bootstrapping*

Although young language learners appear to have an impressive collection of computational tools at their disposal (Section III), it would be computationally impossible to test all of the possible relations that might exist between linguistic elements in utterances. For example, in a corpus of 10-morpheme sentences, a learner would need to compute the 90 relations per sentence just to determine if there were any pairwise dependencies. One type of information that learners might use to restrict their computational space is to first look for reliable relations within prosodically defined phrases and clauses (e.g., Gerken, 1996a; Morgan, 1986; Morgan, Meier, & Newport, 1987). Although learners are clearly sensitive to prosodic cues to major linguistic units (e.g., Hirsh-Pasek *et al.*, 1987; Jusczyk & Kemler Nelson, 1996; Jusczyk *et al.*, 1992; Kemler Nelson *et al.*, 1989; Mandel, Jusczyk, & Kemler Nelson, 1994; Nazzi *et al.*, 2000), research is still needed to demonstrate that they indeed use this sensitivity to limit the domain of their search for statistical regularities in the input.

3. Multiple Cues to Syntactic Categories

Another possible constraint on language learning that I describe in this section concerns the Russian gender paradigm experiment discussed in Section III.F. Recall that infants appeared to discern that words that occurred with *oj* also occurred with *u*, and words that occurred with *ya* also occurred with *yem* (Gerken *et al.*, 2005; see Table I). However, they were only able to do so when a subset of the feminine and masculine words had a double marking of their gender category. Thus, in Table 1, three of the six feminine word stems ended in /k/ and three of the masculine stems ended in /tɛl/. Other studies with adults demonstrate the same requirement for double marking on a subset of words before learners can discern the category structure (Braine, 1987; Gerken *et al.*, 2002; Mintz, 2002). Such a cognitive bias on category formation prevents learners from overgeneralizing based on a single cue. For example, a variety of English nouns can also appear as verbs, as in *the brush* and *can brush*. A learner who noted these two phrases and other similar ones might conclude that words occurring with *the* also occur with *can*, which would lead to a conflation of nouns and verbs. However, if learners only make category generalizations when two or more markers occur in a subset of phrases (e.g., *the brushes*), the likelihood that the input would allow a category overgeneralization error is minute (Gerken *et al.*, 2004).

4. Constraints on Word Learning

The debate concerning the need for language-specific vs. domain neutral constraints on the development of language structure is paralleled by a similar debate concerning word learning. Although some researchers have proposed constraints specific to word learning (Markman, 1989, 1991; Waxman & Markow, 1995), others argue that word learning instead reflects general computational abilities acting on the statistics of the input (e.g., Samuelson, 2002; Smith, 1999; Smith *et al.*, 2002) or is constrained by the learner's theory of other humans' naming intentions (e.g., Akhtar & Tomasello, 2000; Bloom, 1998, 2000).

For example, Markman (1991) proposed that children operate under a mutual exclusivity constraint, causing them to avoid assigning two names to the same object. A test for this constraint is to introduce children to two new objects, name one of them, and then ask for the X, where X is a word that the child has never heard. Children choose the previously unnamed object as the referent for X. However, Markson and Bloom (1997) found that 3- and 4-year-old children perform in the same way if they are shown two new objects, told a fact about one object (e.g., *my uncle gave this to me*), told nothing about the other object, and then asked to select an X. Children select the object about which no information was given. The finding that very young children appear sensitive to others' intentions is in itself impressive. However, the question remains as to whether a namer's intent is simply another (often highly reliable) factor in the input that learners compute, or whether intent has an innately privileged status (Bloom, 2000; Golinkoff *et al.*, 2000). Whether or not perceived intent is a special form of information for word learning, or part of a larger arsenal of general cognitive abilities that children bring to the task, a growing body of data suggests that children may not need language-specific constraints to learn words.

5. Assessment of Constraints

A central question is whether there are constraints specific for learning language, or whether a more general learning mechanism, coupled with rich input, can account for the data. Many of the issues concerning constraints are the same as the ones already raised in the section on creativity. How can a general learning mechanism explain the similarities that exist across languages and our ability to create new languages? However, the data for accounts in which children have preset linguistic values among which they choose (parameter setting and constraint re-ranking) do not account well for the growing body of data indicating that input frequency may be the single most potent factor influencing children's language development. Nevertheless, research on perceptual and cognitive constraints does not yet suggest that these are sufficient to guide learners to the correct generalizations for their language.

VI. Conclusion—What Develops in Language Development?

To understand the mechanism by which language develops in humans, we need to consider both the human learner as a biological and psychological entity and human language as a species characteristic with particular formal properties and distributional characteristics. I have focused primarily on the child learner, which is the typical the domain of psychologists and others taking behavioral approaches to language development. Our understanding of the capabilities of the learner has grown enormously in the past 20–30 years. However, different snapshots of infant and child abilities combine into a whole that is difficult interpret. The previous four sections illuminated three puzzles about the child learner that researchers must solve on our way to having a biologically and psychologically feasible model of language development.

The first puzzle, noted in Section III, is that although children change over time, studies that reveal change typically show a statistically reliable ability at one age and no reliable ability at an earlier age. Seldom are the apparent age-related differences themselves statistically reliable, if they are tested at all. What model of development is consistent with the typical pattern of results over age? One possible model is that children change their basis of organization over development. However, this situation predicts that we would find statistically different behaviors within each age and statistically reliable differences between age groups (e.g., Stager & Werker, 1997). Another possible model is that individual children adopt different bases of organization at early ages and become more similar to each other at later ages. Such a model is consistent with the existing data. In addition, it predicts that individual children at early ages should show consistent performance over multiple tests. A third possible model is that children begin with no organization and converge on similar bases of organization with development. This model predicts little test–retest reliability for individual children. Further research is needed to determine which, if any, of these models is consistent with development of different aspects of language.

Second, the discussions in Section IV revealed that learners store a great deal of the input to which they are exposed, yet they generalize beyond that input, and studies of infant learning in the laboratory suggest that the generalization process can be quite rapid. Furthermore, the degree of generalization varies widely across ages and measures. Exploring the circumstances under which learners of different ages are more likely to respond based on specific stored information vs. more abstract information will help to solve this puzzle.

A third puzzle is that children make many errors in their early productions, and these errors reveal some degree of systematicity. In Section V, I noted that these errors can be interpreted within linguistic frameworks, but the frameworks can seldom predict particular errors *a priori*. The errors also can be interpreted to reflect the statistics of the input, but just how much of children's systematicity can be explained by input frequency has not been fully explored. Each of these current puzzles must be resolved as we attempt to understand how an individual learner masters the language(s) to which she is exposed.

Turning to language as a species characteristic, this is typically the domain of linguists and computer scientists who seek to explore the formal properties of human language and its manifestations across cultures and modalities. In Section II, I raised several points about how language is distributed across humans that must borne in mind as we struggle to grasp how children come to partake of the powerful information transfer system afforded them by their membership in our species. These points are that: languages of the world share some basis units of generalization and entailments (e.g., Head Directionality); humans not only learn but modify and create languages; children have an advantage in language learning and language creation.

In Sections V and III, respectively, I offered discussions of formal properties of language and computational abilities of the learner. I would be so bold as to claim that the formal learning problem for language, although formidable, has been overblown. In particular, researchers need to seriously reconsider the notion that there is no information in the input to infants and children that would allow them to select among possible generalizations. However, the induction problem remains, and the number of possible generalizations that the input allows is indeed infinite. But the problem is not that there is not enough information in the input; rather there is too much (e.g., Saffran *et al.*, 1996). The richness of the input data is nearly matched by the number of computational tools from which the learner might use to analyze it. Furthermore, at least some of the tools are available to non-humans (e.g., Hauser, Newport, & Aslin, 2001; Hauser, Weiss, & Marcus, 2002; Ramus *et al.*, 2000). Therefore, it is unlikely that the computational tools themselves explain how humans acquire language. Rather, it may be how the tools are selected and used that is important.

The considerable computational arsenal available to humans suggests that learners must be constrained to consider only certain input and apply to that input only certain units of analysis and computational procedures. Whether the required constraints are specific to language or more general is very much an open question. However, we must consider both the biological and psychological status of the child learner and the way that language is distributed over the species as we engage in the debate. It would be of little use to find that an appropriately constrained domain neutral learner could acquire human language, only to be at a loss to explain the other data we think are relevant to the problem of language development.

To conclude, the various disciplines that ask what develops in language development are in a similar position to the child learner—we have too much information to be accommodated by any of the existing approaches. Theories that both take into account the range of data and relinquish untenable assumptions are desperately needed. However, the strong and growing interdisciplinarity of the field of language development gives me reason to be hopeful that, as students are exposed to the full range of data that we must explain, these new theories will be forthcoming.

ACKNOWLEDGEMENTS

Manuscript preparation was supported by NIH grant HD042170.The author thanks Rebecca Gómez and Jessica Maye for helpful comments on an earlier draft.

REFERENCES

Akhtar, N., & Tomasello, M. (1997). Young children's productivity with word order and verb morphology. *Developmental Psychology*, *33*, 952–965.

Akhtar, N., & Tomasello, M. (2000). The social nature of words and word learning. In R. M. Golinkoff & K. Hirsh-Pasek (Eds.), *Becoming a word learner: A debate on lexical acquisition* (pp. 115–135). Oxford: Oxford University Press.

Aslin, R. N., Saffran, J. R., & Newport, E. L. (1998). Computation of conditional probability statistics by 8-month-old infants. *Psychological Science*, *9*, 321–324.

Baker, M. C. (2001). *The atoms of language*. New York, NY: Basic Books.
Baker, M. C. (2005). Mapping the terrain of language learning. *Language Learning and Development*, *1*(1), 93–129.
Barlow, J. A., & Gierut, J. A. (1999). Optimality theory in phonological acquisition. *Journal of Speech, Language, & Hearing Research*, *42*, 1482–1498.
Becker, M. (2004). *Children's omission of function morphemes: Syntactic and prosodic influences*. Unpublished manuscript, Chapel Hill, NC.
Beckman, M. E., & Edwards, J. (2000). Lexical frequency effects on young children's imitative productions. In M. Broe & J. Pierrehumbert (Eds.), *Papers in laboratory phonology V: Language acquisition and the lexicon* (pp. 208–217). Cambridge: Cambridge University Press.
Berko, J. (1958). The child's learning of English morphology. *Word*, *14*, 150–177.
Bever, T. (1970). The cognitive basis for linguistic structures. In J. Hayes (Ed.), *Cognition and the development of language* (pp. 279–362). New York: Wiley.
Bloom, L. (1971). Why not pivot grammar. *Journal of Speech & Hearing Disorders*, *36*, 40–50.
Bloom, L. (1998). Language acquisition in its developmental context. In D. Kuhn & R. S. Siegler (Eds.), *Handbook of child psychology*. *Cognition, perception, and language* (5th ed., Vol. 2, pp. 309–370). New York: Wiley.
Bloom, P. (1990). Subjectless sentences in child language. *Linguistic Inquiry*, *21*, 491–504.
Bloom, P. (2000). *How children learn the meanings of words*. Cambridge, MA: MIT Press.
Braine, M. D. S. (1963). The ontogeny of English phrase structure. *Language*, *39*, 1–14.
Braine, M. D. S. (1987). What is learned in acquiring word classes—a step toward an acquisition theory. In B. MacWhinney (Ed.), *Mechanisms of language aquisition* (pp. 65–87). Hillsdale, NJ: Erlbaum.
Brown, R., & Hanlon, C. (1970). Derivational complexity and the order of acquisition in child speech. In J. Hayes (Ed.), *Cognition and the development of language* (pp. 11–54). New York: Wiley.
Carter, A., & Gerken, L. A. (2004). Do children's omissions leave traces? *Journal of Child Language*, *31*, 561–586.
Chomsky, N. (1955). *The logical structure of linguistic theory*. New York: Plenum Press.
Chomsky, N. (1980). The linguistic approach. In M. Piattelli-Palmarini (Ed.), *Language and learning* (pp. 109–116). Cambridge, MA: Harvard University Press.
Chomsky, N. (1981). *Lectures on government and binding*. Dordrecht: Foris.
Clark, E. V. (1996). Early verbs, event types, and inflections. In C. E. Johnson & J. H. V. Gilbert (Eds.), *Children's language* (Vol. 9, pp. 61–73). Hillsdale, NJ: Erlbaum.
Crain, S. (1991). Language in the absence of experience. *Behavioral & Brain Sciences*, *14*, 597–650.
Demuth, K. (1996). The prosodic structure of early words. In J. L. Morgan & K. Demuth (Eds.), *Signal to syntax: Bootstrapping from speech to grammar in early acquisition* (pp. 171–184). Hillsdale, NJ: Erlbaum.
Demuth, K. (2001). A prosodic approach to filler syllables. *Journal of Child Language*, *28*, 246–249.
Dresher, B. E. (1999). Child phonology, learnability, and phonological theory. In W. C. Ritchie & T. K. Bhatia (Eds.), *Handbook of child language acquisition* (pp. 299–346). San Diego, CA: Academic Press.
Ellis, N. C. (2002). Frequency effects in language processing: A review with implications for theories of implicit and explicit language acquisition. *Studies in Second Language Acquisition*, *24*, 143–188, Special Issue: Frequency effects in language processing: A review with commentaries.

Elman, J. L. (1993). Implicit learning in neural networks: The importance of starting small. *Cognition*, *46*, 71–99.

Elman, J. L., Bates, E. A., Johnson, M. H., Karmiloff-Smith, A., Parisi, D., & Plunkett, K. (1996). *Rethinking innateness: A connectionist perspective on development*. Cambridge, MA: The MIT Press.

Feldman, H., Goldin-Meadow, S., & Gleitman, L. (1978). Beyond herotodus: The creation of language by linguistically deprived deaf children. In A. Lock (Ed.), *Action, symbol, and gesture: The emergence of language* (pp. 351–414). New York: Academic Press.

Ferreira, F., Bailey, K., & Ferraro, V. (2002). Good-enough representations in language comprehension. *Current Directions in Psychological Science*, *11*, 11–15.

Fisher, C. L. (2002). Structural limits on verb mapping: The role of abstract structure in 2.5-year-olds interpretation of novel verbs. *Developmental Science*, *5*, 55–64.

Gennari, S., & Demuth, K. (1997). Syllable omission in the acquisition of Spanish. In E. Hughes, M. Hughes, & A. Greenhill (Eds.), *Proceedings of the Annual Boston University Conference on Language Development* (Vol. 1, pp. 182–193). Somerville, MA: Cascadilla Press.

Gerken, L. A. (1994). A metrical template account of children's weak syllable omissions from multisyllabic words. *Journal of Child Language*, *21*, 565–584.

Gerken, L. A. (1996a). Phonological and distributional cues to syntax acquisition. In J. Morgan & K. Demuth (Eds.), *Signal to syntax* (pp. 411–425). Mahwah, NJ: Erlbaum.

Gerken, L. A. (1996b). Prosodic structure in young children's language production. *Language*, *72*, 683–712.

Gerken, L. A. (2002). Early sensitivity to linguistic form. *Annual Review of Language Acquisition*, *2*, 1–36.

Gerken, L. A. (2004a). *Language learning is constrained by nature, but with room for errors*. Unpublished manuscript, Tucson.

Gerken, L. A. (2004b). Nine-nonth-olds extract structural principles required for natural language. *Cognition*.

Gerken, L. A., Wilson, R., & Lewis, W. (2005). 17-month-olds can use distributional cues to form syntactic categories. *Journal of Child Language*, *32*(*2*), 1–20.

Gerken, L. A., Wilson, R., Gómez, R. L., & Nurmsoo, E. (2002). *Linguistic category induction without reference: The importance of correlated cues*. Unpublished manuscript, Tucson.

Gerken, L. A., & Zamuner, T. (in press). Exploring the basis for generalization in language acquisition. In J. Cole & J. Hualde (Eds.), *LabPhon IX: Change in phonology*. The Hague: Mouton de Gruyter.

Gleitman, L. R. (1990). The structural sources of verb meanings. *Language Acquisition*, *1*, 3–55.

Gnanadesikan, A. E. (1996). Child phonology in optimality theory: Ranking markedness and faithfulness constraints. In A. Stringfellow, D. C. Amitay, E. Hughes, & A. Zukowski (Eds.), *Proceedings of the 20th Annual Boston University Conference on Language Development* (pp. 237–248). Somerville, MA: Cascadilla Press.

Goffman, L. (2004). Children's and adults' production of rhythmic speech-motor sequences in content and function words. *Journal of Speech, Language & Hearing Research*, *47*(*5*), 1088–1102.

Goffman, L., & Malin, C. (1999). Metrical effects on speech movements in children and adults. *Journal of Speech, Language, & Hearing Research*, *42*, 1003–1115.

Goffman, L., & Smith, A. (1999). Development and phonetic differentiation of speech movement patterns. *Journal of Experimental Psychology: Human Perception & Performance*, *25*, 649–660.

Goldin-Meadow, S., & Mylander, C. (1998). Spontaneous sign systems created by deaf children in two cultures. *Nature*, *391*, 279–281.

Goldowsky, B. N., & Newport, E. (1993). Modeling the effects of processing limitations on the acquisition of morphology: The less is more hypothesis. In E. Clark (Ed.), *The Proceedings of the 24th Annual Child Language Research Forum* (pp. 124–139). Stanford, CA: CSLI.

Golinkoff, R. M., Hirsh-Pasek, K., Bloom, L., Smith, L. B., Woodward, A. L., Akhtar, N. *et al.* (2000). *Becoming a word learner: A debate on lexical acquisition.* London: Oxford University.

Gómez, R. L. (2002). Variability and detection of invariant structure. *Psychological Science*, *13*, 431–436.

Gómez, R. L., & Gerken, L. A. (1998, April). *Determining the basis of abstraction in artificial language acquisition*. Paper presented at the International Society on Infant Studies, Atlanta.

Gómez, R. L., & Gerken, L. A. (1999). Artificial grammar learning by 1-year-olds leads to specific and abstract knowledge. *Cognition*, *70*, 109–135.

Gómez, R. L., & Gerken, L. A. (2000). Infant artificial language learning and language acquisition. *Trends in Cognitive Sciences*, *4*, 178–186.

Gómez R. L., & LaKusta L. (2004). A first step in form-based category abstraction by 12-month-old infants. *Developmental Science*, *7*(5), 567–580.

Goodman, N. (1954). *Fact, fiction, and forecast*. University of London: Athlone Press.

Greenberg, J. H. (1963). Some universals of grammar with particular reference to the order of meaningful elements. In J. H. Greenberg (Ed.), *Universals of language* (pp. 58–90). Cambridge, MA: MIT Press.

Greenberg, J. H. (1974). *Language typology: A historical and analytic overview*. The Hague: Mouton.

Guest, D. J., Dell, G. S., & Cole, J. S. (2000). Violable constraints in language production: Testing the transitivity assumption of optimal theory. *Journal of Memory & Language*, *42*, 272–299.

Harris, C. L. (1991). Alternatives to linguistic arbitrariness. *Behavioral & Brain Sciences*, *14*, 622–623.

Hauser, M. D., Newport, E. L., & Aslin, R. N. (2001). Segmentation of the speech stream in a non-human primate: Statistical learning in cotton-top tamarins. *Cognition*, *78*, B53–B64.

Hauser, M. D., Weiss, D., & Marcus, G. F. (2002). Rule learning by cotton-top tamarins. *Cognition*, *86*, B15–B22.

Hirsh-Pasek, K., Kemler Nelson, D., Jusczyk, P. W., Wright Cassidy, K., Druss, B., & Kennedy, L. (1987). Clauses are perceptual units for prelinguistic infants. *Cognition*, *26*, 269–286.

Houston, D. M., & Jusczyk, P. W. (2000). The role of talker-specific information in word segmentation by infants. *Journal of Experimental Psychology: Human Perception & Performance*, *26*, 1570–1582.

Houston, D. M., & Jusczyk, P. W. (2003). Infants' long-term memory for the sound patterns of words and voices. *Journal of Experimental Psychology: Human Perception & Performance*, *29*, 1143–1154.

Hudson, C., & Newport, E. L. (1999). Creolization: Could adults really have done it all? In A. Greenhill, H. Littlefeld, & C. Tano (Eds.), *Proceedings of the 23rd Annual Boston University Conference on Language Development* (pp. 265–276). Somerville, MA: Cascadilla Press.

Hudson Kam, C., & Newport, E. (2005). Regularizing unpredictable variation in creole formation: The roles of adult and child learners. *Language Learning and Development*, *1*(2), 151–195.

Hyams, N. (1986). *Language acquisition and the theory of parameters*. Dordrecht: Reidel.

Hyams, N. (1992). A reanalysis of null subjects in child language. In J. Weissenborn & H. Goodluck (Eds.), *Theoretical issues in language acquisition: Continuity and change in development* (pp. 249–267). Hillsdale, NJ: Erlbaum.

Jusczyk, P. W., & Kemler Nelson, D. (1996). Syntactic units, prosody, and psychological reality during infancy. In J. Morgan & K. Demuth (Eds.), *Signal to syntax: Bootstrapping from speech to grammar in early acquisition* (pp. 389–410). Cambridge, MA: MIT Press.

Jusczyk, P. W., Luce, P. A., & Charles-Luce, J. (1994). Infants' sensitivity to phonotactic patterns in the native language. *Journal of Memory & Language*, *33*, 630–645.

Jusczyk, P. W., Hirsch-Pasek, K., Kemler Nelson, D. G., Kennedy, L. J., Woodward, A., & Piwoz, J. (1992). Perception of acoustic correlates of major phrasal units by young infants. *Cognitive Psychology*, *24*, 252–293.

Kemler Nelson, D., Hirsh-Pasek, K., Jusczyk, P. W., & Wright Cassidy, K. (1989). How prosodic cues in motherese might assist language learning. *Journal of Child Language*, *16*, 53–68.

Kemler Nelson, D., Jusczyk, P. W., Mandel, D. R., Myers, J., Turk, A. E., & Gerken, L. A. (1995). The headturn preference procedure for testing auditory perception. *Infant Behavior and Development*, *18*, 111–116.

Kirkham, N. Z., Slemmer, J. A., & Johnson, S. P. (2002). Visual statistical learning in infancy: Evidence for a domain general learning mechanism. *Cognition*, *83*, B35–B42.

Lasnik, H. (1991). Language acquisition and two types of constraints. *Behavioral & Brain Sciences*, *14*, 624–625.

Levelt, W. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.

Lewis, J. D., & Elman, J. L. (2001). A connectionist investigation of linguistic arguments from the poverty of the stimulus: Learning the unlearnable. In J. D. Moore & K. Stenning (Eds.), *Proceedings of the 23rd Annual Conference of the Cognitive Science Society* (pp. 552–557). Mahwah, NJ: Erlbaum.

Lidz, J., Waxman, S., & Freedman, J. (2003). What infants know about syntax but couldn't have learned: Experimental evidence for syntactic structure at 18 months. *Cognition*, *89*, B69–B73.

MacDonald, M., Pearlmutter, N., & Seidenberg, M. (1994). The lexical nature of syntactic ambiguity resolution. *Psychological Review*, *101*, 676–703.

Mandel, D. R., Jusczyk, P. W., & Kemler Nelson, D. G. (1994). Does sentential prosody help infants organize and remember speech information? *Cognition*, *53*, 155–180.

Marcus, G. F., Pinker, S., Ullman, M., Hollander, M., Rosen, T., & Xu, F. (1992). Overregularization in language acquisition. *Monographs of the Society for Research in Child Development*, *57*(4), Serial No. 228.

Marcus, G. F., Vijayan, S., Rao, S. B., & Vishton, P. M. (1999). Rule learning by seven-month-old infants. *Science*, *283*, 77–80.

Markman, E. M. (1989). *Categorization and naming in children: Problems of induction*. Cambridge, MA: MIT Press.

Markman, E. M. (1991). The whole-object, taxonomic, and mutual exclusivity assumptions as initial constraints on word meanings. In S. S. Gelman & J. P. Byrnes (Eds.), *Perspective and language and thought: Interrelations in development* (pp. 72–106). Cambridge, England: Cambridge University Press.

Markson, L., & Bloom, P. (1997). Evidence against a dedicated system for word learning in children. *Nature*, *385*, 813–815.

Maye, J., & Weiss, D. J. (2003). Statistical cues facilitate infants' discrimination of difficult phonetic contrasts. In B. Beachley, A. Brown, & F. Colin (Eds.), *Proceedings of the 27th Annual Boston University Conference on Language Development* (pp. 508–518). Somerville, MA: Cascadilla Press.

Maye, J., Werker, J. F., & Gerken, L. A. (2002). Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition*, *82*, B101–B111.

McCarthy, J. J., & Prince, A. (1993). *Generalized alignment*. Unpublished manuscript, Rutgers University.

McCawley, J. D. (1991). "Negative evidence" and the gratuitous leap from principles to parameters. *Behavioral & Brain Sciences*, *14*, 627–628.

Mintz, T. (2002). Category induction from distributional cues in an artificial language. *Memory & Cognition*, *30*, 678–686.

Morgan, J. L. (1986). *From simple input to complex grammar*. Cambridge, MA: MIT Press.

Morgan, J. L., Meier, R., & Newport, E. (1987). Structural packaging in the input to language learning. *Cognitive Psychology*, *22*, 498–550.

Morgan, J. L., Meier, R. P., & Newport, E. L. (1989). Facilitating the acquisition of syntax with cross-sentential cues to phrase structure. *Journal of Memory & Language*, *28*, 360–374.

Munson, B. (2001). Phonological pattern frequency and speech production in adults and children. *Journal of Speech, Language & Hearing Research*, *44*, 778–792.

Naigles, L. R. (1990). Children use syntax to learn verb meanings. *Journal of Child Language*, *17*, 357–374.

Naigles, L. R. (2002). Form is easy, meaning is hard: Resolving a paradox in early child language. *Cognition*, *86*, 157–199.

Naigles, L. R., & Kako, E. T. (1993). First contact in verb acquisition: Defining a role for syntax. *Child Development*, *64*, 1665–1687.

Nazzi, T., Kemler Nelson, D. G., Jusczyk, P. W., & Jusczyk, A. M. (2000). Six-month-olds' detection of clauses embedded in continuous speech: Effects of prosodic well-formedness. *Infancy*, *1*, 123–147.

Newport, E. (1991). Constraining concepts of the critical period for language. In S. Carey & R. Gelman (Eds.), *The epigenisis of mind: Essays on biology and cognition* (pp. 111–130). Hillsdale, NJ: Erlbaum.

Newport, E., Gleitman, H., & Gleitman, L. (1977). Mother, I'd rather do it myself. In C. Snow & C. Ferguson (Eds.), *Talking to children* (pp. 109–150). Cambridge, England: Cambridge University Press.

Ohala, D. K. (1999). The influence of sonority on children's cluster reductions. *Journal of Communication Disorders*, *32*, 397–422.

Peters, A. M., & Menn, L. (1993). False starts and filler syllables: Ways to learn grammatical morphemes. *Language*, *69*, 742–777.

Pine, J. M., & Martindale, H. (1996). Syntactic categories in the speech of young children: The case of the determiner. *Journal of Child Language*, *23*, 369–395.

Pine, J. M., Lieven, E. V. M., & Rowland, C. F. (1998). Comparing different models of the English verb category. *Linguistics*, *36*, 807–830.

Prince, A., & Smolenksy, P. (1993). *Optimality theory (Technical report)*. New Brunswick, NJ: Rutgers.

Prince, A., & Smolensky, P. (1997). Optimality: From neural networks to universal grammar. *Science*, *275*, 1604–1610.

Ramus, F., Hauser, M. D., Miller, C., Morris, D., & Mehler, J. (2000). Language discrimination by human newborns and by cotton-top tamarin monkeys. *Science*, *288*, 349–351.

Reinhart, T. (1983). *Anaphora and syntactic interpretation*. Croon Helm.

Roeper, T., & Williams, E. (Eds.) (1987). *Parameter setting*. Dordrecht, Netherlands: Reidel.

Rohde, D. L. T., & Plaut, D. (1999). Language acquisition in the absence of explicit negative evidence: How important is starting small? *Cognition*, *72*, 67–109.

Ross, D. S., & Newport, E. L. (1996). The development of language from non-native linguistic input. In A. Stringfellow, D. Cahana-Amitay, E. Hughes, & A. Zukowski (Eds.), *Proceedings of the 20th Annual Boston University Conference on Language Development* (Vol. 2, pp. 634–645). Somerville, MA: Cascadilla Press.

Rumelhart, D., & McClelland, J. (1987). Learning the past tenses of English verbs: Implicit rules or parallel distributed processing? In B. MacWhinney (Ed.), *Mechanisms of language acquisition* (pp. 195–248). Mahwah, NJ: Lawrence Erlbaum Associates.
Sadrzadeh, M. (2002). *Weak syllable deletion: An acoustic analysis of article omissions*. Unpublished MS Thesis, University of Arizona, Tucson.
Saffran, J. R. (2001). Words in a sea of sounds: The output of infant statistical learning. *Cognition*, *81*, 149–169.
Saffran, J. R., Aslin, R. N., & Newport, E. (1996). Statistical learning by 8-month-old infants. *Science*, *274*, 1926–1928.
Saffran, J. R., Johnson, E. K., Aslin, R. N., & Newport, E. L. (1999). Statistical learning of tone sequences by human infants and adults. *Cognition*, *70*, 27–52.
Samuelson, L. K. (2002). Statistical regularities in vocabulary guide language acquisition in connectionist models and 15–20-month-olds. *Developmental Psychology*, *38*, 1016–1037.
Santelmann, L. M., & Jusczyk, P. W. (1998). Sensitivity to discontinuous dependencies in language learners: evidence for limitations in processing space. *Cognition*, *69*, 105–134.
Senghas, A. (2003). Intergenerational influence and ontogenetic development in the emergence of spatial grammar in Nicaraguan sign language. *Cognitive Development*, *18*, 511–531.
Senghas, A., & Coppola, M. (2001). Children creating language: How Nicaraguan sign language acquired a spatial grammar. *Psychological Science*, *12*, 323–328.
Singleton, J. L., & Newport, E. L. (2004). When learners surpass their models: The acquisition of American sign language from inconsistent input. *Cognitive Psychology*, *94*(4), 370–407.
Smith, L. B. (1999). Children's noun learning: How general processes make specialized learning mechanisms. In B. MacWhinney (Ed.), *The emergence of language* (pp. 277–303). Mahwah, NJ: Erlbaum.
Smith, L. B., Jones, S. S., Landau, B., Gershkoff-Stowe, L., & Samuelson, L. K. (2002). Object name learning provides on-the-job training for attention. *Psychological Science*, *13*, 13–19.
Stager, C. L., & Werker, J. F. (1997). Infants listen for more phonetic detail in speech perception than in word-learning tasks. *Nature*, *388*, 381–382.
Tenenbaum, J. B., & Griffiths, T. L. (2001). Generalization, similarity, and Bayesian inference. *Behavioral & Brain Sciences*, *24*, 629–640.
Tomasello, M. (1992). *First verbs*. New York: Cambridge University Press.
Tomasello, M. (2000). Do young children have adult syntactic competence? *Cognition*, *74*, 209–253.
Tomasello, M. (2001). The item-based nature of children's early syntactic development. In M. Tomasello & E. Bates (Eds.), *Language development: The essential readings* (Vol. vii, pp. 169–186). Malden, MA: Blackwell Publishers.
Valian, V. V. (1986). Syntactic categories in the speech of young children. *Developmental Psychology*, *22*, 562–579.
Valian, V. V. (1991). Syntactic subjects in the early speech of American and Italian children. *Cognition*, *40*, 21–81.
Waxman, S. R., & Markow, D. B. (1995). Words as invitations to form categories: Evidence from 12- to 13-month-old infants. *Cognitive Psychology*, *29*, 257–302.
Wijnen, F. (1990). *The development of language production mechanisms*. Unpublished PhD, Catholic University of Nijmegen, Nijmegen.
Zamuner, T. (2003). *Input-based phonological acquisition*. New York, NY: Routledge.
Zamuner, T., Gerken, L. A., & Hammond, M. (2004). Phonotactic probabilities in young children's speech productions. *Journal of Child*, *31*, 515–536.
Zamuner, T., Gerken, L. A., & Hammond, M. (2005). The acquisition of phonology based on input: A closer look at the relation of cross-linguistic and child language data. *Lingua*, *115*(10), 1329–1474.

THE ROLE OF CHILDREN'S COMPETENCE EXPERIENCES IN THE SOCIALIZATION PROCESS: A DYNAMIC PROCESS FRAMEWORK FOR THE ACADEMIC ARENA

Eva M. Pomerantz, Qian Wang and Florrie Ng

DEPARTMENT OF PSYCHOLOGY, UNIVERSITY OF ILLINOIS, CHAMPAIGN, IL 61820, USA

I. Introduction

Since the middle of the 20th century, the question of how parents shape children's development has guided a wealth of theory and research. In answering this question, investigators have generally sought to confirm unidirectional

Advances in Child Development and Behavior
R Kail (Editor)

models of socialization in which parents set children's developmental course. As a consequence, central dimensions of parenting that affect diverse facets of children's psychological functioning have been identified (for reviews, see Darling and Steinberg (1993), Grusec (2002), Parke and Buriel (1998)). In an endeavor to enhance unidirectional models, investigators developed dynamic models of socialization (e.g., Grusec & Goodnow, 1994; Kuczynski, 2003; Sameroff & Macenzie, 2003). In such models, the characteristics that children bring to their interactions with parents play a critical role in the socialization process. A key undertaking to the success of dynamic models of socialization is to identify central dimensions of children that affect their socialization. The goal of this chapter is to make the case that the experiences around competence (e.g., performance in school and perceptions of one's ability to do well in school) that children bring to their interactions with parents represent one such dimension in parents' socialization of children.

Central to children's development is their attainment of a variety of competencies. Beginning at birth and continuing into adulthood, issues of competence arise on a daily basis in children's lives (see Elliot, McGregor, & Thrash, 2002). Although the arenas of most significance vary over the course of development, the attainment of competence is a significant task throughout life (Masten *et al.*, 1995; Roisman *et al.*, 2004). Consequently, children's experiences around competence have the potential to be a fundamental element of the socialization process. Much research indicates that parents shape children's competence experiences (for a review, see Pomerantz, Grolnick, and Price (in press)). However, such experiences may also shape parents' role in children's lives, thereby leading the socialization process to be a dynamic one in which parents and children mutually influence one another in a bidirectional manner. The focus of this chapter is on the academic arena, because this is an arena in which issues of competence are particularly salient. Indeed, once children enter school, they spend a large portion of each day in activities aimed at developing their academic competencies.

We delineate the role of children's competence experiences in the socialization process using the framework presented in Figure 1. A key notion underlying this framework is that important psychological resources accompany children's experiences around competence. As a consequence, such experiences are likely to be influential in the socialization process through two means. First, children's competence experiences may act as an *eliciting* force. Children's experiences around competence may convey their need for psychological resources: the more negative their experiences, the greater their need. Children's competence experiences may thus signal to parents the extent to which parents' provision of psychological resources is required, thereby leading parents to become more involved in children's lives as children's experiences become more negative. Second, children's competence experiences may act as a *moderating* force in

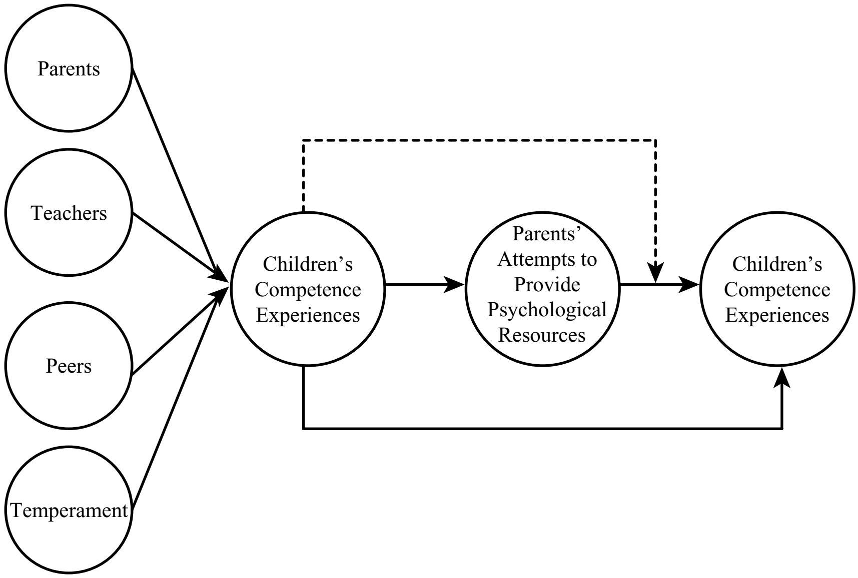

Fig. 1. A dynamic process framework: the role of children's competence experiences in the socialization process. Dashed line represents moderating effect.

the socialization process. In other words, the impact of parents' provision of psychological resources is likely to depend on children's experiences around competence. The more negative children's experiences, the more sensitive they may be to parents' provision of psychological resources because of their heightened need for such resources.

With the overarching goal of explicating the role of children's competence experiences in the socialization process in the academic arena, this chapter has three major aims. The first, which we address in Section II, is to describe children's experiences around competence in the academic arena with particular attention to the psychological resources accompanying such experiences. The second aim is to review how parents' involvement in the academic arena contributes to children's psychological resources and subsequently their competence experiences in this arena (see Section III). In Section IV, we address the third major aim by describing the dynamic socialization framework presented in Figure 1. In this context, we review the extant research in the academic arena supporting the mutual influence of children's competence experiences and parents' involvement. We conclude, in Section V, with a discussion of possible extensions of the proposed framework.

II. Children's Competence Experiences in the Academic Arena

The term "competence" has been assigned various meanings in psychology (see Ford, 1985; Garmezy & Masten, 1991; Masten *et al.*, 1995), including, but not limited to, motivational tendencies (e.g., Ford, 1985), skills and abilities (e.g., White, 1959), and performance (e.g., Masten *et al.*, 1995). We draw from Masten and colleagues' (1995) conceptualization of competence as "a pattern of effective performance in the environment." However, consistent with other investigators (e.g., Elliot, McGregor, & Thrash, 2002; White, 1959), we also include in our definition of competence, the abilities that provide the potential for such a pattern of performance. We now turn to discussing the multi-faceted nature of children's experiences around competence. We then describe the psychological resources that are associated with such experiences.

A. THE NATURE OF CHILDREN'S COMPETENCE EXPERIENCES

Children have both objective and subjective experiences around competence (see Ford, 1985). In the academic arena, children's *attainment of competence* represents a fairly objective experience in that it is often accompanied by tangible feedback. Indeed, children's attainment of competence is reflected in their accomplishment in school, as indicated by their grades for performance in class as well as their scores on achievement tests. Children's attainment of competence is generally a fairly public competence experience in that many of the significant others in children's lives are aware of it. Because teachers are involved in developing and assessing children's competence, they are intimately familiar with children's attainment (for reviews, see Brophy (1983), Jussim, Eccles, and Madon (1996)). Although often positively biased, parents are also conscious of children's attainment of competence (e.g., Miller, Manhal, & Mee, 1991; Pezdek, Berry, & Renno, 2002) as they are generally kept informed of children's progress in school, with many parents playing a role in the development of children's competence in the academic arena. Children's peers are also knowledgeable of their attainment of competence (e.g., Stipek, 1981; Stipek & Tannatt, 1984), perhaps as a consequence of the display of work in the class, teachers' public feedback, and sharing of information between peers (see Stipek & MacIver, 1989).

Children's *perceptions of competence* in the academic arena represent a more subjective experience around competence. Although children base their perceptions of their competence on their attainment, there is also considerable variation in the extent to which they do so (e.g., Cole *et al.*, 1999; Parsons, Kaczala, & Meece, 1982). Children with equivalent attainment often experience it differently. Thus, for example, some children although achieving quite high grades in school, see their competence in school in a negative light (Phillips, 1984; Phillips & Zimmerman, 1990). Given their subjective nature, perceptions of competence are a less public competence experience than that of attainment of competence, particularly when perceptions do not map onto attainment. However, teachers, parents, and peers may be aware of children's perceptions as a consequence of the types of behavior that may be fostered by children's perceptions. For example, children who perceive themselves as incompetent may hesitate to enter situations that are challenging along competence lines, often showing anxiety when faced with such situations.

Children's experiences around competence, particularly their perceptions of their competence, have been characterized as stable dimensions of children that guide cognition, affect, and behavior (e.g., Cole *et al.*, 2001; Demo & Savin-Williams, 1992; Harter, 1998). Much evidence suggests that such a characterization of children's competence experiences in the academic arena is accurate. Across the course of development, children's competence experiences are fairly stable, becoming more stable as children get older (Cole *et al.*, 2001). Over periods of less than a year up to 5 years, correlations for children's attainment of academic competence (i.e., grades and achievement test scores) generally range from 0.48 to 0.86 (e.g., Altermatt & Pomerantz, 2003; Kowaleski-Jones & Duncan, 1999; Marsh & Yeung, 1998; Pomerantz & Eaton, 2001; Steinberg, Elmen, & Mounts, 1989). Similar findings are evident for children's perceptions of their academic competence, with correlations ranging from approximately 0.29 to 0.86 (e.g., Cole *et al.*, 2001; Guay, Boggiano, & Vallerand, 2001; Marsh & Yeung, 1998; Pomerantz & Saxon, 2001). Research spanning longer durations and using multiple indices of children's competence experiences in the academic arena yields correlations of 0.43–0.54 (e.g., Kowaleski-Jones & Duncan, 1999; Masten *et al.*, 1995).

Although children's competence experiences have been characterized as stable, they have simultaneously been described as sensitive to social and cognitive influences, thereby changing over the course of development (e.g., Cooley, 1902; Eccles, 1983; Markus & Wurf, 1987; Stipek & MacIver, 1989). Indeed, several lines of evidence indicate that despite their stability, children's experiences around competence change. First, although the effect sizes for the stability of children's academic competence experiences over the course of development indicate that such experiences are stable, they are not so large as to suggest that change does not occur. Second, a wealth of research indicates that children's competence experiences in school change over the course of development in response to various social and cognitive influences, such as school transitions and the understanding of constancy (e.g., Cole *et al.*, 2001; Kowaleski-Jones & Duncan, 1999; Stipek & MacIver, 1989; Wigfield *et al.*, 1991). Third, although children may progress through development on quite stable academic competence experience trajectories, some children may be "shocked off" of these trajectories (Kowaleski-Jones & Duncan, 1999). Thus, for example, children who are on a trajectory of decreased attainment may sometimes move onto a trajectory of increased attainment.

B. PSYCHOLOGICAL RESOURCES ACCOMPANYING COMPETENCE EXPERIENCES

Children's experiences around competence play a significant role in their development for several related reasons. Beginning at birth and continuing into adulthood, children have a fundamental need to feel competent (see Deci & Ryan, 1985; Elliot, McGregor, & Thrash, 2002; White, 1959). Such a need represents a major motivation underlying much of children's behavior (see White, 1959). When it goes unfulfilled, children are likely to suffer in terms of their psychological functioning (see Deci & Ryan, 1987, 2000). In a related, but somewhat different vein, the case has been made that at each period of development, children are faced with attaining competence in salient arenas, with attainment in such arenas representing an important milestone with implications for children's psychological functioning in the next period of development (Masten *et al.*, 1995; Sroufe, Fox, & Pancake, 1983). Due in part to these issues, important psychological resources accompany children's competence experiences. Such resources shape the experiences children have around competence which then influence their subsequent psychological resources. We focus here on three key psychological resources: strategy-related, motivational, and affective. These resources may lead children to become engaged in challenging tasks in a manner that fosters positive competence experiences.

1. Strategy-Related Resources

Children with positive experiences around competence possess strategy-related resources that children with negative experiences lack. Strategy-related resources include highly automatic to highly effortful strategies that allow children to regulate their learning successfully. We focus on three such strategies. First, children's self-regulated learning strategies include planning and organizing (e.g., how to study for an exam or complete an assignment), managing time spent on learning, maintaining attention and effort, and monitoring comprehension (see Pintrich & De Groot, 1990). A second strategy-related resource is delaying immediate gratification for later benefits. This involves postponing a desire that may be instantly fulfilled (e.g., hanging out with friends) to work toward fulfilling a more important goal in the longer term (e.g., doing well in school). This often means giving up engagement in pleasurable activities for engagement in less pleasurable ones. Third, an important strategy-related resource is actively seeking assistance from teachers, peers, or others who can provide assistance when it is needed. Once children have attempted to solve problems on their own, they may ask knowledgeable others for assistance on how to best solve the problems so that they may complete the task independently.

Children's strategy-related resources may cultivate positive competence experiences because they facilitate the acquisition of the material being taught, thereby easing the attainment of competence (Alexander, Entwisle, & Dauber, 1993). Such facilitation may also affect perceptions of competence: when children experience the learning process as one that may be successfully navigated with strategy-related resources, they may develop positive perceptions of their competence. Children's competence experiences may further shape their strategy-related resources: after children have had positive experiences around competence, they may heighten their investment in attaining competence (see Tesser, 1988), which may increase the effort they exert in developing their skills (Pomerantz, Saxon, & Oishi, 2000; Pomerantz & Shim, in press).

The empirical evidence is consistent with the notion that children's competence experiences are accompanied by strategy-related resources. Children with positive competence experiences, whether it be in terms of attainment or perceptions of competence, are particularly likely to maintain their attention in class, as well as use other self-regulated learning strategies, such as planning, organizing, and comprehension monitoring (Corno, 1994; Pintrich & De Groot, 1990; Zimmerman & Martinez-Pons, 1990). High-achieving children are also less likely than their low-achieving counterparts to avoid seeking help from their peers when they need it (Ryan, Patrick, & Shim, in press). Research on the direction of effects between children's competence experiences and strategy-related resources is consistent with the notion that children's possession of such resources, both shape their competence experiences and, in turn, are shaped by them. For example, a study using a longitudinal design to examine children's math achievement over the transition to junior high school revealed a reciprocal relation between children's use of self-regulated learning strategies and their grades in school (Kenney-Benson *et al.*, 2003): children's heightened self-regulated learning strategies predict their enhanced grades 2 years later, adjusting for their earlier grades (see also Alexander, Entwisle, & Dauber, 1993). Similar effects have been found in research examining the effects of preschool children's ability to delay immediate gratification on their attainment of academic competence during adolescence (Mischel, Shoda, & Peake, 1988; Mischel, Shoda, & Rodriguez, 1989). Moreover, children's help seeking is predictive of their subsequent achievement (Ryan, Patrick, & Shim, in press). Children's grades also predict their use of self-regulated learning strategies 2 years later, even when their earlier grades have been taken into account (Kenney-Benson *et al.*, 2003).

2. Motivational Resources

Children's competence experiences are tied not only to strategy-related resources, but also to motivational resources. There are three such resources.[1] The extent of children's investment in the pursuit of competence is one motivational resource. This involves children seeing the attainment of competence as a personally important endeavor to which they are highly committed (see Pomerantz & Shim, in press). A second motivational resource is the reasons children's have for pursuing competence. Various types of reasons have been implicated as underlying children's pursuit of competence (see Eccles & Wigfield, 2002). Perhaps most central are two related sets of reasons. Children may be concerned with mastery (i.e., developing their competence) vs. performance (i.e., demonstrating competence) (Dweck & Leggett, 1988). Intrinsic reasons (i.e., motives within children, such as enjoyment) vs. extrinsic reasons (i.e., motives outside of children, such as punishment) may also underlie children's investment in the pursuit of competence (Ryan & Connell, 1989). Although not completely overlapping, mastery and intrinsic orientation are similar in their focus on the process of attaining competence, whereas performance and extrinsic orientations are similar in their focus on the outcome of doing so. Third, children's feelings of control over their attainment of competence also represent a motivational resource. When children feel in control of their attainment of competence, they experience themselves as influential in the process of attainment.

Children's possession of motivational resources is likely to promote positive experiences around competence as such resources lead children to constructively engage themselves in activities that promote positive competence experiences (see Dweck & Leggett, 1988; Eccles, Wigfield, & Schiefele, 1998; Grolnick & Slowiaczek, 1994). Such experiences are also likely to influence children's subsequent motivational resources, as they may cause the pursuit of competence to be more important, enjoyable, and under children's control (see Deci & Ryan, 1985; Elliot, McGregor, & Thrash, 2002; Skinner, Zimmer-Gembeck, & Connell, 1998; Tesser, 1988).

A wealth of evidence is consistent with the notion that children's competence experiences are accompanied by motivational resources. Children with positive competence experiences, whether it be in terms of their attainment or perceptions of competence, are more invested in the pursuit of competence (e.g., Berndt & Miller, 1990; Eccles & Wigfield, 1995; Pomerantz, Saxon, & Oishi, 2000), concerned with mastery (e.g., Grant & Dweck, 2003; Kenney-Benson *et al.*, 2003;

[1]Children's perceptions of competence have typically been considered a key motivational resource (e.g., Eccles, Wigfield, & Schiefele, 1998; Grolnick, Deci, & Ryan, 1997; Pomerantz, Grolnick, & Price, in press). However, given that in this chapter we are interested in the motivational resources accompanying children's perceptions, we do not discuss them.

Wolters, 2004), and intrinsically motivated (e.g., d'Ailly, 2003; Grolnick, Ryan, & Deci, 1991; Guay, Boggiano, & Vallerand, 2001) than are children with negative competence experiences. In addition, they feel more in control of their attainment of competence (e.g., Grolnick, Ryan, & Deci, 1991; Skinner, Zimmer-Gembeck, & Connell, 1998).

Research employing longitudinal designs is consistent with the notion that children's motivational resources shape their competence experiences which in turn contribute to their psychological resources. In terms of investment, Pomerantz and Shim (in press), for example, found that the more invested children were in academics on one day, the more positive their competence experiences the next day, even when their earlier experiences were taken into account. Moreover, children's endorsement of mastery goals predicts enhanced performance in school over time, when adjusting for children's initial performance (e.g., Dweck & Sorich, 1999; Kaplan & Maehr, 1999; Roeser, Midgeley, & Urdan, 1996). Children's intrinsic motivation also predicts their subsequent attainment and perceptions of competence (e.g., Guay, Marsh, & Boivin, 2003; Miserandino, 1996). Similar findings are evident for children's feelings of control (e.g., Skinner, Zimmer-Gembeck, & Connell, 1998). Conversely, the days on which children perceive themselves as highly competent in school are followed by increased investment in school among children over and above their prior investment (Pomerantz & Shim, in press). Although children's competence experiences do not appear to foreshadow the reasons they become invested in school (Guay, Boggiano, & Vallerand, 2001; Kenney-Benson *et al.*, 2003), the more positive children's experiences around competence, both in terms of their attainment and perceptions of competence, the more children later feel in control of their competence (e.g., Pomerantz & Saxon, 2001; Skinner, Zimmer-Gembeck, & Connell, 1998).

3. Affective Resources

Affective resources also accompany children's competence experiences. Such resources include the increased experience of positive emotions, such as joy, happiness, and pride, and the decreased experience of negative emotions, such as sadness, disappointment, and shame. Children's affective resources are likely to manifest themselves in terms of children's emotional responses to academic endeavors, such as homework and tests. Affective resources may also take the form of more general, ongoing emotional experiences, such as the lack of depression or anxiety.

Children's emotional experiences are often thought of as a by-product of their competence experiences. For example, in a number of theories of depression, children's negative perceptions of competence have been posited to be a risk factor (e.g., Abramson, Metalsky, & Alloy, 1989; Beck, 1967; Garber & Hilsman, 1992). However, emotional experiences have also been identified as a resource that can shape children's competence experiences (e.g., Cole *et al.*, 1999; Fredrickson, 2001; Pomerantz & Rudolph, 2003; Rudolph, in press). Investigators have suggested that positive affect facilitates constructive engagement with one's environment, whereas negative affect inhibits such engagement (e.g., Cacioppo, Gardner, & Bernston, 1999; Carver & Scheier, 1990). When children are engaged with their environment, they are likely to build skills that promote positive competence experiences. Moreover, children's positive emotional experiences may also lead them to view the world in a positive fashion, whereas their negative emotional experiences may foster negative world views, which may contribute to their competence experiences (Pomerantz & Rudolph, 2003). Thus, similar to children's strategy-related and motivational resources, children's affective resources may be reciprocally linked to their competence experiences, shaping these experiences and in turn being shaped by them.

A wealth of research links children's competence experiences in the academic arena to their affective resources: children with positive experiences are likely to report better emotional functioning than do their counterparts with negative experiences (for a review, see Rudolph (in press)). Evidence from longitudinal research is consistent with the notion that children's affective resources are both an antecedent and consequence of children's competence experiences. In several studies, depressive symptoms predict decreases in children's attainment of competence as manifested in their grades over time, even after adjusting for their earlier attainment (e.g., Chen & Li, 2000; Roeser, Eccles, & Sameroff, 1998). Similarly, the more depressive and anxiety symptoms children have during the elementary and junior high school years, the more children experience decrements in their perceptions of competence over time, adjusting for their earlier perceptions (e.g., Cole *et al.*, 1999; Pomerantz & Rudolph, 2003; Roeser, Eccles, & Sameroff, 1998). Importantly, children's competence experiences also predict their emotional functioning over time periods as little as one day to as long as a year later (e.g., Cole *et al.*, 1999; Dong & Pomerantz, 2004; Pomerantz & Rudolph, 2003; Roeser, Eccles, & Sameroff, 1998).

4. *Summary*

In sum, children's competence experiences are accompanied by important psychological resources. Children with positive competence experiences are likely to possess enhanced strategy-related resources, motivational resources, and affective resources. In contrast, children with negative competence experiences are deficient in such resources. The available evidence suggests that the link between children's competence experiences and their psychological resources reflects a transactional process in which children's psychological resources shape their competence experiences which in turn shape their psychological resources.

C. IMPLICATIONS FOR A DYNAMIC SOCIALIZATION PROCESS BETWEEN CHILDREN AND PARENTS

Children's experiences around competence have several key characteristics that may lead them to be instrumental to children's socialization by parents. First, children's competence experiences are stable, but changeable, dimensions of children. Given the stability of children's competence experiences, on an ongoing basis, children are likely to bring them to their interactions with parents. Thus, such experiences have the potential to shape the socialization process. However, because children's competence experiences are changeable as well, there is also the potential that over time, parents may exert an influence on such experiences. Second, children's competence experiences are significant because they reflect children's fulfillment of fundamental needs as well as their accomplishment of important developmental tasks. Moreover, children's competence experiences are accompanied by psychological resources that may be fundamental to their development in terms of their subsequent experiences around competence. As a consequence, children's competence experiences may determine their need for the psychological resources that parents have the potential to provide children. Taken together, such qualities may lead children's competence experiences to act as an eliciting force in the socialization process by influencing the extent to which parents attempt to provide children with psychological resources. Subsequently, children's competence experiences may act as a moderating force determining the effects of parents' attempts (see Figure 1).

III. Parents' Involvement in the Academic Arena

Although children's competence experiences may be a central source of psychological resources, parents may also be such a source (see Pomerantz, Grolnick, & Price, in press). By becoming involved in children's lives, parents have the potential to provide children with the psychological resources of import to the development of children's competence experiences. Drawing on several diverse lines of theory and research, Grolnick and Slowiaczek (1994) defined parents' involvement in the academic arena as parents' commitment of resources to this arena of children's life. Such involvement can take place on the school front. For example, parents may make contact with teachers, attend school open houses, and volunteer in children's classrooms. Parents' involvement can also take place on the home front. Practices in this vein include, but are not limited to, assisting children with homework, discussing children's activities at school, and responding to children's performance in school. In this section, we first delineate how parents' involvement provides children with psychological resources, reviewing the research to date on the effects of parents' involvement in the academic arena. We then make the case that how parents become involved is critical to maximizing the benefits for children.

A. PSYCHOLOGICAL RESOURCES ACCOMPANYING INVOLVEMENT

Parents' involvement in children's school lives may enhance children's competence experiences in the academic arena through the provision of strategy-related, motivational, and affective resources (see Grolnick, Kurowski, & Gurland, 1999; Pomerantz, Grolnick, & Price, in press). In terms of strategy-related resources, parents' involvement may provide them with useful information about how and what children are learning in school; such information aids parents in helping children to develop strategies that are useful to the development of competence (see Baker & Stevenson, 1986). In addition, when parents are involved in children's school lives, they may gain accurate information about children's skills and useful strategies for helping children, enabling them to assist children at a level that fosters maximal development of self-regulated learning strategies among children (see Pomerantz, Grolnick, & Price, in press). However, even when parents do not have such knowledge, their involvement may provide children with opportunities to develop such strategies because of the instruction and practice it affords children (see Senechal & LeFevre, 2002).

Parents' involvement may also cultivate motivational resources in children. First, when parents are involved in children's school lives, they may highlight the value of school to children. This allows children themselves to view school as valuable, cultivating their investment in the pursuit of competence in the academic arena (Epstein, 1988). Over time, children may internalize the value of school, so that their academic engagement is driven by intrinsic rather than extrinsic reasons (Grolnick & Slowiaczek, 1994). Second, parents' involvement in children's schooling represents an active strategy for dealing with schools and the challenges it presents. Grolnick and Slowiaczek (1994) have argued specifically that by being involved, parents are modeling an approach in which they take control of the situation, often to create positive change. Such a strategy may convey to children that they also have control over their attainment of competence.

Parents' involvement in children's school lives may afford affective resources as well. First, when parents are involved in children's school lives, they may provide children with support leading them to feel that they are worthwhile (Grolnick & Slowiaczek, 1994). Such support may not only lead children to develop positive perceptions of competence (Grolnick & Slowiaczek, 1994), but may also enhance children's emotional functioning. Second, parents' involvement may bring them closer to children as they share an important arena of children's lives with them. Such closeness may in turn lead to positive emotional functioning in children (Pomerantz, Wang, & Ng, in press). Third, parents' involvement in children's school lives may cultivate affective resources through the other resources it provides: for example, as a consequence of the strategies and motivation afforded by parents' involvement, children may experience increased positive and decreased negative emotions during their engagement in academic activities.

A wealth of research is consistent with the idea that parents' involvement in children's schooling promotes positive competence experiences among children. A number of studies reveal concurrent associations between heightened involvement on parents' part and enhanced competence experiences, both in terms of attainment and perceptions of competence, on children's part (e.g., Grolnick & Ryan, 1989; Reynolds, 1989; Stevenson & Baker, 1987). Importantly, much longitudinal research indicates that parents' involvement actually foreshadows children's competence experiences (e.g., Epstein, 1983; Gutman & Eccles, 1999; Pomerantz & Eaton, 2001; Senechal & LeFevre, 2002; Steinberg *et al.*, 1992). For example, Izzo and colleagues (1999) found that parents' involvement in elementary school children's academic lives predicts enhanced classroom behavior and school performance among children 2 years later, even when children's initial classroom behavior and school performance are taken into account. In a similar vein, longitudinal research over the transition from elementary school to junior high school demonstrates that parents' involvement is predictive of children's perceptions of competence after the transition, adjusting for children's earlier perceptions (Grolnick *et al.*, 2000).

These effects appear to be due to the psychological resources that parents' involvement provides children. In support of the provision of strategy-related resources, Hill and Craft (2003) demonstrated that parents' involvement is associated with heightened self-regulated learning strategies among children in kindergarten. In terms of motivational resources, Grolnick and Slowiaczek (1994) showed that mothers' involvement in elementary school children's schooling is associated with heightened feelings of control over the attainment of competence, which accounts in part for the effects of mothers' involvement on children's actual attainment of competence as manifested in children's grades (see also Steinberg *et al.*, 1992). Additional evidence is consistent with the notion that parents' involvement provides affective resources to children. Studying children in the late years of elementary school and the early years of junior high school, Pomerantz, Ng, and Wang (2004) found that the more involved mothers were in providing assistance with children's homework, the lower children's negative emotional functioning (i.e., negative emotions, depressive symptoms, and anxiety symptoms) 6 months later, taking into account children's earlier negative emotional functioning. In sum, parents' involvement appears to enhance children's competence experiences. Notably, evidence suggests that this is due to the strategy-related, motivational, and affective resources that parents' involvement can provide.

B. CENTRAL DIMENSIONS OF INVOLVEMENT

Although becoming involved in children's school lives is an important first step to providing children with psychological resources that enhance competence experiences, *how* parents become involved is likely to determine just how beneficial their involvement is (see Grolnick, 2003; Pomerantz, Grolnick, & Price, in press). Indeed, Darling and Steinberg (1993) have proposed that the effects on children of parents' practices are determined by the style with which such practices are used. Drawing from this perspective, we focus on three dimensions of parents' involvement that have emerged in theory and research on parenting as important in terms of providing children with psychological resources (see Pomerantz, Grolnick, & Price, in press). The degree to which parents' involvement is characterized by autonomy support vs. control, focused on the process of learning vs. the ability or performance of the child, and accompanied by positive vs. negative affect.

1. Autonomy-Support vs. Control

Early on in the study of parents' socialization of children, parents' use of autonomy support vs. control was identified as a key dimension. Hence, an extensive body of theory and research has focused on multiple forms of parental autonomy support and control (see Pomerantz and Ruble (1998a), Rollins and Thomas (1979), Steinberg (1990), for reviews). Here, we draw on Deci and Ryan's (1987) self-determination theory (see also Grolnick & Ryan, 1989; Ng, Kenney-Benson, & Pomerantz, 2004; Pomerantz & Eaton, 2001), in which autonomy support is defined as allowing children to explore their own environment, initiate their own behavior, and take an active role in solving their own problems. Controlling behavior, in contrast, involves the exertion of pressure by parents to channel children toward particular outcomes, such as doing well in school.

When parents become involved in their children's school lives in an autonomy-supportive rather than controlling manner, they may provide children with the experience of solving challenges on their own which may aid children in developing strategy-related resources (e.g., Ng, Kenney-Benson, & Pomerantz, 2004; Nolen-Hoeksema *et al.*, 1995; Pomerantz & Ruble, 1998b). In addition, when parents support children's autonomy rather than attempting to control them, they allow children to take initiative, which cultivates motivational resources as it promotes the feeling in children that they are engaging in their behavior for intrinsic rather than extrinsic reasons (e.g., Deci & Ryan, 1985; Grolnick, Deci & Ryan, 1997).

A fairly large body of research using a variety of methods suggests that parents' autonomy support enhances children's competence experiences, whereas parents' control detracts from such experiences (for reviews, see Grolnick (2003), Pomerantz, Grolnick, and Price (in press)). These effects appear to begin early on in children's lives and extend into the adolescent years (e.g., Grolnick & Ryan, 1989; Ng, Kenney-Benson, & Pomerantz, 2004; Steinberg *et al.*, 1992). For example, mothers' controlling behavior, particularly appeals to authority, with 4-year old children is associated not only with children demonstrating poor school readiness a year or two later, but also with children doing poorly in school 8 years later (Hess & McDevitt, 1984). Moreover, Grolnick *et al.* (2002) created a situation in the laboratory designed to simulate the homework situation. In this context, the more autonomy-supportive and less controlling mothers were, the better their elementary school children's performance in the laboratory. A number of studies have also shown that the more autonomy-supportive and less controlling parents are, the more positive children are in their perceptions of academic competence (e.g., Grolnick & Ryan, 1989; Steinberg *et al.*, 1994; Wagner & Phillips, 1992).

Apparently, a key reason that parents' use of autonomy support rather than control confers benefits on children in terms of their competence experiences is because it provides psychological resources. In terms of strategy-related resources, both concurrent and longitudinal research indicate that when parents are autonomy-supportive rather than controlling, children are particularly likely to use self-regulated learning strategies and to delay gratification (e.g., Houck & Lecuyer-Maus, 2004; Silverman & Ippolito, 1995; Steinberg, Elmen, & Mounts, 1989). Several studies provide support for the idea that autonomy-supportive vs. controlling parenting provides motivational resources (e.g., Grolnick & Ryan, 1989; Nolen-Hoeksema *et al.*, 1995). For example, children of parents who are autonomy-supportive rather than controlling are intrinsically rather than extrinsically motivated in the school context (e.g., d'Ailly, 2003; Ginsburg & Bronstein, 1993; Grolnick & Ryan, 1989). Importantly, there is evidence directly indicating that such heightened motivational resources underlie the tendency for parents' orientation toward autonomy support rather than control to promote the attainment of competence among children (Grolnick, Deci & Ryan, 1991; Steinberg, Elmen, & Mounts, 1989).

2. *Process- vs. Person-Focus*

Several lines of theory and research suggest that the extent to which parents are focused on the process of learning rather than the innate ability or performance of children is an important dimension of parents' involvement in children's schooling (e.g., Dweck & Lennon, 2001; Gottfried, Fleming, & Gottfried, 1994; Hokoda & Fincham, 1995). A process-focused orientation emphasizes the importance and pleasure of effort and learning (Gottfried, Fleming, & Gottfried, 1994;

Kamins & Dweck, 1999; Mueller & Dweck, 1998). A person-focus, in contrast, emphasizes the importance of stable attributes, such as intelligence and performance.

When parents adopt a process- rather than a person-focus in the context of their involvement in children's school lives, they may afford strategy-related resources as they highlight the importance of effort. Moreover, in the context of focusing on the process of learning, parents emphasize the significance of effort and learning in the achievement process, which may lead children to develop motivational resources such as a mastery orientation (Hokoda & Fincham, 1995; Mueller & Dweck, 1998). Parents may also afford affective resources by attuning children to the pleasure of learning (Pomerantz, Ng, & Wang, 2004).

A growing body of research indicates that parents' process- vs. person-focused involvement provides children with psychological resources that enhance their competence experiences. Dweck and colleagues (Kamins & Dweck, 1999; Mueller & Dweck, 1998) have looked at the effects of process- and person-focused practices by manipulating the type of feedback children are given by a previously unknown adult. Children given process-focused feedback are likely to perform better in the face of failure, perceive their competence more positively, adopt mastery over performance goals to a greater extent, and express more positive affect than are children given person-focused feedback. Parents' use of process- and person-focused practices have a similar effect. Using observational methods in the context of a laboratory task with qualities similar to that of homework, Hokoda and Fincham (1995) found that mothers who responded to their elementary school children's performance-oriented behavior (e.g., concentrating on how much time is left) with process-focused practices ("That's okay; you did your best.") were particularly likely to have mastery-oriented children (see also Gottfried, Fleming, & Gottfried, 1994). Research in which mothers reported daily on their responses to their elementary school children's academic successes indicates that when mothers refrain from using person-focused praise, 6 months later, children embrace challenging tasks (Kempner & Pomerantz, 2003).[2]

3. *Positive vs. Negative Affect*

As Dix (1991) has emphasized, parenting is an inherently affective endeavor (see also Larson & Gillman, 1999). This may be particularly true of parents' involvement in children's schooling. On the positive side, parents' involvement may afford a structured context in which to spend time with children, gain

[2]Kelley, Brownwell, and Campbell (2000) found no evidence of negative effects of mothers' use of person-focused praise in the laboratory on 2-year-olds' mastery motivation. This may be because children of this young age do not yet have a full understanding of ability and effort (see Dweck, 2002).

knowledge about a significant area of children's lives, and work together with children to overcome obstacles. Thus, when involved in children's schooling, many parents may attempt to maintain positive affect by making their interactions with children fun, loving, and supportive. On the negative side, children often experience negative affect while working on academic tasks, such as homework (Fuligni, Yip, & Tseng, 2002; Leone & Richards, 1989). This may lead parents to experience negative affect which may manifest itself in irritation and annoyance as well as more extreme forms such as hostility and criticism. When parents accompany their involvement in children's school lives with more positive than negative affect they may afford motivational resources by conveying to children that although schoolwork can be frustrating, it is an enjoyable endeavor (see Estrada *et al.*, 1987; Hokoda & Fincham, 1995; Nolen-Hoeksema *et al.*, 1995; Pomerantz, Wang, & Ng, in press). Parents' positive affect may also signal their support of children during times of difficulty, enabling children to confront challenge constructively. In addition, parents may directly transmit their affect to children, thereby providing affective resources (Larson & Gillman, 1999; Pomerantz, Wang, & Ng, in press).

Relatively little research has examined the link between parents' affect and children's competence experiences in the academic arena; however, that which exists is consistent with the idea that parents' affect plays a role in such experiences. Mothers whose relationship with their children is characterized by positive affective qualities when children are of preschool age have children with enhanced school readiness at 5–6 years of age and heightened achievement at 12 years of age (Estrada *et al.*, 1987). This may be because the affective nature of parents' involvement influences children's motivational and affective resources. In laboratory research in which children just entering elementary school and their mothers worked on an unsolvable task, when mothers expressed negative affect (e.g., hostility and criticism) toward their children during this task, children were less mastery-oriented in the face of challenge in the lab and in school (Nolen-Hoeksema *et al.*, 1995). Hokoda and Fincham (1995) found similar effects of mothers' negative affect (e.g., pouting and anger) in their laboratory research with elementary school children. They also demonstrated that children with mothers who responded to particularly difficult tasks with positive affect (e.g., enjoyment and laughter) exhibited less of a helpless orientation than did children of mothers who failed to respond with such affect. Using mothers' daily reports of their affect, Pomerantz, Wang, and Ng (in press) showed that when mothers experienced high negative affect (i.e., irritation and annoyance) and low positive affect (i.e., fun and love) while assisting children with homework, children reported dampened mastery orientation, intrinsic motivation, and positive emotional functioning 6 months later. Notably, these effects were unique to mothers' affect while assisting with homework and did not stem from their affect in other contexts.

4. *Summary*

In sum, parents' involvement in children's school lives appears to serve as a significant source of psychological resources for children, with how parents become involved being of much import. Parents' involvement is particularly likely to afford children strategy-related, motivational, and affective resources when it is autonomy-supportive, process-focused, or affectively positive. However, it may actually undermine such resources among children if it is controlling, person-focused, or affectively negative.

C. IMPLICATIONS FOR A DYNAMIC SOCIALIZATION PROCESS BETWEEN CHILDREN AND PARENTS

Given that children's competence experiences are accompanied by the same psychological resources that parents' involvement may influence, such experiences likely shape how involved parents become in children's lives, as well as the impact of parents' involvement on children's subsequent competence experiences. Parents may use children's competence experiences as cues for the extent to which they need to provide children with psychological resources. Consequently, such experiences may act to elicit parents' involvement. Children's competence experiences may also serve as a moderating force: children with negative competence experiences may be particularly sensitive to the quality of parents' involvement because of their need for the psychological resources that such involvement can provide (see Figure 1).

IV. A Dynamic Socialization Framework for the Academic Arena

Drawing from the theory and research on children's competence experiences and parents' involvement, we delineate how the two work together in the socialization process in the academic arena. As is evident from our review of the effects of parents' involvement, children's experiences around competence are likely to be influenced by parents' involvement (for other parent influences, see Halle, Kurtz-Costes, and Mahoney (1997), Jodl *et al.* (2001)). However, forces outside the home shape such experiences as well. Indeed, there is much evidence that significant others in children's lives, such as teachers (e.g., Madon, Jussim, & Eccles, 1997; Pintrich & Blumenfeld, 1985) and peers (e.g., Altermatt *et al.*, 2002; Ryan, 2001), play a role in children's competence experiences. Moreover, children's biological predispositions may be influential in the development of such experiences. Thus, children enter into their interactions with parents with established competence experiences. In this section, we make the case that such experiences play a role in two stages of the socialization process: they act initially as a force that triggers parents' attempts to provide psychological resources and subsequently as a force that moderates the effects of such attempts.

A. CHILDREN'S COMPETENCE EXPERIENCES AS AN ELICITING FORCE

With the movement away from unidirectional models of socialization, investigators have focused on how children's characteristics shape the role parents take in children's lives (e.g., Belsky, 1984; Plomin, 1994; Scarr, 1992), with empirical evidence supporting the notion that such characteristics are influential (for a review, see Sameroff and Macenzie (2003)). Although several investigators have argued that the parenting elicited by children's characteristics acts to maintain children's attributes (e.g., Lytton, 1990; Scarr, 1992), others have suggested that such parenting may also act to produce change in children (e.g., Eder & Mangelsdorf, 1997; Grusec, 2002; Hart *et al.*, 1997). In line with the transactional perspective, as shown in Figure 1, a key component of the dynamic framework of socialization we present is that children's competence experiences play a role in determining the extent to which parents attempt to provide children with psychological resources—that is, the extent to which they become involved in children's lives. Children's competence experiences are likely to convey children's need for psychological resources. They may thus signal to parents the extent to which their provision of such resources is required by children, thereby leading parents to become more involved in children's lives as children's experiences become more negative. Unfortunately, parents may not always become involved in a manner that actually provides resources because they may become anxious, and even frustrated and angry, when children have negative competence experiences; such negative emotions may lead parents to become involved in a controlling, performance-oriented, or affectively negative manner (see Pomerantz & Eaton, 2001; Pomerantz, Wang, & Ng, in press). Thus, children's subsequent competence experiences may either be enhanced or undermined, depending in part on how parents become involved.

In line with the notion that parents may heighten their involvement in children's schooling when children have negative competence experiences in the academic arena, several concurrent investigations reveal that parents are more likely to assist low- than high-achieving children with homework (e.g., Chen & Stevenson, 1989; Cooper, 1989; Levin *et al.*, 1997). Moreover, mothers offer heightened assistance with homework on the days they perceive children as frustrated with their homework (Pomerantz, Wang, & Ng, in press). Parents also have heightened contact with teachers when children are doing poorly in school (Izzo *et al.*, 1999). Although such associations may reflect the negative effects of parents' involvement on children's attainment of competence, research suggests this is unlikely. Pomerantz and Eaton (2001) showed that elementary school children's poor performance in school foreshadowed mothers' frequent assistance with homework when children did not request it 6 months later. Mothers apparently increased their assistance because they picked up on cues from children that they lacked psychological resources. The effect over time of children's attainment of competence on mothers' assistance with homework was accounted for in part by the tendency of low-achieving children to experience uncertainty about how to do well in school. Importantly, when children's initial attainment of competence was taken into account, mothers' assistance with homework predicted an increase in attainment over 6 months (Pomerantz & Eaton, 2001).

Although evidence suggests that children's attainment of competence elicits involvement from parents, there is little evidence that children's perceptions of competence do so. Research to date has not revealed a substantial link between children's perceptions of competence and mothers' provision of assistance with homework (Pomerantz, 2001; Pomerantz, Ng, & Wang, 2004). Because children's attainment of competence represents an objective experience, whereas their perceptions of competence represent a subjective experience, parents may be more aware of the former than the latter. Parents may also respond differently to the two competence experiences. Children's lack of attainment may be met with heightened involvement in tasks such as homework as parents attempt to provide children with effective skills to remedy the situation. However, children's negative perceptions of competence may be met with heightened involvement on a more emotional level (e.g., providing encouragement) as parents attempt to enhance children's feelings about themselves.

B. CHILDREN'S COMPETENCE EXPERIENCES AS A MODERATING FORCE

In line with interactional models of socialization in which the effects of parents on children are influenced by what children themselves bring to their interactions with parents (e.g., Grusec, 2002; Kochanska, 1993), after parents become involved in children's school lives, children's experiences around competence may serve to moderate the effects of how parents become involved. In essence, the impact of parents' provision of psychological resources is likely to depend on children's competence experiences. The more negative children's experiences, the more sensitive they may be to parents' provision of psychological resources because of their heightened need for such resources. Consequently, they derive particular benefit when parents become involved in their school lives in an autonomy-supportive, process-oriented, or affectively positive manner. Under such conditions, the psychological resources, and consequently the subsequent competence experiences, of children with negative competence experiences may profit. Unfortunately, however, children with negative experiences may be particularly vulnerable when parents become involved in a controlling, person-oriented, or affectively negative manner because such involvement deprives these children of the resources of which they are in much need. In contrast, children with positive competence experience may not be as sensitive to how parents become involved because of their possession of psychological resources.

Several longitudinal studies are consistent with the notion that children with negative competence experiences are more sensitive than are children with positive experiences to how parents become involved. Children with negative experiences are particularly likely to benefit when parents become involved in their school lives in an autonomy-supportive rather than a controlling manner. In one study, mothers' responses to their elementary and middle school children's failure in a variety of areas, including the academic, were assessed with a daily checklist (Ng, Kenney-Benson, & Pomerantz, 2004, Study 2). Mothers' autonomy-supportive responses (i.e., discussing children's failure with them) predicted increased performance and their controlling responses (i.e., reprimanding children for their failure or punishing children for their failure) predicted decreased performance the next day and 6 months later more for low- than high-achieving children. In a second study, mothers' involvement with their elementary school children was observed in the laboratory in the context of a challenging task designed to reflect the homework situation (Ng, Kenney-Benson, & Pomerantz, 2004, Study 1). Over the course of their interactions with children, mothers' autonomy support predicted enhanced subsequent performance and their control predicted diminished subsequent engagement more for low- than high-achieving children (see also Pomerantz, 2001).

A similar pattern is evident for mothers' process-focused involvement in children's schooling. In a daily telephone interview, mothers' process orientation (e.g., encouraging children to understand their work) was examined in the context of their assistance with elementary and middle school children's homework (Pomerantz, Ng, & Wang, 2004). When mothers adopted a process orientation, children with negative perceptions of their academic competence were more likely than children with positive perceptions to benefit in terms of their perceptions of competence 6 months later. Consistent with the notion that parents' process-oriented involvement fosters positive competence experiences among children because it provides them with psychological resources, mothers' process orientation was linked to children's mastery orientation and positive emotional functioning over time. As anticipated, only children initially perceiving themselves as lacking competence were likely to benefit in terms of these two types of psychological resources—so much so that when mothers were highly process-oriented, such children were no more likely to lack motivational or affective resources than were their counterparts with positive perceptions of competence.

When mothers' involvement in children's school lives is accompanied by positive affect, children with negative competence experiences also appear to be particularly likely to benefit. Using the daily interview method, Pomerantz, Wang, and Ng (in press) focused on mothers' affect on the days their elementary and middle school children had homework. In this study, the focus was on children's competence experiences as manifested in helplessness (i.e., frustration and giving up) in competing homework. As noted earlier, mothers were particularly likely to assist children who demonstrated helplessness with their homework (see also Pomerantz & Eaton, 2001). However, this heightened assistance only benefited these children in terms of their mastery orientation when mothers accompanied it with positive affect. Specifically, when mothers' affect was positive on days they were involved in children's homework, children demonstrating high levels of helplessness while completing homework experienced heightened mastery orientation over the course of 6 months to a greater extent than did children demonstrating low levels of helplessness. The benefit was to such a great extent that when mothers were particularly high in positive affect, helpless children's mastery orientation was not lower than that of children who were not helpless. Mothers' heightened positive affect was also particularly beneficial to helpless children when children had homework, but mothers refrained from assisting: helpless children were intrinsically motivated and functioning positively emotionally 6 months later when their mothers maintained positive affect on days they did not provide assistance, perhaps because they were positively supporting such children's autonomy.

C. SUMMARY AND FUTURE DIRECTIONS

In sum, the research to date supports the dynamic socialization framework we have proposed (see Figure 1). For one, parents become more involved in their children's academic lives when their children have difficulty in school. In addition, after parents are involved in children's lives, children with negative competence experiences are particularly sensitive to the nature of this involvement. Such children are highly likely to experience benefits when parents are autonomy-supportive, process-focused, and affectively positive, but highly likely to experience costs when parents are not involved in such ways.

There are several key directions for future research. First, it will be important to fully capture the dynamic nature of socialization with designs that can simultaneously examine the eliciting and moderating role of children's competence experiences. Second, more attention to the mechanisms by which children's competence experiences shape and are shaped by parents' involvement is necessary. Third, the use of within-family designs will be critical to fully teasing apart the effects of children and parents in the dynamic socialization process we have proposed. Fourth, attention to modalities other than parents' involvement (e.g., their perceptions of their children's competence) by which parents provide children with psychological resources is of import.

Finally, a significant endeavor will be to identify the times when parents' provision of psychological resources is important for children's positive competence experiences. For example, during periods of particular difficulty and uncertainty, parents' provision of psychological resources may benefit children regardless of their competence experiences because of an increased need among children for psychological resources.

V. Concluding Remarks

There is an increasing trend for theories of socialization to be highly contextualized in that they focus on specific arenas, socialization agents, or types of functioning (see Maccoby, 1992; O'Connor, 2002). This trend does justice to the complex nature of the socialization process. Moreover, it is quite functional as it yields concrete recommendations regarding positive parenting to address specific problems children may have. However, creating contextualized theories may also impede progress by leading key principles underlying the socialization process in general to be overlooked (for a similar argument in social psychology, see Higgins (2004)). Although the dynamic framework presented here was created to elucidate *parents'* role in children's development in the *academic* arena, it has the potential to be useful to understanding socialization more broadly. Indeed, children's competence experiences may represent a central dimension of the socialization process that cuts across arenas as well as agents of socialization. In this concluding section, we first discuss the potential for children's competence experiences to play a role in the socialization process in arenas other than the academic one. We then focus on the possibility that children's competence experiences may influence how socialization unfolds with agents of socialization other than parents.

A. BEYOND THE ACADEMIC ARENA

The academic arena is not the only one in which children's competence experiences are salient. In fact, over the course of development, such experiences are quite salient in other arenas (Elliot and Dweck, in press; Masten *et al.*, 1995). Moreover, the same psychological resources that accompany children's competence experiences in the academic arena accompany their experiences in other arenas (for a review of the social arena, see Rudolph and Asher (2000)). In a parallel fashion, parents' involvement in children's lives in arenas other than the academic is likely to provide psychological resources, particularly when it is autonomy-supportive rather than controlling, process- rather than person-focused, and characterized by positive rather than negative affect (see Eisenberg & Valiente, 2002; Grolnick & Farkas, 2002; Ladd & Pettit, 2002).

As a consequence, the framework we have presented (see Figure 1) may apply to socialization in arenas beyond the academic. In several arenas of import to children's development, research is consistent with the ideas we have put forth here. We focus on two such arenas: children's social relations and their emotional regulation.

The social arena is a key context for the development of competence, particularly after children enter school (e.g., Masten *et al.*, 1995), but even before this time (e.g., Ladd, 2003). As a consequence, children's competence experiences in this arena are likely to play a role in the socialization process. Although longitudinal research is still needed, concurrent research is consistent with the idea that children's social competence may act as an eliciting force. Indeed, as is the case when children lack academic competence, when children lack social competence, parents apparently attempt to provide them with psychological resources. For example, Mize, Pettit, and Brown (1995) found that mothers were particularly involved in supervising the play of their children and peers when they perceived their children as lacking social competence. In the equally important arena of children's emotional regulation, there is some suggestion that at least among younger children, parents attempt to provide psychological resources to children who they perceive as having poor emotional regulation abilities (e.g., Casey & Fuller, 1994; Fabes *et al.*, 1994). However, more often than not such involvement is controlling rather than autonomy-supportive and laden with negative rather than positive affect (see Grolnick & Farkas, 2002).

Children's competence experiences in the social arena also act as a moderating force in the socialization process. In a framework similar to ours, Schwartz and colleagues (2000) proposed that children's positive interactions with friends might buffer the effects of negative parenting on subsequent victimization by peers by enhancing children's self-regulation skills and facilitating the development of other core social skills. In two longitudinal studies following children beginning in the preschool and kindergarten years into the middle elementary school years, these investigators tested their hypothesis that children's friendships (a competence experience in the social arena) would moderate the influence of parents' control (i.e., harsh discipline) and negative affect (i.e., hostility). Using structured interview and observational techniques to assess mothers' control and negative affect, across the two studies, the negative effects of these aspects of parenting on children's subsequent victimization were stronger among children with negative social competence experiences—that is, with few reciprocated friendships—than among children with positive social competence experiences.

Several studies suggest that children's competence at regulating their emotions moderates the role that parents play in their subsequent development. In general, this research suggests that children, particularly boys, who have problems regulating their emotions are more sensitive to how their parents become involved in their lives than are their counterparts who do not have such problems. In longitudinal research on boys during the toddler years, Belsky and colleagues (Belsky, Hsieh, & Crnic, 1998; Park *et al.*, 1997) found that negative parenting (e.g., intrusiveness and negative affect) as assessed in observations around dinner time is more predictive over time of subsequent externalizing and inhibition problems among boys high in negative emotionality (e.g., fearfulness and negative emotional responses to frustrating or limiting tasks) than among boys low in negative emotionality (see also Rubin *et al.*, 1998). Moreover, when children in the early elementary school years view their parents' as controlling, those with emotional regulation problems—that is, with a high susceptibility to anger and distress—are more likely to have internalizing symptoms than are their counterparts without such problems (Morris *et al.*, 2002).

B. BEYOND PARENTS

Although parents are clearly important in the socialization process, as noted earlier, significant others, such as teachers and peers, outside of the family are also important. These agents of socialization may be just as likely as parents to provide children with psychological resources. As a consequence, children's competence experiences may influence how they are shaped by the significant others outside of the family. Because there has been considerably little research on the role of significant others outside the family in children's socialization in the academic arena, tests of dynamic models of socialization have been limited, particularly in terms of examining children as an eliciting force. However, some evidence indicates that children's competence experiences may shape involvement among teachers. For example, although not all teachers report responding to children having writing difficulties during elementary school with increased involvement, the majority of teachers report doing so (Graham *et al.*, 2003).

Teachers' involvement may provide children with psychological resources in much the same way that parents' involvement does. Indeed, research suggests that children benefit in terms of their motivational resources, as well as their competence experiences, when teachers are autonomy-supportive rather than controlling (e.g., Deci *et al.*, 1981; Wentzel, 2002). Although no research to date has examined whether this effect is moderated by children's competence experiences, there is some evidence that children's motivational resources may serve as moderators. Elementary school children's perceptions of their teachers as autonomy-supportive are more likely to predict enhanced feelings of competence a year later among extrinsically motivated children than among intrinsically motivated children (Guay, Boggiano, & Vallerand, 2001). In a similar vein, among college students, the positive effects on performance of perceiving instructors as autonomy-supportive are stronger for students pursuing their studies for controlled reasons than for students pursuing their studies for autonomous reasons (Black & Deci, 2000). Another way that both parents and teachers may provide children with competence resources is through their perceptions of children's competence (Pomerantz, Grolnick, & Price, in press). The effects of teachers' perceptions of children's competence on children's subsequent performance are moderated by children's competence attainment: such perceptions have stronger effects among initially low-achieving children than among initially high-achieving children (Madon, Jussim, & Eccles, 1997).

Although the ways in which peers may provide children with psychological resources may differ from that of parents and teachers, aspects of the framework presented here may also be useful in understanding the peer socialization process. Unlike parents and teachers, peers are often not in the role of looking out for children's well-being. Thus, children's competence experiences may not serve as an eliciting force in their interactions with peers. However, such experiences may serve as a moderating force because children with negative competence experiences may be particularly likely to benefit from the psychological resources that their peers provide. Future research examining this possibility could take several forms. One is that the tendency for children to reap benefits both academically and emotionally from having friendships (e.g., Ladd, Birch, & Buhs, 1999) may be particularly pronounced among children who lack academic or social competence. Another, albeit related, possibility is that children with negative competence experiences in the academic arena may be particularly likely to benefit from supportive interactions with peers around schoolwork (e.g., the provision of help in class and doing homework with friends).

C. CONCLUSIONS

As investigators turn their attention from testing unidirectional models of parents' socialization of children to testing bidirectional models characterized by dynamic interplay between parents and children, a key goal is to identify central dimensions of children that affect the socialization process. Although contextualized models of socialization are of import to understanding children's development, the identification of general principles is fundamental. The framework presented here, along with the accompanying research, suggests that children's competence experiences may represent a central dimension of children that affects the socialization process not only in the academic arena, but also in other arenas. Indeed, a general principle may be that the need for psychological resources that children bring to their interactions with others has a significant influence in determining the course of the socialization process.

ACKNOWLEDGEMENTS

Preparation of this chapter was supported by a grant from the National Institute of Mental Health (R01 MH57505). While writing this chapter, the first author was a visiting scholar at the School of Psychology at Victoria University, Wellington, New Zealand. We are grateful for the constructive feedback provided by Liz Moorman and Karen Rudolph.

REFERENCES

Abramson, L. Y., Metalsky, G. I., & Alloy, L. B. (1989). Hopelessness depression: A theory-based subtype of depression. *Psychological Review*, *96*, 358–372.

Alexander, K. L., Entwisle, D. R., & Dauber, S. L. (1993). First-grade classroom behavior: Its short- and long-term consequences for school performance. *Child Development*, *64*, 801–814.

Altermatt, E. R., & Pomerantz, E. M. (2003). The development of competence-related and motivational beliefs: An investigation of similarity and influence among friends. *Journal of Educational Psychology*, *95*, 111–123.

Altermatt, E. R., Pomerantz, E. M., Ruble, D. N., Frey, K. S., & Greulich, F. (2002). Predicting changes in children's self-perceptions of academic competence: A naturalistic examination of evaluative discourse among classmates. *Developmental Psychology*, *38*, 903–917.

Baker, D. P., & Stevenson, H. W. (1986). Mothers' strategies for children's school achievement: Managing the transition to high school. *Sociology of Education*, *59*, 156–166.

Beck, A. T. (1967). *Depression: Causes and treatment*. Philadelphia, PA: University of Pennsylvania.

Belsky, J. (1984). The determinants of parenting: A process model. *Child Development*, *55*, 83–96.

Belsky, J., Hsieh, K., & Crnic, K. (1998). Mothering, fathering, and infant negativity as antecedents of boys' externalizing problems and inhibition at age 3 years: Differential susceptibility to rearing experience? *Development and Psychopathology*, *10*, 301–319.

Berndt, T. J., & Miller, K. E. (1990). Expectancies, values, and achievement in junior high school. *Journal of Educational Psychology*, *82*, 319–326.

Black, A. E., & Deci, E. L. (2000). The effects of instructors' autonomy support and students' autonomous motivation on learning organic chemistry: A self-determination theory perspective. *Science Education*, *84*, 740–756.

Brophy, J. (1983). Research on the self-fulfilling prophecy and teacher expectations. *Journal of Educational Psychology*, *75*, 631–661.

Cacioppo, J. T., Gardner, W. L., & Bernston, G. G. (1999). The affect system has parallel and integrative processing components: Form follows function. *Journal of Personality and Social Psychology*, *76*, 839–855.

Carver, C. S., & Scheier, M. F. (1990). Origins and functions of positive and negative affect: A control-process view. *Psychological Review*, *97*, 19–35.

Casey, R. J., & Fuller, L. L. (1994). Maternal regulation of children's emotions. *Journal of Nonverbal Behavior*, *18*, 57–89.

Chen, X., & Li, B. (2000). Depressed mood in Chinese children: Developmental significance for social and school adjustment. *International Journal of Behavioral Development*, *24*, 472–479.

Chen, C., & Stevenson, H. W. (1989). Homework: A cross-cultural examination. *Child Development*, *60*, 551–561.

Cole, D. A., Martin, J. M., Peeke, L. A., Seroczynski, A. D., & Fier, J. (1999). Children's over- and underestimation of academic competence: A longitudinal study of gender differences, depression, and anxiety. *Child Development*, *70*, 459–473.

Cole, D. A., Maxwell, S. E., Martin, J. M., Peeke, L., Seroczynski, A. D., Tram, J. M., Hoffman, K. B., Ruiz, M. D., Jacquez, F., & Maschman, T. (2001). The development of multiple domains of child and adolescent self-concept: A cohort sequential longitudinal design. *Child Development*, *72*, 1723–1746.

Cooley, C. H. (1902). *Human nature and the social order*. New York: Charles Schribner's Sons.

Cooper, H. (1989). *Homework*. White Plains, NY: Longman.

Corno, L. (1994). Student volition and education: Outcomes, influences, and practices. In D. H. Schunk & B. J. Zimmerman (Eds.), *Self-regulation of learning and performance: Issues and educational applications* (pp. 229–251). Hillsdale, NJ: Erlbaum.

d'Ailly, H. (2003). Children's autonomy and perceived control in learning: A model of motivation and achievement in Taiwan. *Journal of Educational Psychology*, *95*, 84–96.

Darling, N., & Steinberg, L. (1993). Parenting style as context: An integrative model. *Psychological Bulletin*, *113*, 487–496.

Deci, E. L., & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior*. New York: Plenum Press.

Deci, E. L., & Ryan, R. M. (1987). The support of autonomy and the control of behavior. *Journal of Personality and Social Psychology*, *53*, 1024–1037.

Deci, E. L., & Ryan, R. M. (2000). The "what" and "why" of goal pursuits: Human needs and the self-determination of behavior. *Psychological Inquiry*, *11*, 227–268.

Deci, E. L., Schwartz, A. J., Sheinman, L., & Ryan, R. M. (1981). An instrument to assess adults' orientations toward control versus autonomy with children: Reflections on intrinsic motivation and perceived competence. *Journal of Educational Psychology*, *73*, 642–650.

Demo, D. H., & Savin-Williams, R. C. (1992). Self-concept stability and change during adolescence. In R. P. Lipka & T. M. Brinthaupt (Eds.), *Self-perspectives across the life span* (pp. 116–150). Albany, NY: State University of New York Press.

Dix, T. (1991). The affective organization of parenting: Adaptive and maladaptive processes. *Psychological Bulletin*, *110*, 3–25.

Dong, W., & Pomerantz, E. M. (2004). *Affect in the academic context: Transactions between children's daily functioning around academics and their affective experience*. Unpublished manuscript.

Dweck, C. S. (2002). The development of ability conceptions. In A. Wigfield & J. S. Eccles (Eds.), *Development of achievement motivation* (pp. 57–88). San Diego, CA: Academic Press.

Dweck, C. S., & Leggett, E. L. (1988). A social-cognitive approach to motivation and personality. *Psychological Review*, *95*, 256–273.

Dweck, C. S., & Lennon, C. (2001). *Person vs. process focused parenting: Impact on achievement motivation*. Paper presented at the Society for Research on Child Development, Minneapolis, MN.

Dweck, C. S., & Sorich, L. (1999). Mastery-oriented thinking. In C. R. Snyder (Ed.), *Coping* (pp. 232–251). New York: Oxford University Press.

Eccles, J. S. (1983). Expectancies, values and academic behaviors. In J. T. Spence (Ed.), *Achievement and achievement motives* (pp. 75–146). San Francisco: W.H. Freeman.

Eccles, J. S., & Wigfield, A. (1995). In the mind of the actor: The structure of adolescents' achievement task values and expectancy-related beliefs. *Personality and Social Psychology Bulletin*, *21*, 215–225.

Eccles, J. S., & Wigfield, A. (2002). Motivational beliefs, values, and goals. *Annual Review of Psychology*, *53*, 109–132.

Eccles, J. S., Wigfield, A., & Schiefele, U. (1998). Motivation to succeed. In N. Eisenberg (Ed.), *Handbook of child psychology: Vol. 3. Social, emotional, and personality development* (5th Ed., pp. 1017–1095). New York: Wiley.

Eder, R. A., & Mangelsdorf, S. C. (1997). The emotional basis of early personality development: Implications for the emergent self-concept. In R. Hogan, J. Johnson, & S. Briggs (Eds.), *Handbook of personality psychology* (pp. 209–240). Orlando, FL: Academic Press.

Eisenberg, N., & Valiente, C. (2002). Parenting and children's prosocial and moral development. In M. Bornstein (Ed.), *Handbook of parenting: Vol. 5: Practical issues in parenting* (2nd Ed., pp. 111–142). Mahwah, NJ: Erlbaum.

Elliot, A. J., & Dweck, C. S. (in press). Competence and motivation. In A. J. Elliot & C. S. Dweck (Eds.), *The handbook of competence and motivation*. New York: Guilford.

Elliot, A. J., McGregor, H. A., & Thrash, T. M. (2002). The need for competence. In E. L. Deci & R. M. Ryan (Eds.), *Handbook of self-determination research* (pp. 361–387). Rochester, NY: University of Rochester Press.

Epstein, J. L. (1983). Longitudinal effects of family–school–person interactions on student outcomes. In A. Kerckhoff (Ed.), *Research in sociology of education and socialization* (Vol. 4, pp. 101–128). Greenwich, CT: JAI Press.

Epstein, J. L. (1988). How do we improve programs for parental involvement? *Educational Horizons*, *66*, 75077.

Estrada, P., Arsenio, W. F., Hess, R. D., & Holloway, S. D. (1987). Affective quality of the mother–child relationship: Longitudinal consequences for children's school-relevant cognitive functioning. *Developmental Psychology*, *23*, 210–215.

Fabes, R. A., Eisenberg, N., Karbon, M., Bernzweig, J., Speer, A. L., & Carlo, G. (1994). Socialization of children's vicarious emotional responding and prosocial behavior: Relations with mothers' perceptions of children's emotional reactivity. *Developmental Psychology*, *30*, 44–55.

Ford, M. E. (1985). The concept of competence: Themes and variations. In H. A. Marlowe & R. B. Weinberg (Eds.), *Competence development: Theory and research in special populations* (pp. 3–49). Hillsdale, NJ: Erlbaum.

Fredrickson, B. L. (2001). The role of positive emotions in positive psychology: The broaden and build theory of positive emotions. *American Psychologist*, *56*, 218–226.

Fuligni, A. J., Yip, T., & Tseng, V. (2002). The impact of family obligation on the daily activities and psychological well-being of Chinese American adolescents. *Child Development*, *73*, 302–314.

Garber, J., & Hilsman, R. (1992). Cognitions, stress, and depression in children and adolescents. *Child and Adolescent Psychiatric Clinics of North America*, *1*, 129–167.

Garmezy, N., & Masten, A. S. (1991). The protective role of competence indicators in children at risk. In E. M. Cummings, A. L. Greene, & K. H. Karraker (Eds.), *Life span developmental psychology: Perspectives on stress and coping* (pp. 151–174). Hillsdale, NJ: Erlbaum.

Ginsburg, G. S., & Bronstein, P. (1993). Family factors related to children's intrinsic/extrinsic motivational orientation and academic performance. *Child Development*, *64*, 1461–1474.

Gottfried, A. E., Fleming, J. S., & Gottfried, A. W. (1994). Role of parental motivational practices in children's academic intrinsic motivation and achievement. *Journal of Educational Psychology*, *86*, 104–113.

Graham, S., Harris, K. R., Fink-Chorzempa, B., & MacArthur, C. (2003). Primary grade teachers' instructional adaptations for struggling writers: A national survey. *Journal of Educational Psychology*, *95*, 279–292.

Grant, H., & Dweck, C. S. (2003). Clarifying achievement goals and their impact. *Journal of Personality and Social Psychology*, *85*, 541–553.

Grolnick, W. S. (2003). *The psychology of parental control: How well-meant parenting backfires*. Hillside, NJ: Lawrence Erlbaum.
Grolnick, W. S., Deci, E. L., & Ryan, R. M. (1997). Internalization within the family: The self-determination theory perspective. In J. Grusec & L. Kuczynski (Eds.), *Parenting and children's internalization of values: A handbook of contemporary theory* (pp. 135–161). New York: Wiley.
Grolnick, W. S., & Farkas, M. (2002). Parenting and the development of children's self-regulation. In M. Bornstein (Ed.), *Handbook of parenting: Vol. 5. Practical issues in parenting* (2nd Ed., pp. 89–110). Mahwah, NJ: Erlbaum.
Grolnick, W. S., Gurland, S. T., DeCourcey, W., & Jacob, K. (2002). Antecedents and consequences of mothers' autonomy support: An experimental investigation. *Developmental Psychology*, *38*, 143–154.
Grolnick, W. S., Kurowski, C. O., Dunlap, K. G., & Hevey, C. (2000). Parental resources and the transition to junior high. *Journal of Research on Adolescence*, *10*, 465–488.
Grolnick, W. S., Kurowski, C. O., & Gurland, S. T. (1999). Family processes and the development of children's self-regulation. *Educational Psychologist*, *34*, 3–14.
Grolnick, W. S., & Ryan, R. M. (1989). Parent styles associated with children's self-regulation and competence in school. *Journal of Educational Psychology*, *81*, 143–154.
Grolnick, W. S., Ryan, R. M., & Deci, E. L. (1991). Inner resources for school achievement: Motivational mediators of children's perceptions of their parents. *Journal of Educational Psychology*, *83*, 508–517.
Grolnick, W. S., & Slowiaczek, M. L. (1994). Parents' involvement in children's schooling: A multidimensional conceptualization and motivational model. *Child Development*, *64*, 237–252.
Grusec, J. E. (2002). Parenting socialization and children's acquisition of values. In M. H. Bornstein (Ed.), *Handbook of parenting: Vol. 5. Practical issues in parenting* (2nd Ed., pp. 143–167). Mahwah, NJ: Erlbaum.
Grusec, J. E., & Goodnow, J. J. (1994). Impact of parental discipline methods on the child's internalization of values: A reconceptualization of current points of view. *Developmental Psychology*, *30*, 4–19.
Guay, F., Boggiano, A. K., & Vallerand, R. J. (2001). Autonomy support, intrinsic motivation, and perceived competence: Conceptual and empirical linkages. *Personality and Social Psychology Bulletin*, *27*, 643–650.
Guay, F., Marsh, H. W., & Boivin, M. (2003). Academic self-concept and academic achievement: Developmental perspectives on their causal ordering. *Journal of Educational Psychology*, *95*, 124–136.
Gutman, L. M., & Eccles, J. S. (1999). Financial strain, parenting behaviors, and adolescents' achievement: Testing model equivalence between African American and European American single- and two-parent families. *Child Development*, *70*, 1464–1476.
Halle, T. G., Kurtz-Costes, B., & Mahoney, J. L. (1997). Family influences on school achievement in low-income African–American children. *Journal of Educational Psychology*, *89*, 527–537.
Hart, C. H., Olsen, S. F., Robinson, C. C., & Mandleco, B. L. (1997). The development of social and communicative competence in childhood: Review and a model of personal, familial, and extrafamilial processes. *Communication Yearbook*, *20*, 305–373.
Harter, S. (1998). The development of self-representations. In N. Eisenberg (Ed.), *Handbook of child psychology: Vol. 3. Social, emotional, and personality development* (5th Ed., pp. 102–132). New York: Wiley.
Hess, R. D., & McDevitt, T. M. (1984). Some cognitive consequences of maternal intervention: A longitudinal study. *Child Development*, *55*, 2017–2030.

Higgins, E. T. (2004). The eighth Koan of progress in social psychology: A variable anointed as "special" will demand special treatment. In J. T. Jost, M. R. Banaji, & D. A. Prentice (Eds.), *Perspectivism in social psychology: The yin and yang of scientific progress* (pp. 305–317). Washington, DC: American Psychological Association.

Hill, N. E., & Craft, S. A. (2003). Parent-school involvement and school performance: Mediated pathways among socioeconomically comparable African American and Euro-American families. *Journal of Educational Psychology*, *95*, 74–83.

Hokoda, A., & Fincham, F. D. (1995). Origins of children's helpless and mastery achievement patterns in the family. *Journal of Educational Psychology*, *87*, 375–385.

Houck, G. M., & Lecuyer-Maus, E. A. (2004). Maternal limit setting during toddlerhood, delay of gratification, and behavior problems at age five. *Infant Mental Health Journal*, *25*, 24–86.

Izzo, C. V., Weissberg, R. P., Kasprow, W. J., & Fendrich, M. (1999). A longitudinal assessment of teacher perceptions of parent involvement in children's education and school performance. *American Journal of Community Psychology*, *27*, 817–839.

Jodl, K. M., Michael, A., Malanchuk, O., Eccles, J. S., & Sameroff, A. (2001). Parents' roles in shaping early adolescents' occupational aspirations. *Child Development*, *72*, 1247–1265.

Jussim, L., Eccles, J. S., & Madon, S. (1996). Social perception, social stereotypes, and teacher expectations: Accuracy and the quest for the powerful self-fulfilling prophecy. In M. Zanna (Ed.), *Advances in experimental social psychology* (pp. 281–388). San Diego, CA: Academic Press.

Kamins, M. L., & Dweck, C. S. (1999). Person versus process praise and criticism: Implications for contingent self-worth and coping. *Developmental Psychology*, *35*, 835–847.

Kaplan, A., & Maehr, M. L. (1999). Achievement goals and well-being. *Contemporary Educational Psychology*, *24*, 330–358.

Kelley, S. A., Brownwell, C. A., & Campbell, S. B. (2000). Mastery motivation and self-evaluative affect in toddlers: Longitudinal relations with maternal behavior. *Child Development*, *71*, 1061–1071.

Kempner, S., & Pomerantz, E. M. (2003). *Mothers' use of praise in their everyday interactions with their children: The moderating role of children's gender*. Paper presented at the Society for Research on Child Development, Tampa, FL.

Kenney-Benson, G. A., Pomerantz, E. M., Ryan, A., & Patrick, H. (in press). Sex differences in math performance: The role of how children approach school. *Developmental Psychology*.

Kochanska, G. (1993). Toward a synthesis of parental socialization and child temperament in early development of conscience. *Child Development*, *64*, 325–347.

Kowaleski-Jones, L., & Duncan, G. J. (1999). The structure of achievement and behavior across middle childhood. *Child Development*, *70*, 930–943.

Kuczynski, L. (Ed.) (2003). *Handbook of dynamics in parent–child relations*. Thousand Oaks, CA: Sage.

Ladd, G. W. (2003). Probing the adaptive significance of children's behavior and relationships in the school context: A child by environment perspective. In R. Kail (Ed.), *Advances in child development and behavior* (Vol. 31, pp. 43–104). San Diego, CA: Academic Press.

Ladd, G. W., Birch, S. H., & Buhs, E. (1999). Children's social and scholastic lives in kindergarten: Related spheres of influence? *Child Development*, *70*, 1373–1400.

Ladd, G. W., & Pettit, G. S. (2002). Parenting and the development of children's peer relationships. In M. H. Bornstein (Ed.), *Handbook of parenting: Vol. 5. Practical issues in parenting* (pp. 269–309). Mahwah, NJ: Erlbaum.

Larson, R. W., & Gillman, S. (1999). Transmission of emotions in the daily interactions of single-mother families. *Journal of Marriage and the Family*, *61*, 21–37.

Leone, C. M., & Richards, M. H. (1989). Classwork and homework in early adolescence: The ecology of achievement. *Journal of Youth and Adolescence*, *18*, 531–548.

Levin, I., Levy-Shiff, R., Appelbaum-Peled, T., Katz, I., Komar, M., & Meiran, N. (1997). Antecedents and consequences of maternal involvement in children's homework: A longitudinal analysis. *Journal of Applied Developmental Psychology*, *18*, 207–227.

Lytton, H. (1990). Child and parent effects in boys' conduct disorder: A reinterpretation. *Developmental Psychology*, *26*, 683–697.

Maccoby, E. E. (1992). The role of parents in the socialization of children: An historical overview. *Developmental Psychology*, *28*, 1006–1017.

Madon, S., Jussim, L., & Eccles, J. S. (1997). In search of the powerful self-fulfilling prophecy. *Journal of Personality and Social Psychology*, *72*, 791–809.

Markus, H. R., & Wurf, E. (1987). The dynamic self-concept: A social psychological perspective. *Annual Review of Psychology*, *38*, 299–337.

Marsh, H. W., & Yeung, A. S. (1998). Longitudinal structural equation models of academic self-concept and achievement: Gender differences in the development of math and English constructs. *American Educational Research Journal*, *35*, 705–738.

Masten, A. S., Coatsworth, D., Neeman, J., Gest, S. D., Tellegen, A., & Garmezy, N. (1995). The structure and coherence of competence from childhood through adolescence. *Child Development*, *66*, 1635–1659.

Miller, S. A., Manhal, M., & Mee, L. L. (1991). Parental beliefs, parental accuracy, and children's cognitive performance: A search for causal relations. *Developmental Psychology*, *27*, 267–276.

Mischel, W., Shoda, Y., & Peake, P. K. (1988). The nature of adolescent competencies predicted by preschool delay of gratification. *Journal of Personality and Social Psychology*, *54*, 687–696.

Mischel, W., Shoda, Y., & Rodriguez, M. L. (1989). Delay of gratification in children. *Science*, *244*, 933–938.

Miserandino, M. (1996). Children who do well in school: Individual differences in perceived competence and autonomy in above-average children. *Journal of Educational Psychology*, *88*, 203–214.

Mize, J., Pettit, G. S., & Brown, E. G. (1995). Mothers' supervision of their children's peer play: Relations with beliefs, perceptions, and knowledge. *Developmental Psychology*, *31*, 311–321.

Morris, A. S., Silk, J. S., Steinberg, L., Sessa, F. M., Avenevoli, S., & Essex, M. J. (2002). Temperamental vulnerability and negative parenting as interacting predictors of child adjustment. *Journal of Marriage and the Family*, *64*, 461–471.

Mueller, C. M., & Dweck, C. S. (1998). Praise for intelligence can undermine children's motivation and performance. *Journal of Personality and Social Psychology*, *75*, 33–52.

Ng, F. F., Kenney-Benson, G. A., & Pomerantz, E. M. (2004). Children's achievement moderates the effects of mothers' use of control and autonomy support. *Child Development*, *75*, 764–780.

Nolen-Hoeksema, S., Wolfson, A., Mumme, D., & Guskin, K. (1995). Helplessness in children of depressed and nondepressed mothers. *Developmental Psychology*, *31*, 377–387.

O'Connor, T. G. (2002). Annotation: The 'effects' of parenting reconsidered: Findings challenges, and applications. *Journal of Child Psychology and Psychiatry*, *43*, 555–572.

Park, S.-Y., Belsky, J., Putnam, S., & Crnic, K. (1997). Infant emotionality, parenting, and 3-year inhibition: Exploring stability and lawful discontinuity in a male sample. *Developmental Psychology*, *33*, 218–227.

Parke, R. D., & Buriel, R. (1998). Socialization in the family: Ethnic and ecological perspectives. In N. Eisenberg (Ed.), *Handbook of child psychology: Vol. 3. Social, emotional, and personality development* (5th Ed., pp. 463–552). New York: Wiley.

Parsons, J. E., Kaczala, C. M., & Meece, J. L. (1982). Socialization of achievement attitudes and beliefs: Classroom influences. *Child Development*, *53*, 322–339.

Pezdek, K., Berry, T., & Renno, P. A. (2002). Children's mathematics achievement: The role of parents' perceptions and their involvement in homework. *Journal of Educational Psychology*, *94*, 771–777.

Phillips, D. A. (1984). The illusion of incompetence among academically competent children. *Child Development*, *55*, 2000–2016.

Phillips, D. A., & Zimmerman, M. (1990). The developmental course of perceived competence and incompetence among competent children. In R. J. Sternberg & J. Kolligian (Eds.), *Competence considered* (pp. 41–66). New Haven, CT: Yale University Press.

Pintrich, P. R., & Blumenfeld, P. C. (1985). Classroom experience and children's self-perceptions of ability, effort, and conduct. *Journal of Educational Psychology*, *77*, 646–657.

Pintrich, P. R., & De Groot, E. V. (1990). Motivational and self-regulated learning components of classroom academic performance. *Journal of Educational Psychology*, *82*, 33–40.

Plomin, R. (1994). *Genetics and experience: The interplay between nature and nurture*. Thousand Oaks, CA: Sage.

Pomerantz, E. M. (2001). Parent × Child socialization: Implications for the development of depressive symptoms. *Journal of Family Psychology*, *15*, 510–525.

Pomerantz, E. M., & Eaton, M. M. (2001). Maternal intrusive support in the academic context: Transactional socialization processes. *Developmental Psychology*, *37*, 174–186.

Pomerantz, E. M., Grolnick, W. S., & Price, C. E. (in press). The role of parents in how children approach school: A dynamic process perspective. In A. J. Elliot, & C. S. Dweck (Eds.), *The handbook of competence and motivation*. New York: Guilford.

Pomerantz, E. M., Ng, F., & Wang, Q. (2004). *Mothers' mastery-oriented involvement in children's homework: Implications for the well-being of children with negative perceptions of competence*. Unpublished manuscript.

Pomerantz, E. M., & Ruble, D. N. (1998a). The multidimensional nature of control: Implications for the development of sex differences in self-evaluation. In J. Heckhausen & C. S. Dweck (Eds.), *Motivation and self-regulation across the life-span* (pp. 159–184). New York: Cambridge University Press.

Pomerantz, E. M., & Ruble, D. N. (1998b). The role of maternal control in the development of sex differences in child self-evaluative factors. *Child Development*, *69*, 458–478.

Pomerantz, E. M., & Rudolph, K. D. (2003). What ensues from emotional distress? Implications for competence estimates. *Child Development*, *74*, 329–346.

Pomerantz, E. M., & Saxon, J. L. (2001). Conceptions of ability as stable and self-evaluative processes: A longitudinal examination. *Child Development*, *72*, 152–173.

Pomerantz, E. M., Saxon, J. L., & Oishi, S. (2000). The psychological tradeoffs of goal investment. *Journal of Personality and Social Psychology*, *79*, 617–630.

Pomerantz, E. M. & Shim, S. (in press). The regulatory function of goal investment: A look at the benefits. In J. Shah & W. Gardner (Eds.), *Handbook of motivational science*. New York: Guilford.

Pomerantz, E. M., Wang, Q., & Ng, F. F. (in press). Mothers' affect in the homework context: The importance of staying positive. *Developmental Psychology, 41*.

Reynolds, A. J. (1989). A structural model of first-grade outcomes for an urban, low socioeconomic status, minority population. *Journal of Educational Psychology*, *81*, 594–603.

Roeser, R. W., Eccles, J. S., & Sameroff, A. (1998). Academic and emotional functioning in early adolescence: Longitudinal relations, patterns, and prediction by experience in middle school. *Development and Psychopathology*, *10*, 321–352.

Roeser, R. W., Midgeley, C., & Urdan, T. (1996). Perceptions of the school psychological environment and early adolescents' psychological and behavioral functioning in school: The mediating role of goals and belonging. *Journal of Educational Psychology*, *88*, 408–422.

Roisman, G., Masten, A. S., Coatsworth, D., & Tellegen, A. (2004). Salient and emerging developmental tasks in the transition to adulthood. *Child Development*, *75*, 123–133.

Rollins, B. C., & Thomas, D. L. (1979). Parental support, power, and control techniques in the socialization of children. In W. R. Burr, R. Hill, F. I. Nye, & I. L. Reiss (Eds.), *Contemporary theories about the family* (pp. 317–364). New York: Free Press.

Rubin, K. A., Hastings, P., Chen, X., Stewart, S., & McNichol, K. (1998). Intrapersonal and maternal correlates of aggression, conflict, and externalizing problems in toddlers. *Child Development*, *69*, 1614–1629.

Rudolph, K. D. (in press). Adolescent depression in the school context. In T. Urdan & F. Pajares (Eds.), *Adolescence and education* (Vol. 4). Greenwich, CT: Information Age Publishing.

Rudolph, K. D., & Asher, S. R. (2000). Adaptation and maladaptation in the peer system: Developmental processes and outcomes. In A. J. Sameroff, M. Lewis, & S. M. Miller (Eds.), *Handbook of developmental psychopathology* (2nd Ed., pp. 157–175). New York: Plenum Press.

Ryan, A. (2001). The peer group as a context for the development of young adolescent motivation and achievement. *Child Development*, *72*, 1135–1150.

Ryan, R. M., & Connell, J. P. (1989). Perceived locus of causality and internalization: Examining reasons for acting in two domains. *Journal of Personality and Social Psychology*, *57*, 749–761.

Ryan, A., Patrick, H., & Shim, S. (in press). Differential profiles of students identified by their teacher as having avoidant, appropriate or dependent help-seeking tendencies in the classroom. *Journal of Educational Psychology, 97*.

Sameroff, A., & Macenzie, M. J. (2003). Research strategies for capturing transactional models of development: The limits of the possible. *Development and Psychopathology*, *15*, 613–640.

Scarr, S. (1992). Developmental theories for the 1990s: Development and individual differences. *Child Development*, *63*, 1–19.

Schwartz, D., Dodge, K. A., Petit, G. S., Bates, J. E., & The Conduct Problems Prevention Research Group. (2000). Friendship as a moderating factor in the pathway between early harsh home environment and later victimization in the peer group. *Developmental Psychology*, *36*, 646–662.

Senechal, M., & LeFevre, J. (2002). Parental involvement in the development of children's reading skill: A five year longitudinal study. *Child Development*, *73*, 445–460.

Silverman, I. W., & Ippolito, M. F. (1995). Maternal antecedents of delay ability in young children. *Journal of Applied Developmental Psychology*, *16*, 569–591.

Skinner, E. A., Zimmer-Gembeck, M. J., & Connell, J. P. (1998). Individual differences and the development of perceived control. *Monographs of the Society for Research in Child Development*, *63*, Serial No. 254.

Sroufe, L. A., Fox, N. E., & Pancake, V. R. (1983). Attachment and dependency in developmental perspective. *Child Development*, *54*, 1615–1627.

Steinberg, L. (1990). Autonomy, conflict, and harmony in the family relationship. In S. Feldman & G. R. Elliott (Eds.), *At the threshold: The developing adolescent* (pp. 255–277). Cambridge: Harvard University Press.

Steinberg, L., Elmen, J. D., & Mounts, N. S. (1989). Authoritative parenting, psychosocial maturity, and academic success among adolescents. *Child Development*, *60*, 1424–1436.

Steinberg, L., Lamborn, S. D., Darling, N., Mounts, N. S., & Dornbusch, S. (1994). Over-time changes in adjustment and competence among adolescents from authoritative, authoritarian, indulgent, and neglectful homes. *Child Development*, *65*, 754–770.

Steinberg, L., Lamborn, S. D., Dornbusch, S. M., & Darling, N. (1992). Impact of parenting practices on adolescent achievement: Authoritative parenting, school involvement, and encouragement to succeed. *Child Development*, *63*, 1266–1281.

Stevenson, H. W., & Baker, D. P. (1987). The family–school relations and the child's school performance. *Child Development*, *58*, 1348–1357.

Stipek, D. J. (1981). Children's perceptions of their own and their classmates' ability. *Journal of Educational Psychology*, *73*, 404–410.

Stipek, D. J., & MacIver, D. (1989). Developmental change in children's assessment of intellectual competence. *Child Development*, *60*, 521–538.

Stipek, D. J., & Tannatt, L. M. (1984). Children's judgments of their own and their peers' academic competence. *Journal of Educational Psychology*, *76*, 75–84.

Tesser, A. (1988). Toward a self-evaluation maintenance model of social behavior. In L. Berkowitz (Ed.), *Advances in experimental social psychology, Vol. 21. Social psychological studies of the self: Perspectives and programs* (pp. 181–227). New York: Academic Press.

Wagner, B. M., & Phillips, D. A. (1992). Beyond beliefs: Parent and child behaviors and children's perceived academic competence. *Child Development*, *63*, 1380–1391.

Wentzel, K. R. (2002). Are effective teachers like good parents? Teaching styles and student adjustment in early adolescence. *Child Development*, *73*, 287–301.

White, R. H. (1959). Motivation reconsidered: The concept of competence. *Psychological Review*, *66*, 297–333.

Wigfield, A., Eccles, J. S., MacIver, D., Reuman, D. A., & Midgley, C. (1991). Transitions during early adolescence: Changes in children's domain-specific self-perceptions and general self-esteem across the transition to junior high school. *Developmental Psychology*, *27*, 552–565.

Wolters, C. A. (2004). Advancing achievement goal theory: Using goal structures and goal orientation to predict students' motivation, cognition, and achievement. *Journal of Educational Psychology*, *96*, 236–250.

Zimmerman, B. J., & Martinez-Pons, M. (1990). Student differences in self-regulated learning: Relating grade, sex, and giftedness to self-efficacy and strategy use. *Journal of Educational Psychology*, *82*, 51–59.

This Page Intentionally Left Blank

THE INFANT ORIGINS OF INTENTIONAL UNDERSTANDING

Amanda L. Woodward

DEPARTMENT OF PSYCHOLOGY, COMMITTEE ON HUMAN DEVELOPMENT,
UNIVERSITY OF CHICAGO, CHICAGO, IL 60637, USA

I. Introduction

Children develop immersed in sea of human action. All around them, people traverse complex paths, attend to and act on objects, and interact with one another. Becoming a functional member of our species depends on being able to represent these actions not as purely physical motions through space, but rather as the manifestation of a person's psychological life. When this ability is seriously impaired, as is the case in autism, the effects on cognition and social life are pervasive and devastating (Baron-Cohen, 1995; Repacholi & Slaughter, 2003).

Adults possess a system of knowledge sometimes termed "folk psychology" which explains observed behavior with reference to internal states such as beliefs, perceptions, emotions, and intentions. This system of knowledge takes years to fully emerge in human ontogeny (Wellman, 1992; Astington, 1993; Flavell & Miller, 1998). Even so, elements of intentional action knowledge are in place by the second year of life. Some of the strongest evidence for this point comes from studies of social learning.

To illustrate, imagine the following scene: preparing to head to the park, a man turns to his 1-year-old daughter and says "Now, where are your socks?" He proceeds to open a dresser drawer, peer inside, and pull out one item after another, tossing each aside with a sigh. Finally, he pulls out the socks and says "There we go". An astute child would understand her father's instrumental actions (opening the drawer, digging through the clothing), and attentional behaviors (peering into the drawer) as evidence of his intention (to obtain a particular item), and thus have a basis for understanding his subsequent actions. She would also be able to infer the meaning of the word "socks", even though she did not experience a direct pairing of the word and its referent. By 14–18 months, children do exactly this; that is, they use behavioral evidence of a speaker's attentional states and goals to interpret his or her subsequent actions, as well as his or her emotional expressions and the words that he or she utters (see Tomasello, 1999; Baldwin & Moses, 2001 for reviews). Moreover, when young children imitate the actions of adults, they reproduce the apparent intention of the adult, not his or her exact motor patterns (Meltzoff, 1995; Carpenter, Akhtar, & Tomasello, 1998; Bellagamba & Tomasello, 1999; Gergely, Bekkering, & Kiraly, 2001; Hauf, Elsner, & Aschersleben, in press).

Thus, by the second year of life children's learning from social partners is mediated by an analysis of the intentional structure of action. By 18 months of age, children draw on a person's inferred goals, plans, and states of attention to glean new information, both about the person and about the environment. This conclusion highlights two important points. First, the ability to analyze the intentional structure of actions is foundational to social learning and cognitive development more generally. Therefore, it is critical to understand the developmental origins of this ability. Second, aspects of this foundational ability are present by 18 months of age. Therefore, the origins of these abilities should be sought still earlier in ontogeny. In this chapter, I review recent findings from studies that seek these origins. I begin by outlining what infants' seem to know about intentional action during the first year of life, and then I consider the question of how this knowledge originates.

II. What Infants Know About Action

A. CERTAIN ACTIONS ARE ORGANIZED BY THE RELATION BETWEEN AGENT AND GOAL

Mature observers interpret actions not as purely physical motions through space but rather as directed at particular objects or outcomes. This object-directedness is a perceived property of many intentional actions, both at the level of individual actions and at the level of sequences of action. This is evident in adults' event memory and narratives (Zacks & Tversky, 2001), and in children's responses to the actions of others (Bekkering, Wohlschlaeger, & Gattis, 2000). Even a simple, concrete action like the one depicted in Figure 1 is most readily described in terms of the relation between agent and goal ("She grasped the bear") rather than in terms of the strictly physical properties of the person's motion (e.g., "She moved her arm up and to the left"). We can perceive and represent the physical features of actions, of course. But, to adult eyes, the physical attributes are less central than the relation between the agent and his or her goal. A first question then, is whether, and under what conditions infants represent actions in terms of their goal structure.

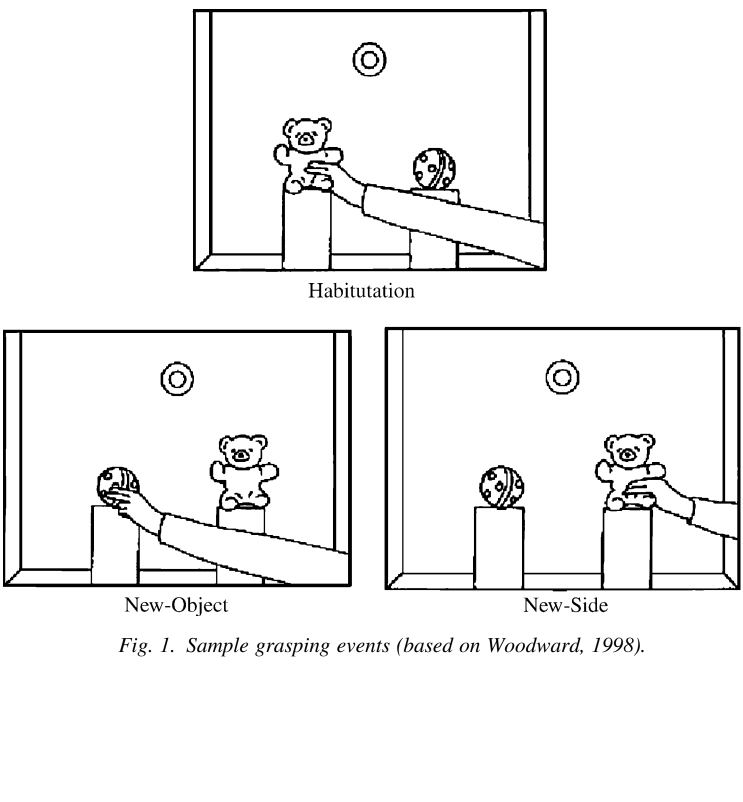

Fig. 1. Sample grasping events (based on Woodward, 1998).

The visual habituation paradigm offers a way to test whether infants, like adults, represent actions in terms of the relation between agent and object. The logic is to use infants' visual responses as evidence about the features they weight heavily in their mental representations of events. To illustrate, having habituated infants to one action, we present test events that either (1) vary the surface properties of the event while maintaining the relation between agent and object, or (2) preserve many of the surface properties while varying the relation between the agent and the object. Following full habituation, infants are predicted to look longer at stimuli that they perceive to be novel compared to the habituation stimulus. Therefore, longer looking on the latter trials than the former indicates that infants represented the original event primarily in terms of the relation between agent and object.

Our findings indicate that infants are sensitive to the goal-directed structure of one action, grasping, by the time they are 5–6 months of age (Woodward, 1998, 1999, 2003; Guajardo & Woodward, 2004). To illustrate, in one study (Woodward, 1998) infants were habituated to an event in which a person grasped one of two toys mounted on a stage (see Figure 1). After habituating infants to one event, we reversed the objects' positions and showed infants test events that either disrupted the spatial properties of the reach while maintaining the same goal relation (new path events) or maintained the spatial properties of the reach while disrupting the goal relation (new-object trials). In our studies, infants at 6, 7, 9, and 12 months have shown a strong novelty response (i.e., longer looking) on new-object trials than on new path trials. This finding has been replicated in several other laboratories (Wellman & Phillips, 2001; Jovanovic *et al*., 2004; Sodian & Thoermer, 2004; Spaepen & Spelke, 2004).

The initial findings also suggested that infants' propensity to encode actions as object-directed is specific to familiar *human* actions—infants did not respond in the same way to events in which inanimate claws or ambiguous agents grasp objects (Meltzoff, 1995; Woodward, 1998; Guajardo & Woodward, 2004; see also Jovanovic *et al*., 2004), or to unfamiliar human actions on objects (Woodward, 1999). However, several researchers have suggested that infants interpret novel actions or the motions of unusual agents as goal-directed under certain conditions (Gergely *et al*., 1995; Johnson, 2000; Kiraly *et al*., 2003; Luo & Baillargeon, 2004; Shimizu & Johnson, in press). I return to this possibility later in this chapter.

B. SHIFTING ATTENTION TO OBJECTS VS. REPRESENTING AGENT–OBJECT RELATIONS

In the studies just summarized, as in many studies of infant cognition, the goal was to assess infants' representations of event structure via their looking times on test trials. Like other researchers who use this paradigm, we were careful to control the factors that could have contributed to infants' visual responses. The same agent and objects were present in each habituation and test trial, with the key variations in the events concerning the relation between the agent and the objects. This, we reasoned, would enable us to determine whether infants' represented the agent–object relation. However, there was a remaining concern: infants' responses on test trials might be a by-product of the way the events entrained their attention on the objects rather than reflecting their representation of the agent–object relation *per se*. This was of particular concern because the events involved the actions of social partners, and these actions have been shown to entrain infants' attention in other contexts. Infants may look "smarter" than they are because adults' actions support their attention to the relevant aspects of the environment. Indeed, debates about infant social cognition hinge on the question of whether infants really understand the relevant intention or instead are simply led to look in the right place at the right time (e.g., Moore & Corkum, 1994; Tomasello, 1995). In the case of the habituation studies we worried that the motion of the hand may have drawn infants' attention to the object, as if it were a spotlight on the object, and then, when the spotlight was directed at a new object during test trials infants looked longer not because the agent–object relation had changed but because they were led to look at a different object.

One way to assess this possibility was to conduct matched control conditions involving events that were not goal-directed. For example, in various control conditions we showed infants rods that touched the objects, hand-shaped cardboard cutouts that partially occluded the objects, mechanical claws that grasped the objects, or apparently purposeless manual contact with the objects. In none of these cases did infants look longer on new-object than new-side trials (Woodward, 1998, 1999; see also Jovanovic *et al.*, 2004). These findings suggest that infants' responses were not driven by the motion of the hand toward the object or its contact with the object—because this motion and contact were present in each of the control events. It still might be the case, however, that grasping hands are more potent spotlights for infants than are inanimate objects and inert hand postures. To evaluate this possibility, we coded infants' attention to each of the toys during test trials. We found that rods, flat cutouts, claws, purposeless hands, and grasping hands were all equally effective in directing infants' attention to the contacted toy. Because these attentional effects were uniform across conditions, they cannot account for infants' differential responses to new-object vs. new-side events across conditions.

These analyses show that infants' overall looking times on test events were driven by their representation of event structure rather than by the effects of the action on their attention to the objects in the display. There are many ways to lead infants to look at an object, including grasping it, touching it with a rod, grasping it with a claw, and dropping one's hand onto it. But directing attention in this way does not determine whether infants represent the event in terms of the relation between actor and object. Infants only encode events in terms of the actor–object relation in the case of intentional human actions such as grasping, and, as I review next, certain other intentional actions.

C. THERE IS A CONNECTION BETWEEN A PERSON AND THE OBJECT OF HIS OR HER ATTENTION

Instrumental actions, such as grasping, carry concrete indicators of the goals they express. In everyday life, these actions have observable effects, for example, moving a desired object closer to the agent, and these effects may support infants' ability to extract the goal structure of these actions. In contrast, the relation between a person and the object of her attention can only be inferred, and seems, to non-scientist adults and psychologists alike, to be uniquely psychological. Following on this intuition, researchers have long been interested in when and how infants come to understand the invisible connection between a person and the object of his or her attention.

For many years, the main means for investigating this issue was to assess infants' propensity to look in the direction of an adult's gaze shifts. Infants systematically follow adult gaze shifts during the first year of life (e.g., Scaife & Bruner, 1975; Schaffer, 1984; Butterworth & Jarrett, 1991). Researchers have often assumed that if infants turn to follow the adult's line of regard, then they must do so because they understand that the adult is looking at something. The prior discussion illustrates the problem with this assumption. Shifting attention to a location need not indicate that infants have understood the action as object-directed. Indeed, there has long been debate about the significance of infants' gaze-following, with a number of researchers pointing out that gaze-following could result from processes that do not involve a conceptual representation of the "seeing" relation (e.g., Moore & Corkum, 1994).

This debate about the significance of gaze-following indicates that an alternative source of evidence is needed. With this need in mind, we recruited the experimental logic from our studies of grasping to ask whether infants represent the invisible connection between a person and the object of his or her attention (Woodward & Guajardo, 2002; Woodward, 2003). We showed infants events in which a person turned to look at a toy (Woodward, 2003—see Figure 2), in which a person looked at and pointed to a toy (Woodward & Guajardo, 2002), and in which only the person's arm was visible as she pointed to the toy (Woodward & Guajardo, 2002). In each case, infants were habituated to an event in which a person pointed toward (or gazed at) one of two toys. Then, the toys' positions were reversed and infants viewed test events that disrupted either the object to which the person directed pointing (or gaze) (new-object trials) or the person's physical motions (maintaining the same object as the target) (new-side trials). For all the events, the person's actions drew infants' attention to

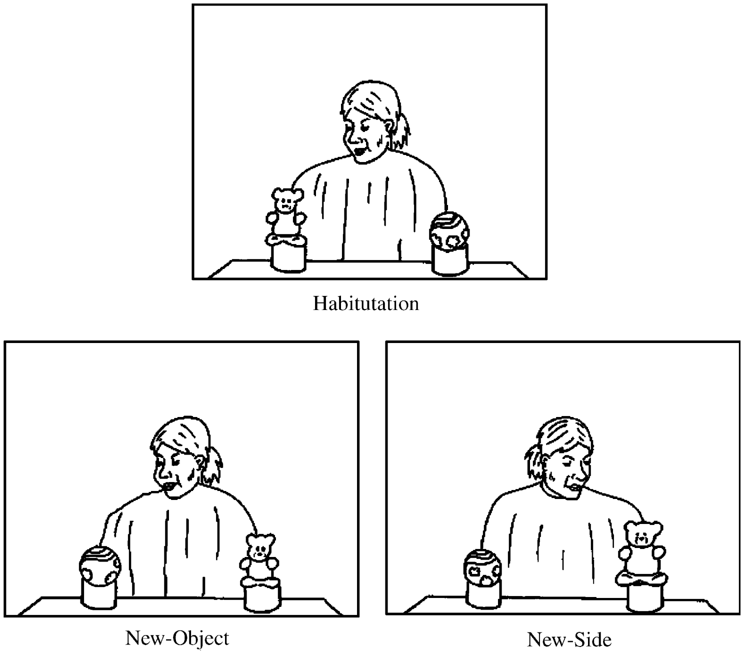

Fig. 2. Sample gaze events for (based on Woodward, 2003).

the indicated toy at all of the ages we studied (7-, 9-, and 12-month-olds). However, only 12-month-old infants responded to a change in the actor–object relation for gaze and pointing: they showed a reliably greater novelty response on new-object trials than on new side trials. Infants younger than 12 months shifted their attention in response to the experimental events, but seemed not to comprehend the significance of the action to which they had just responded.

Observers have long noted that between 9 and 12 months infants seem to "tune in" to their social partners, engaging in more shared attention with their parents and producing as well as responding to communicative gestures (Bakeman & Adamson, 1984; Schaffer, 1984; Tomasello, 1995). Our findings provide evidence that this change in social responsiveness is accompanied by a change in infants' social cognition, specifically, an emerging sensitivity to the relational structure of attentional behaviors. Furthermore, our findings suggest that specific relations exist between infants' social responsiveness and their social cognition. We found that infants' own pointing status was related to their sensitivity to the object-directed structure of pointing. In Woodward and Guajardo (2002) we tested 48 infants between the ages of 8 and 11 months using the habituation paradigm for pointing. For each infant, we also established (via parental interview and observations in the laboratory) whether the infant produced clear points that were directed toward objects. Eighteen of the infants had begun to produce object-directed points and 30 had not. These two groups of infants did not differ in terms of their age, overall attentiveness or habituation rates. They did differ, however, in terms of their responses on test trials: infants who pointed looked reliably longer on new-object trials than on new-side trials, whereas infants who did not point looked equally on the two kinds of test trials.

The results across our studies of gaze and pointing converge with findings from other laboratories in indicating infants' growing awareness of attentional relations beginning at around 12 months (Phillips, Wellman, & Spelke, 2002; Onishi & Baillargeon, 2004; Sodian & Thoermer, 2004). To illustrate, in one study, Phillips, Wellman, and Spelke (2002) tested whether infants use the inferred relation between a person and the object of her attention to predict her subsequent actions. During habituation, infants saw a woman first look at and then pick up one of two toys. Then, infants viewed test events in which the woman looked at and then picked up the other toy (consistent actions), or in which she looked at the first toy but picked up the other one (inconsistent actions). At 12 months (but not 8 months), infants looked longer at the latter than the former, indicating that they detected the inconsistent relation between the woman's gaze and her subsequent actions. At 14 months, infants expected the woman to act on the prior object of her attention even when they were not habituated to a full "look and then grab" sequence.

Experiments that capitalize on infants' social responding provide further evidence for an emerging understanding of attentional relations at the end of the first year. As noted previously, infants follow gaze from early in the first year of life, and our findings suggest that this early gaze-following does not rest on an understanding of attentional relations. However, at 12–14 months, infants begin to modulate their gaze-following in ways that suggest they understand the connection between a person and the object at which her gaze is directed as well as the physical constraints governing this connection: specifically, they refrain from following an adult's gaze when the adult's eyes are closed or when there is a barrier between the adult's eyes and the object (Brooks & Meltzoff, 2002; Dunphy-Lelii & Wellman, in press; see also Butler, Caron, & Brooks, 2000 for similar evidence with older infants). Moreover, 12- to 18-month-old infants have been shown to use an adult's gaze direction to interpret his or her referential expressions (e.g., Baldwin, 1995; Moses *et al*., 2001; Tomasello & Haberl, 2003; Woodward, 2004a). To illustrate, Moses and colleagues (2001) found that 12-month-old infants relate an adult's expression of disgust to the object at which the adult was looking, even when infants were unable to view the object themselves.

D. THE SAME MOTION MIGHT OR MIGHT NOT BE GOAL-DIRECTED IN DIFFERENT CONTEXTS

The evidence reviewed so far shows that during the first year infants represent several common actions, grasping, looking, and pointing, as being object-directed. These actions are potent signs of goals and intentions for adults, so much so that they have become metaphors for more abstract intentional relations (e.g., "The prize was just beyond my grasp" or "I see what you mean"). However, mature observers are not limited to understanding certain canonical actions as expressing intentions. Rather, we can flexibly interpret actions online, using the context to infer the goals or intentions behind ambiguous or novel actions. This ability is based in the knowledge that goals or intentions do not reside in the particular actions that they drive. To the extent that infants can also flexibly represent actions as goal-directed or not, then, this indicates that they may also understand goals as being distinct from particular actions.

Studies from several distinct paradigms support the conclusion that by 9–12 months of age, infants interpret actions based on the context in which they occur, including the physical context (such as whether the action is a rational means to attain the goal given the physical obstacles present) as well as the other actions the agent produces (such as facial and vocal expressions of frustration or surprise). In a striking demonstration of the first of these, Gergely *et al.* (1995) found that 12-month-old infants responded to a computer-animate shape traversing a looping path as being goal-directed when it circumvented a wall to approach another shape, but did not respond in this way to the same path of motion when there was no wall present. The looping path was apparently a rational route to the other shape when the wall was present, but not when it was absent. Gergely and colleagues concluded that infants evaluated the rationality of the shape's motion, and responded to the rational path as evidence of goal-directedness (see also Csibra *et al.*, 2003). Csibra and colleagues (1999) obtained a similar result at 9 months, and other studies have confirmed that infants respond in the same way when the moving entity is a person rather than a computer animation (Sodian, Schoeppner, & Metz, 2004).

Behne and colleagues (in press) reported converging evidence from a paradigm that manipulated infants' social responses. They engaged infants in a game in which an experimenter handed the infant a series of small toys. After several exchanges, the experimenter failed to hand the infant a toy, in some cases acting as if she was unwilling to complete the transfer, and in other cases acting as if she was unable to do so. These two cases were designed to involve similar movements. For example, the adult would hold out a toy and then teasingly pull it out of reach in one case, and hold it out and then "accidentally" drop it in the other. Behne and colleagues found that infants as young as 9 months of age responded with more frustration when the adult was unwilling to hand them the toy than when she was apparently unable to do so. Across items, infants seemed to recruit several different kinds of information to make sense of the two kinds of actions, including the causal constraints present (e.g., the toy was out of reach) as well as the adults' facial and vocal behaviors indicating an intent to tease. Though methodologically quite different from the work by Gergely and Csibra, this finding supports a similar conclusion: by 9 months of age, infants represent very similar motions as being object-directed or not based on contextual information.

Infants younger than 9 months did not respond systematically in either the habituation paradigm developed by Gergely, Csibra and colleagues or in the paradigm developed by Behne and colleagues. One possibility is that although younger infants are sensitive to the goal-directedness of some familiar actions, they are not able to flexibly interpret actions based on contextual information. Alternatively, younger infants may be unable to draw on the particular kinds of contextual information provided in these studies, but able to use other aspects of the situation to interpret actions as goal-directed. In line with this possibility, several researchers have proposed situational and behavioral cues that are hypothesized to support young infants' interpretation of actions as goal-directed. Kiraly and colleagues (2003), for example, propose that when an action has an observable causal effect on the object (i.e., it causes an object to move), or when repeated actions apparently pursue the same goal via different routes, then infants interpret the action as goal-directed. The evidence for these proposals is considered subsequently because it bears on debates about the innate contributors to infants' action knowledge.

E. ACTIONS CAN BE ASSEMBLED IN SERVICE OF OVERARCHING GOALS

Adults are not limited to understanding the goal structure of single actions such as grasps or glances, but can understand sequences of actions as being organized by overarching goals (Schank & Abelson, 1977; Searle, 1983; Zacks & Tversky, 2001). For example, seeing someone walk to the cupboard, grasp the knob, pull open the door, and then grasp a box of cookies inside, we understand not only the goal of each component action (e.g., opening the cupboard), but also the overarching goal that drives the sequence (getting something to eat). Zacks and Tversky (2001) describe this aspect of action structure in terms of *partonomic hierarchies*: actions organized by subgoals form the parts of a sequence organized by a higher order goal. Analyzing action sequences in this way is integral to mature event representation (e.g., Searle, 1983), and this hierarchical structure is evident in adults' and children's memories for and descriptions of complex events (Trabasso *et al.*, 1992; Baldwin & Baird, 2001; Bekkering, Wohlschlaeger, & Gattis, 2000; Zacks, Tversky, & Iyer, 2001). Moreover, this aspect of action knowledge structures imitative learning by the second year of life (Wenner & Bauer, 1999; Gergely, Bekkering, & Kiraly, 2001).

Because of its centrality to mature conceptions of intention, developmental psychologists have considered hierarchical action representation to be a hallmark of both having intentions (Piaget, 1953) and representing the intentions of others (Gergely *et al.*, 1995; Meltzoff, 1995; Tomasello, 1999). The findings reviewed so far indicate that some of the prerequisites for this ability emerge during the first year. Infants can detect the goal structure of single actions early in the first year, and, by 9–12 months, infants can interpret a single action as goal-directed or not based on contextual information. Furthermore, 9- to 12-month-old infants sometimes assume that sequential actions, such as looking and grasping, will be directed at the same object (Phillips, Wellman, & Spelke, 2002; Sodian & Thoermer, 2004).

These findings raise the question of whether infants could also interpret the same action as being directed at goals at differing hierarchical levels. To investigate this question, we have introduced infants to action sequences in which a person acts on one object in order to gain access to another object. These sequences have a simple hierarchical structure in which the action on the first object is interpretable either as directed at that object (the proximal goal) or at the object obtained a the end of the sequence (the ultimate goal). To illustrate, in one study (Woodward & Sommerville, 2000), 12-month-old infants saw an actor reach toward and grasp the lid of one of two transparent boxes, each of which contained a toy (see Figure 3). The actor proceeded to open the box and then grasp the toy inside it. The question of interest was whether infants interpreted the first action, the grasp of the box lid, as directed at the box itself or

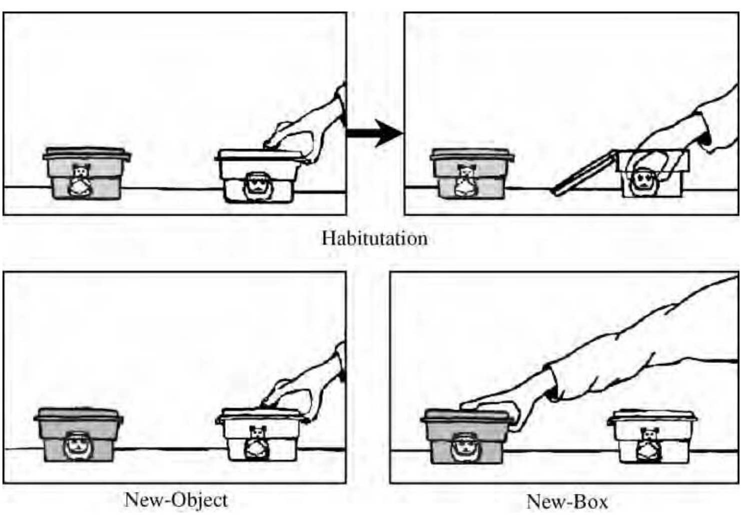

Fig. 3. Sample means-end events (based on Woodward & Sommerville, 2000).

instead at the toy inside the box. Infants were habituated to one box-opening sequence (see Figure 3 for an example). After habituation, the positions of the toys were reversed, and infants saw new-object test trials (in which the actor grasped the same box as during habituation, which now contained a different toy) and new-box trials (in which the actor grasped the other box, which now contained the toy that had been in the contacted box during habituation). In each case, the adult grasped the box lid but did not complete the sequence. Thus, these trials provided a test of how infants interpreted the first action in the sequence (grasping the box lid). If they interpreted it as directed at the box itself, they should look longer on new-box trials; if they interpreted the initial grasp as directed at the toy within the box, they should look longer on new-toy trials. Twelve-month-old infants showed the latter pattern, looking longer when the actor grasped the same box that now contained a different toy. Thus, they "read ahead" in the sequence, and responded to a change in the ultimate goal rather than a change in the proximal goal, even though the actor did not go on to complete the sequence.

On what basis did infants read ahead? One possibility is that infants relate actions to ultimate goals based solely on their order of occurrence—that is they may assume that actions are directed at the goals, which follow them. Infants are adept at extracting sequential patterns in temporally ordered stimuli (Saffran, Aslin, & Newport, 1996) and this has been hypothesized to contribute to their action analysis (Baldwin & Baird, 2001). However, adults do not analyze action based only on sequential ordering. In relating actions to higher order plans, we draw on contextual evidence including information about the causal constraints on action. When an action physically enables the attainment of a goal, then we may interpret the action as directed at that goal. As described previously, Gergely, Csibra and others have shown that infants use such evidence to infer whether or not a particular path of motion is goal-directed (Gergely *et al*., 1995; Csibra *et al*., 1999), and this raises the possibility that infants may also be able to use causal relations to relate subgoals to higher order goals.

To address this question, we conducted a follow-up to the box-opening study (Woodward & Sommerville, 2000, Study 2), testing whether infants related the actions based only on their sequential ordering, or, instead drew on the causal relation between them (opening the box enabled the actor to grasp the toy). The events were the same as in the first study except that now the toy sat outside the box rather than in it. Thus, although the temporal relation between the actions was maintained, the causal relation between them was disrupted. Under these conditions, infants looked marginally longer on new-box trials, indicating that when there was no causal relation between the actions, infants did not interpret the grasp of the box lid as directed at the toy. In other words, infants, like adults, draw on causal relations as evidence about the higher order goals at which actions may be directed.

In a later series of studies, we replicated these results at 12 months using a different means-end problem, pulling a cloth to obtain a toy (Sommerville & Woodward, 2005). We then went on to test younger infants, 10-month-olds, and found that as a group they responded randomly to the test events. However, additional measures and analyses revealed underlying individual variation in infants' responses at this age. Each infant also completed an action task, in which they were presented with a toy out of reach on a cloth, and their task was to bring the toy within reach. There was a positive correlation between the extent to which infants produced planful responses on the action task (maintaining eye-contact with the toy while pulling it into reach and then grasping the toy as soon as it came near) and their preference for the new-object event over the new cloth event in the habituation task. Moreover, an analysis of the habituation performance of the upper and lower 25% of infants in the action task revealed two distinct patterns of response: the most planful infants looked reliably longer on new-object trials, indicating that they interpreted the actor's grasp of the cloth as directed to the toy, and the least planful infants looked longer on new cloth trials, indicating that they interpreted the grasp of the cloth as directed toward the cloth. Thus, at 10 months, some infants endorsed one interpretation of the action—grasping the box was directed at the box itself—whereas others endorsed the more abstract interpretation—grasping the box was directed at the toy inside.

In summary, by 10–12 months, infants attend not only to the local relations between actions and objects, but also to relations between actions and ultimate outcomes. An action, grasping, that can readily be interpreted as directed at the object grasped can be re-interpreted as directed a goal further along in the sequence, so long as a causal chain connects the actions. By the end of the first year, then, infants have begun to extract the partonomic hierarchical structure of action.

F. CONCLUSIONS: WHAT INFANTS KNOW

During the first year of life infants begin to analyze action in terms of its intentional structure. This analysis goes beyond the surface level of motions and contact, reflecting meaningful components of human behavior. This analysis is first evident in 5- to 6-month-old infants' propensity to relate agents to goals for actions that appear purposeful, in particular, grasping. Between 9 and 12 months, infants begin to relate agents to the objects of their attention, and also begin to relate actions in a sequence to one another in situations in which these relations are likely to be meaningful. Converging evidence for these conclusions has emerged from studies across several laboratories and from different paradigms, including not only visual habituation measures of infants' event representation, but also experiments that manipulate infants' overt social responses.

These findings raise the question of how infants understand the intentions behind the structure. Infants represent certain actions in terms of the relation between the agent and the object at which her actions are directed, but what do they understand about the nature of this relation? Adults both represent the behavioral regularities associated with action and infer the mental states that underlie these regularities. One possibility is that infants do the former but not the latter. For example, infants may understand the relation between agent and object in terms of likely subsequent actions (e.g., a person is likely to act on the object at which he or she has just looked), and this could then lead them to preferentially encode agent–object relations (see Woodward, 1998, 2003, in press).

Even if infants begin with a purely behavioral analysis of action structure, there are reasons to believe that this could be a step in the construction of mature conceptions of intention. A behavior-based analysis of agent–object relations that leads infants to attend to relations between agents and goals, and agents and the objects of their attention highlights the aspects of events that are likely to be useful for constructing richer knowledge about intentions. In other words, a behavioral analyses of action could provide the foundation for insights about the psychological correlates of action (*cf.* Whiten, 1994). As another example of this possibility, Gergely and Csibra (1998, 2003) hypothesized that infants' analysis of actions as rationally organized toward goals (teleological representations) exists prior to and provides a foundation for the construction of knowledge about the mental causes action: infants' initial teleological representations specify the relations between agents and observable goals given observable states of affairs. These representations are hypothesized to provide the foundation for the conceptual insight that goals are mental entities (desires) that are pursued with respect to mentally represented states of affairs (beliefs).

Thus, infants might hold relatively abstract expectations about certain kinds of motion, and yet not conceptualize these expectations in terms of mental states. Alternatively, infants may understand something about the inner states that drive action. In considering this possibility, I first raise two caveats. One is that there are aspects of mental life that even preschool-aged children do not yet fully understand, including the representational nature of beliefs (Flavell & Miller, 1998; Wellman, Cross, & Watson, 2001), and these are fundamental to mature folk conceptions of mental life. Therefore it seems extremely unlikely that infants possess the explicit and elaborate system of knowledge that is evident in later in childhood.

The other caveat is that the kinds of data that can be obtained from infants may never completely resolve this question. Older children's talk about the mind has been an invaluable source of evidence for their mentalistic understanding (Astington, 1993; Bartsch & Wellman, 1995), and of course such evidence is not available from infants. Because infants do not understand or use the words that name mental states, our best evidence will necessarily be their behavioral responses to observed actions, and these responses are often, if not always, open to the interpretation that infants represent (important) behavioral regularities, but not mental states (see Johnson, 2000).

Even given these concerns, there is evidence consistent with the conclusion that infants understand something about the inner correlates of observable actions. Mature folk psychology represents mental states as existing independent of immediate physical actions or connections, as residing within the individual agent, and as having the same form in oneself and in others. In their tracking of action information, infants evidence understanding of each of these aspects of mental experience by the end of the first year of life.

First, as discussed earlier, infants represent the non-physical relation between a person and the object of his or her attention, and further, represent goals as being independent of particular actions. Each of these abilities indicates that infants represent something more abstract than the immediate physical connections between agents and objects. Indeed, these two abilities have been widely viewed as evidence that infants understand mental states of attention (Leslie, 1993; Barresi & Moore, 1996; Tomasello, 1999; Johnson, 2000) and plans held in mind (Behne *et al.*, in press; Meltzoff, 1995; Carpenter, Nagell, & Tomasello, 1998).

Second, infants represent goals as attributes of individual agents. We have found that 9- and 13-month-old infants represent the identity of the agent as integral to the goal of an action (Sootsman & Woodward, 2004). We used a habituation paradigm, similar to our prior studies, in which infants viewed a person directing actions toward a particular goal object, and then saw test events in which the physical context was changed and the actor acted either in accordance with the prior goal or a new goal. Changing the identity of the actor between habituation and test disrupted infants' propensity to respond to the change in goal. In other words, infants did not to attribute the goal of the first actor to the second actor. Kuhlmeier, Wynn, and Bloom (2003) addressed a similar question from a different vantage point, asking whether 12-month-old infants would attribute enduring dispositions to individual agents. Infants viewed geometric shapes interacting in a animated film. One of these shapes repeatedly moved in a way that suggested it was harassing another. Infants expected the victim to subsequently selectively avoid the harasser when given a choice. Thus, infants apparently attributed a particular (negative) disposition to the victim that provided a basis for inferring the victim's actions in a novel context.

Third, as I review in more detail subsequently, there is a tight relation between infants' own experiences as agents and their understanding of others' goal-directed actions. Infants' understanding of familiar actions as goal-directed emerges at the same time as their own mastery of these actions becomes robust (Woodward, Sommerville, & Guajardo, 2001; see also Molina *et al.*, 2004), there are correlations between action production and action comprehension during these periods of emergence (Woodward & Guajardo, 2002; Brune, 2004; Sommerville & Woodward, 2005), and interventions that alter infants' own agentive experience also alter their representation of others' actions (Sommerville, Woodward, & Needham, in press). In addition, there are well-documented self-other connections in the other direction—infants shape their own goal-directed actions to match those of others from very early in life (Meltzoff & Moore, 1977; Meltzoff, 2002). These findings all suggest that infants represent their own and others' goals in similar formats, and further, that to the extent that infants are aware of their own inner intentional states, they may be able to attribute similar states to others.

Although none of these findings conclusively proves that infants understand mental states, they converge in suggesting that central aspects of mental state knowledge have begun to emerge in infancy. Ultimately, the question of whether infants "really" understand mental states may elude a definitive answer, in part because of limits on infants' channels of knowledge expression, and in part because the question will have different answers depending on which aspects of mature mental state knowledge are taken to be criterial. Nevertheless, and I think more importantly, we have learned a great deal about infants' emerging action knowledge. During the first year, infants' action knowledge includes some of central elements of mature systems of knowledge. Because infants represent the relevant aspects of action structure for understanding the intentions of others, their action knowledge is likely to be generative in at least two ways: First, it would provide the foundation for the subsequent development of more abstract and differentiated folk psychological knowledge. Second, it could account for the "smart" social learning described at the start of this chapter. Representing goal-directed actions, attentional relations, and higher order plans would lead infants to focus on the right aspects of situations to learn words, infer the referents of emotional expressions, and extract the meaningful components of others' actions (see Woodward, 2004a).

III. Origins of Infants' Action Knowledge

The findings just summarized provide an initial framework for the emergence of infants' action knowledge, and thereby raise the question of how this knowledge originates. Among the enduring debates in the fields of cognitive science and developmental psychology is the question of whether foundational aspects of conceptual structure are the expression of innate abstract knowledge systems, or instead the product of bottom–up learning and conceptual construction. This debate has been particularly active in the domain of social cognition, especially because the evidence for intentional action knowledge in infancy raises the possibility that there are innate contributors to this system of knowledge. As in the field in general, the initial nature-nurture debate on this issue has become more nuanced, focusing not on whether there are innate contributors to this system of knowledge, but instead, on what these might be and the relative contributions of innate and experiential factors.

Within this discussion, three general proposals about the ontogeny of intentional action knowledge have been elaborated, and I review them below. Although in their strongest form these proposals seem incompatible, ultimately it is possible that elements of each may be shown to contribute to infants' emerging action knowledge. At this point, the field has just begun to gather the data that test the limits of each proposal.

A. INNATE ABSTRACT CONCEPTS

A well-known proposal, articulated in an influential theoretical paper by Premack (1990), is that infants begin life with abstract systems for interpreting observed events as intentional. Following Premack's proposal, a number of similar proposals have emerged, each differing in important ways, but all beginning from the assumption that the innate intentionality-detection system exists independent of experience with real-world agents, and is triggered by particular patterns of motion. Kiraly and colleagues (2003) summarized these proposals as follows:

> Several theories propose innately based, abstract, and domain-specific representational systems specialized for identifying intentional agents…While these models differ in several important respects, they all assume an initially wide scope of entities…that infants can recognize as goal-directed from very early on (including unfamiliar actions of humans or unfamiliar agents with no human features). This generality in scope is due to the fact that these theories all postulate or imply sensitivity to abstract behavioural cues…that indicate agency…irrespective of previous experience with the types of agents or actions that exhibit these cues (2003, p. 753).

Premack proposed that the triggering cue is self-propelled motion, suggesting that any self moving object would be identified as an agent, and its behaviors then be interpreted as manifesting intentions, preferences, the capacity to learn, and other psychological phenomena (see also Luo & Baillargeon, 2004). In this case, a single behavioral cue triggers a system of innate beliefs. In a similar account, Johnson (2000; Shimizu and Johnson, 2004) have hypothesized that when an entity engages in contingent social interaction, infants infer that it is an agent that can attend to distant entities and act in goal-directed ways.

In other accounts, the behavioral triggers are also the embodiment of the hypothesized innate concept (*cf.* Leslie, 1993; Gergely & Csibra, 2003; Kiraly *et al.*, 2003). In other words, the innate concept explains particular behavioral patterns, and the system fires to any event that manifests these patterns. Gergely, Csibra, and colleagues (Gergely *et al.*, 1995; Csibra *et al.*, 1999, 2003; Gergely & Csibra, 2003) proposed seeing an entity move with apparently rational motion in pursuit of a goal triggers innate conceptions of rational action, and leads infants to expect that the entity will act rationally in new situations. Kiraly and colleagues (2003) have proposed that two additional cues have this effect (1) motion that causes a salient outcome, and (2) repeated, varied (equipotential) motions to the same goal. Each of these, organization of actions in terms of outcomes and equipotentiality, are hypothesized to be part of the content of innate conceptions of intentional action.

All of these accounts predict that infants possess relatively abstract conceptions of intentional action that they apply to any object that moves in the critical way, regardless of whether the object bears any resemblance to a real-world agent. A number of findings provide support for this prediction in older infants. One source of evidence derives from studies like those of Gergely, Csibra, and colleagues (Gergely *et al*., 1995; Csibra *et al*., 1999, 2003) and Kuhlmeier, Wynn, and Bloom (2003), that were summarized earlier. In these studies, 12- and sometimes 9-month-old infants respond to the motions of abstract figures on a computer screen in ways that suggest they have interpreted these motions as goal-directed.

A second source of evidence are findings by Johnson and colleagues suggesting that viewing an entity engage in socially contingent behaviors influences infants' subsequent propensity to regard the entity as an agent (Johnson, Slaughter, & Carey, 1998; Johnson, Booth, & O'Hearn, 2001; Shimizu & Johnson, 2004). To illustrate, Shimizu and Johnson (2004) introduced 12-month-old infants to a fuzzy green block. One group of infants first viewed the block responding contingently to the social bids of an experimenter (by beeping and moving). A comparison group viewed the block and heard it beep, but did not see an interaction between the object and the experimenter. Then both groups of infants saw the block approach (with apparently self-propelled motion) and make contact with one of two target objects. Infants were habituated to this event. Then, as in the procedure developed in Woodward (1998), the target objects' positions were reversed and infants were shown new-object test trials (the block moved to the same location to contact the other object) and new-location test trials (the block moved to the other location to contact the same object as during habituation). Infants who had not viewed the contingent social interaction did not differentiate between the test events, suggesting that they did not interpret the block's motions as goal-directed. Infants who had viewed the interaction, in contrast, looked reliably longer on new-object than new-location trials, suggesting that they did interpret the block's motions as goal-directed. Thus, these findings suggest that the social interaction led infants to construe the block as an agent.

These findings indicate that by the time infants are 9–12 months of age, they possess relatively abstract conceptions of goal-directed action that can be extended to unusual agents in some circumstances (but not all, see Meltzoff, 1995; Guajardo & Woodward, 2004). However, the further claim that this abstract knowledge is innate is largely untested. Because nearly all of the relevant findings pertain to infants who are 12 months of age or older, they may be the products of a developing system rather than an innate endowment (see Woodward, Sommerville, & Guajardo, 2001; Guajardo & Woodward, 2004). Infants' action knowledge may derive from experience with real-world agents and their actions, and become more abstract and flexible with development. As infants acquire knowledge about specific actions and the relations between them, they may extract regularities across actions, and then use these regularities to inform their interpretation of novel or ambiguous events.

If infants' initial conceptions of goal-directed action were abstract, then infants would broadly attribute goals from the start, either from birth, or from the first point at which they attribute goals to any entity. Several sources of evidence suggest that this is not the case. To start, as reviewed earlier, our prior findings showed that although infants interpret the familiar actions of people as goal-directed beginning as early as 5 or 6 months of age, they do not interpret the motions of inanimate objects or unfamiliar human actions in this way (Woodward, 1998, 1999). Furthermore, we found that infants' propensity to view a grasping event as goal-directed was modulated by the extent to which infants could identify the hand as part of a person (Guajardo & Woodward, 2004). It has been suggested (see Kiraly *et al.*, 2003; Shimizu & Johnson, 2004) that these failures are due to the fact that the unfamiliar actions and object motions in those studies, though carefully matched to the familiar human actions, did not possess the behavioral cues hypothesized to trigger infants' goal attribution. However, when these cues have been provided, younger infants still fail to interpret unfamiliar events as goal-directed. Csibra and colleagues (1999) tested 6-month-old infants with computer animations of rational motion around barriers, and these infants, unlike older infants, failed to respond systematically. Kiraly and colleagues (2003) tested 6-, 8-, and 10-month-olds in a paradigm designed to assess whether providing a salient action effects would lead infants to interpret an unusual action as goal-directed. Their findings suggested that 8- and 10-month-olds, but not 6-month-olds, responded to the events as goal-directed (but see Heineman-Pieper & Woodward (2003) for an alternative interpretation of the older infants' responses). Jovanovic and colleagues (2004) conducted similar studies with 6-month-old infants and reported mixed findings. Infants apparently interpreted a human hand gesture as goal-directed, but did not respond to grasping by a mechanical claw in this way.

Thus, the available findings are most consistent with the conclusion that infants begin with local understandings of goal-directed action that become broader over the course of the first year of life. Further research is needed to thoroughly test this conclusion, however. There are several hypotheses concerning the potential behavioral triggers to innate knowledge, and current studies may not have presented young infants with the right one. As a possible case in point, Luo and Baillargeon (2004) reported that 6-month-old infants apparently interpreted a self-propelled box as a goal-directed agent. However, this report is inconsistent with the findings of Shimizu and Johnson (2004) who found that infants did not treat a self-propelled block as goal-directed.

B. SOCIALLY BASED COGNITIVE LEARNING

The counterpoint to strongly nativist proposals are accounts that consider conceptual structure to be the product of bottom–up learning and conceptual abstraction. Researchers have made important strides in investigating the role of these processes in older children's cognitive development (e.g., Gentner & Medina, 1998; McClelland & Siegler, 2001; Samuelson & Smith, 2000) and have begun to consider the role of these processes during infancy (Mandler, 1998; Baillargeon, 2002; Cohen, Chaput & Cashon, 2002; Rakison & Oakes, 2003). It has been hypothesized that associative learning (Rakison & Poulin-Dubois, 2001), statistical learning, and structure mapping (Baldwin & Baird, 2001) each contribute to infants' emerging social cognition, but these proposals have not been directly tested.

The proposal that intentional action knowledge is the product of cognitive learning has also been developed by researchers who focus on the influence of social experiences on infants' social cognition (Tomasello, 1995; Barresi & Moore, 1996; Carpenter, Nagell & Tomasello, 1998; Carpendale & Lewis, 2004). These accounts generally focus on 9–12 months of age, a period during which infants begin to engage in more robust and well-structured interactions with adults, including shared attention, communicative gestures, game playing, and imitation. The onset of these new ways of interacting is striking, and has been taken by many to signal a "social-cognitive revolution", and, in particular, newly emerged understandings of intentional action (Tomasello, 1995).

These emerging interactive patterns are taken not only as the sign of new social-cognitive abilities, but also as the means by which these abilities arise. Specifically, it has been hypothesized that engagement in triadic interactions sets the conditions for infants' discovery of others' intentions (Barresi & Moore, 1996; Carpenter, Nagell & Tomasello, 1998; Carpendale & Lewis, 2004). In these interactions, adult and child attend to (and sometimes act on) the same aspect of the environment, thus providing the opportunity for infants to align their own actions and intentional states with the observed actions of others. This alignment of self and other has been suggested to provide infants with the structural components necessary to infer intentional relations between others and the objects of their attention. Some propose that infants begin with a merged representation of their own intentional relations and those of others, from which they then construct differentiated concepts of themselves and others as independent agents (Barresi & Moore, 1996; Carpendale & Lewis, 2004). Other theorists suggest something more like a process of analogical extension from self to other; that is, that infants infer that others have internal experiences analogous to their own (Carpenter, Nagell, & Tomasello, 1998; Meltzoff, 2002).

These accounts are both plausible and potentially deeply informative about cognitive development in infancy. They take seriously the roles of rich environmental structure and cognitive learning processes in infants' developing social cognition. Moreover, they are consistent with what is known about the subsequent development of this system of knowledge: there are widely documented effects of social environments on the development of folk psychological knowledge in older children (see Dunn, 1999; Repacholi & Slaughter, 2003; Carpendale & Lewis, 2004 for reviews). However, little headway has been made in empirical tests of these accounts with respect to infant social cognition.

One reason for this is that these accounts often rely on social responsiveness as an index of intentional understanding, and this compromises their ability to draw strong conclusions about the nature of infants' social cognition. There has long been debate about whether and when children's social responses can serve as evidence about their comprehension of others' intentions. Indeed, even among those who take social responsiveness as evidence of underlying action representations, there is serious disagreement about which behaviors "count". For example, Tomasello (1995) has suggested that the gaze-following and joint attention behaviors of 9- to 12-month-olds reflect an understanding of others' intentions. But Barresi and Moore (1996) (see also Moore & Corkum, 1994) point out that these behaviors might be shaped by reinforcement or supported by low-level processes that do not require an understanding of attentional relations. More generally, overt social behaviors are likely to be influenced by processes at several levels, and therefore there is no straightforward relation between the behavior and a particular mental representation. In addition, reliance on organized triadic behaviors as evidence for organized social cognition is likely to underestimate what young infants know. Infants younger than 9–12 months of age do not yet participate in well-organized triadic interactions, however, as summarized previously, they do understand critical aspects of intentional action (see Wilson & Woodward, 2003). Clearer evidence about infants' social cognition is needed, and, as described earlier, habituation paradigms have begun to yield such evidence.

A second barrier to testing the social-construction theories is that essentially the same behaviors are considered to be both cause and effect. Aspects of social responsiveness are seen as both contributors to infants' emerging action knowledge and evidence for the existence of this knowledge. For example, it has been suggested that engaging in triadic interactions, in which infant and adult share attention on an object, both provides a means by which infants discover the adult's attentional relations, and constitute evidence that infants understand these relations. It is entirely plausible that social experience contributes to infants' underlying knowledge about action, but to test this possibility (and determine the specific ways in which it may be true), an independent measure of social knowledge is required.

Brune and I addressed this empirical gap by combining measures of social responsiveness with habituation measures of infants' action understanding (Brune, 2004). Our goal was to determine whether infants' action understanding is correlated with their social responsiveness during a period when both are undergoing important changes, that is between 9 and 12 months. If they are correlated, then this provides initial evidence that experience may contribute to the development of infants' social cognition. This result would then pave the way for investigations that pinpoint the direction of causation.

We tested 10-month-old infants, who, based on our prior findings, are just beginning to understand the attentional relations expressed by gazing and pointing. Each infant was tested in two habituation procedures on different days, one assessing their understanding of gaze, the other their understanding of pointing (as described in Section II.C). The infants were variable in their responses on test trials, showing no systematic group level pattern. Our question was whether this variability correlated with infants' social responsiveness. To this end, we also tested each infant in laboratory procedures assessing several aspects of their social responsiveness, including gaze-following, engagement in shared attention with parents, and ability to point at objects.

We found relations between infants' habituation responses and their social behaviors, but these relations were different than might be expected. The theories outlined previously predict relatively global relations between infants' actions and social cognition—all of the behaviors marking the social-cognitive revolution are supposed to lead to (and express) a unified concept of intentional action. Our findings tell a different story. The relations between infants' social behaviors and social cognition appear to relatively action-specific. Infants' understanding of pointing was significantly related to their own ability to point, replicating our earlier work (Woodward & Guajardo, 2002), but was not correlated with the other measures of social responsiveness. In contrast, infants' understanding of gaze correlated with the extent to which they engaged in shared attention with caretakers, but not with point production or the other social behaviors we assessed. Neither aspect of action understanding related to infants' propensity to follow gaze, consistent with our prior findings that orienting to actions does not always travel with understanding the action as object-directed (Woodward & Guajardo, 2002; Woodward, 1998, 2003).

Documenting concurrent correlations between social cognition and social responsiveness breaks new ground, and indicates that visual habituation data may (at least in this case) be reliable enough to use in further investigations. However, this evidence alone cannot tell us whether social experiences contribute to the development of social cognition. Longitudinal work is needed to determine the direction(s) of causation that accounts for the relations we observed. It seems likely that the influence is bi-directional by early childhood. However, at the earliest points in development the influence may run in only one direction. Infants' first insights into the intentional states of others may derive from social experiences that are organized by factors other than the infants' social cognition, including the structured behavior of parents and lower level processes that subserve infants' responses to them.

Our findings also begin to shed light on the nature of the relations between social interaction and social cognition during early development, suggesting that they exist at the level of particular aspects of social cognition and particular kinds of social behaviors. Producing points relates to (and may therefore contribute to) infants' understanding the significance of other people's points (see Woodward, in press; Woodward & Guajardo, 2002). Engaging in shared attention relates to (and may therefore contribute to) understanding the significance of others' gaze. These findings raise the question of when and how infants' action knowledge becomes more general.

C. EMBODIED ACTION REPRESENTATIONS

It has long been hypothesized that the experience of being an intentional agent contributes fundamentally to the development of concepts of intention. This general proposal plays a role in several theoretical accounts, including those reviewed in the previous section. Independent of, but relevant to, these accounts, there has been renewed interest in the possibility that experience as an agent informs understanding of other agents because recent findings indicate the existence of shared neuro-cognitive representations for action production and action perception (Gallese & Goldman, 1998; Blakemore & Decety, 2001; Meltzoff & Prinz, 2002).

Evidence for these shared representations has emerged from several diverse research programs. Single cell recordings in monkeys have revealed a class of motor neurons that fire both when the animal is about to produce a particular goal-directed action and when the animal observes a person produce that action (Rizzolatti, Fogassi, & Gallese, 2000). These "mirror neurons" respond to specific natural goal-directed actions (e.g., grasping or tearing), and also have been found for novel goal-directed actions, in particular tool use, following training (Rizzolatti & Arbib, 1998; Rizzolatti, Fogassi, & Gallese, 2000). Using neuroimaging techniques, other researchers have found areas in the adult human brain that have a similar mirroring function (Iaccobini *et al*., 1999; Grezes & Decety, 2001). Furthermore, behavioral studies with adults have revealed overlapping cognitive representations that subserve the perception and production of simple actions, as evidenced by interference across these two modalities (Hommel *et al.*, 2001).

These shared representations may exist primarily to monitor self-produced actions, a critical function for the prospective control of action, and because of this function they also fire in response to actions produced by others (Rizzolatti, Fogassi, & Gallese, 2000). It has been hypothesized that these systems contribute to the subjective sense of one's own intentionality (Frith, 2002) as well as to the perception of others' intentional actions (Gallese, 2001; Frith, 2002), and, by extension, to mind reading (Gallese & Goldman, 1998; Blakemore & Decety, 2001; Meltzoff & Prinz, 2002). This speculation is in line with more general proposals about embodied cognition, in particular, the suggestion that sensorimotor representations provide structure for "off-line" cognition, in this case, the interpretation of observed actions (see Wilson, 2001).

Although there is no direct evidence from infants that is comparable to the single cell work with monkeys or the fMRI work with adults, there are reasons to suspect that mirroring systems exist in infants. First, infants imitate others' actions from birth, thus suggesting an automatic resonance between their own actions and those of others (Meltzoff & Moore, 1977; Meltzoff, 2002; Meltzoff & Prinz, 2002). Second, if mirroring systems are important for monitoring and prospective control of actions, then there is every reason to expect their existence in organisms who engage in complex prospective action, and infants do this (see Hofsten, 2004). Beginning early in the first year, infants become able to control actions, such as reaching and grasping, that are organized with respect to external goals (Bertenthal & Clifton, 1998; Clearfield & Thelen, 2001; Hofsten, 2004), and it is possible that the systems for monitoring these actions involve mirror representations.

If mirroring systems exist in infancy, then infants' emerging ability organize their own actions in service of goals would create representations that could in turn structure their perception of others' actions. On the basis of neonatal imitation, Meltzoff has proposed that this is the case (Meltzoff, 2002). However, a thorough test of this hypothesis requires measures of infants' action representations as well as their actions. Visual habituation measures like the ones we have developed provide a tool for such investigations.

Indeed, several findings from our laboratory suggested to us that infants' initial action representations derive from their own actions. To start, infants' initial sensitivity to object-directed action appears to be limited to human actions, and actions in infants' own repertoires (e.g., grasping) (Woodward, 1998, 1999; Guajardo & Woodward, 2004; but see Jovanovic *et al.*, 2004). In addition, infants become sensitive to the goal-structure of actions during the age periods that they are mastering production of these same actions (Woodward, Sommerville, & Guajardo, 2001; see also Molina *et al.*, 2004). Finally, at these transitional periods, infants' mastery of a particular action is related to their sensitivity to the goal-structure of that same action in others, as was reviewed previously for infants' understanding of object-directed points and means-end sequences (Woodward & Guajardo, 2002; Brune, 2004; Sommerville & Woodward, 2005).

As a strong test of whether motor experience affects infants' responses to observed actions, Sommerville, Needham and I conducted an intervention study (Sommerville, Woodward, & Needham, in press). Our goal was to scaffold infants' ability to produce a new goal-directed action, and then test whether this experience affected their response to an observed action. We tested 3-month-old infants. At this age, infants are very limited in their ability to reach for objects, and our prior findings suggest that infants at this age do not represent observed grasps as object-directed. To support infants' ability to apprehend objects, we used velcro-bearing mittens, developed by Needham and her colleagues, which enable infants to pick up toys by swiping at them (Needham, Barrett, & Peterman, 2002). Needham and colleagues have found that with practice 3-month-old infants began to use the mittens in an apparently planful manner, and that this experience had enduring effects on infants' object exploration behaviors (Needham, Barrett, & Peterman, 2002).

We gave infants in the experimental group a few minutes of practice using the mittens. Small toys were placed on a surface in front of the infant, and he or she was allowed to swipe at the toys until one was picked up by the mitten. Infants generally found this game to be highly engaging. They swiped eagerly at the toys, and watched closely as their mittened hands moved the object. Then infants were tested in a habituation paradigm like the one depicted in Figure 1, except that the actor wore a mitten that matched the infants, and the toys were larger replicas of the ones the infant had acted on. Our goal was to maximize the similarity between the infants' experience and the observed events because similarity has been shown to facilitate mental comparison (Gentner & Medina, 1998). A control group of infants participated in the habituation paradigm before engaging in the mittens task.

As Needham and her colleagues had found, the mittens facilitated infants' manipulation of the toys. Infants spent a greater proportion of time in coordinated gaze and manual contact with the toys when they were wearing the mittens vs. when they were not. Critically, mittens experience also affected infants' responses to the habituation events. Infants in the experimental condition looked reliably longer on new-object trials than on new-side trials, that is, they attended to the relation between the actor and her goal. In contrast, infants in the control condition did not differentiate between the two kinds of test trials. Moreover, in the experimental condition infants' relative preference for the new-object event was correlated with the amount of coordinated gaze and manual contact on the toys while wearing mittens, but not with their total amount of visual contact with the toys, or their amount of coordinated gaze and manual contact when they were given the chance to act on the toys without the mittens. Thus, infants' responses to the habituation events reflected the extent to which they had engaged in organized mittened actions on the toys rather than perceptual highlighting of the toys or individual differences in motor development.

In a later study, we replicated the finding that active mittens practice leads infants to represent observed mittened reaches as goal-directed (Woodward, 2004b). We also found that infants did not show this effect when they viewed a barehanded reach rather than a mittened reach. In other words, the few minutes of mittens practice infants received seemed to have relatively circumscribed effects. Infants did not readily generalize from the mittened action to other manual actions. This finding suggests that infants build relatively specific action representations to start. A question for future studies is whether infants would generalize more broadly as they gain expertise with an action.

Certainly, additional studies are needed to investigate the effects of acting on action perception across ages and across actions. Even so, these findings begin to support a developmental account in which infants' action knowledge is experience-driven, constrained by developmental progressions in the motor domain, and dependent on innate pathways for establishing mirroring systems. If this hypothesis is correct, as aspects of intentional structure emerge in the behavioral control systems of the infant, they may become available for action perception. The critical questions then become how (and whether) these effects on action perception contribute to subsequent conceptual development.

D. CONCLUSIONS: ORIGINS

The proposals outlined here highlight distinct hypotheses concerning the origins of intentional understanding and the relation between real-world knowledge and abstract concepts. One hypothesis is that core components of mature knowledge systems are innately specified in the form abstract principles (e.g., Premack, 1990; Gergely *et al.*, 1995; Csibra *et al.*, 1999). On this view, infants also accrue knowledge about real-world actions. This knowledge may come to inform infants' application of the innate principles in some circumstances, but it does not contribute to the formation of abstract knowledge about intentional action (see Gergely *et al.*, 1995; Kiraly *et al.*, 2003 for discussions).

A second hypothesis is that real-world action knowledge provides the developmental basis for more abstract conceptions of intentional action (Guajardo & Woodward, in press; Woodward, Sommerville, & Guajardo, 2001; see also Baldwin & Baird, 2001). Under this view, learning, cognitive comparison, and conceptual abstraction contribute fundamental structure to intentional understanding. Cognitive development in this domain may draw heavily on social experiences, as well as the linking of information about ones own and others' actions (Carpenter, Akhtar, & Tomasello, 1998; Meltzoff, 2002; Carpendale & Lewis, 2004). On this account, the innate contributors to infants' action knowledge might not be abstract concepts, but instead shared neuro-cognitive representations for perceiving and producing actions.

Either kind of account could, in principle, explain the findings of relatively abstract expectations about goal-directed motions at the end of the first year of life. Although these findings are often taken as evidence for innate processing modules, they could also reflect the products of learning and abstraction. The two accounts differ in their predictions about the initial form of infants' action knowledge. As reviewed previously, much of the current evidence supports the conclusion that infants begin with relatively local representations of goal-directed action, only later showing signs of more abstract expectations. These findings are most consistent with the view that infants begin by tracking regularities in real-world actions, including their own, and from these regularities construct more abstract expectations about intentional actions. However, new findings are beginning to test the limits of this conclusion (e.g., Luo & Baillargeon, 2004). Further empirical investigations are needed to determine the scope and generality of young infants' action knowledge.

In addition, it may be useful to consider alternative models of the relation between abstract and real-world knowledge in development. One possibility is that these kinds of knowledge are independent contributors to infants' conceptual structure early in life. For important aspects of development, nature may provide a broad arsenal, including in this case, pattern detection systems, a rich database of observed real-world regularities, and the ability to relate representations of one's own actions to those of others. Each of these may emerge independently, and contribute to infants' abilities to represent the intentional structure of certain kinds of events. In this case, the open questions include how these components interact at different points in development. When abstract expectations contradict real-world regularities (e.g., when inanimate objects move in agent-like ways), on which basis will infants respond and why? Do these components remain independent throughout life, or do they become integrated with development?

An alternative possibility is suggested by considering discussions about the innate foundations of grammar. Infants might possess abstract, innate conceptions of intentional action, but these conceptions may only be expressed as infants discover the real-world cases that embody them. This is analogous to the hypothesis that children are born with innate universal grammar, which is only expressed once they have acquired the linking rules that specify how this grammar is embodied in their native language. If this were the case, then the findings that suggest initial specificity in infants' attributions of goal-directedness would reflect the gradual emergence of the "linking rules" rather than the gradual emergence of abstract knowledge. It has proven difficult to distinguish between these possibilities in the case of language (e.g., Tomasello, 2001; Fisher, 2002), and similar debates may well arise in the domain of infants' action knowledge.

IV. Final Remarks

To understand the development of any system it is necessary to describe the states of the system at different points in time and the processes by which the system changes over time. The study of infant social cognition is relatively new, yet significant progress has been made toward the first of these goals, and first steps exist in the pursuit of the second.

As has been illustrated throughout this review, there are several points of broad consensus among researchers of infant intentional understanding, despite strong differences in theoretical assumptions and methodology. Converging evidence from across the field has shown that infants represent other people's actions in ways that are important for intentional understanding, that is, in terms of the relations between agents and their goals, agents and the objects of their attention, and subgoals and higher order plans. By the end of the first year of life, infants can flexibly interpret observed motions as goal-directed or not, or as directed at goals at differing hierarchical levels, based on contextual cues, and they track goals as the attributes of individual agents. These general conclusions about infant cognition dovetail with work on social cognition during the second year of life. By 18–24 months of age, children's responses to and learning from the actions of others is mediated by a rich system of intentional action knowledge. Elements of this system have now been traced to the first year of life.

There are also areas of heated debate, in particular those concerning the origins and initial development of infants' action knowledge. These issues currently focus much of the research on infant social cognition. In pursuing these issues, we should keep in mind an important insight from research on intentional action knowledge at later points in life: folk psychology is a system of interrelated concepts, and its application to real-world events involves structural principles at several levels of analysis (Baldwin & Baird, 2001). The questions for infancy research should be framed not in terms of when infants get "it", but instead in terms of the emergence of the system of concepts that becomes mature folk psychology. In later childhood, the folk psychological knowledge system emerges, at least in part, as a function of children's social experiences, including their conversations with parents, interactions with siblings and peers, and participation in a cultural context (see Dunn, 1999). It is reasonable to hypothesize that similar experiences contribute to infants' knowledge.

Infant cognition is often assumed to be qualitatively discontinuous from cognition in early childhood, but research in the social domain may reveal continuity, both in the environmental contributors to knowledge and in the nature of knowledge itself. Indeed, findings from Wellman and his colleagues (in press) provide compelling support for this possibility. These researchers followed infants they had tested in a visual habituation paradigm (the Phillips, Wellman, and Spelke study, described previously), and tested these children at age 4 on an explicit theory of mind scale. Infants' visual responses at 12 months predicted their theory of mind performance at 4 years, independent of general intelligence. These findings are among the first to demonstrate continuity in a knowledge system from infancy to preschool, and they suggest that the social domain may provide an opportunity to develop an account that bridges the longstanding theoretical divide between infancy and the rest of life.

ACKNOWLEDGEMENTS

Preparation of this chapter was supported by a James McKeen Cattell sabbatical fellowship.

REFERENCES

Astington, J. W. (1993). *The child's discovery of the mind*. Cambridge, MA: Harvard University Press.

Baillargeon, R. (2002). The acquisition of physical knowledge in infancy: a summary in eight lessons. In U. Goswami (Ed.), *Blackwell handbook of childhood cognitive development* (pp. 47–83). Malden, MA: Blackwell.

Bakeman, R., & Adamson, L. B. (1984). Coordinating attention to people and objects in mother-infant and peer-infant interaction. *Child Development*, *55*, 1278–1289.

Baldwin, D. A. (1995). Understanding the link between joint attention and language. In P. Dunham (Ed.), *Joint attention: Its origins and role in development* (pp. 131–158). Hillsdale, NJ: Erlbaum.

Baldwin, D. A., & Baird, J. A. (2001). Discerning intentions in dynamic human action. *Trends in Cognitive Sciences*, *5*, 171–178.

Baldwin, D. A., & Moses, J. A. (2001). Links between social understanding and early word learning: Challenges to current accounts. *Social Development*, *10*, 311–329.

Baron-Cohen, S. (1995). *Mindblindness: An essay on autism and theory of mind*. Cambridge, MA: MIT Press.

Barresi, J., & Moore, C. (1996). Intentional relations and social understanding. *Behavioral and Brain Sciences*, *19*, 107–154.

Bartsch, K., & Wellman, H. M. (1995). *Children talk about the mind*. New York: Oxford University Press.

Behne, T., Carpenter, M., Call, J., & Tomasello, M. (in press). Unwilling versus unable? Infants' understanding of intentional action. *Developmental Psychology*.

Bekkering, H., Wohlschlaeger, A., & Gattis, M. (2000). Imitation of gestures in children is goal-directed. *Quarterly Journal of Experimental Psychology*, *53*, 153–164.

Bellagamba, F., & Tomasello, M. (1999). Re-enacting intended acts: Comparing 12- and 18-month-olds. *Infant Behavior and Development*, *22*, 277–282.

Bertenthal, B., & Clifton, R. K. (1998). Perception and action. In W. Damon, D. Kuhn, & R. Siegler (Eds.), *Handbook of child psychology, Volume 2: Cognition, perception and language* (pp. 51–102). New York: Wiley.

Blakemore, S., & Decety, J. (2001). From the perception of action to the understanding of intention. *Nature Reviews: Neuroscience*, *2*, 561–567.

Brooks, R., & Meltzoff, A. M. (2002). The importance of eyes: How infants interpret adult looking behavior. *Developmental Psychology*, *38*, 958–966.

Brune, C. W. (2004). *The origins of joint attention: Relations between social knowledge, social responsiveness, and attentional control*. Unpublished Doctoral Dissertation. University of Chicago.

Butler, S., Caron, A., & Brooks, R. (2000). Infant understanding of the referential nature of looking. *Journal of Cognition and Development*, *1*, 359–377.

Butterworth, G., & Jarrett, N. (1991). What minds have in common is space: Spatial mechanisms serving joint visual attention in infancy. *British Journal of Developmental Psychology*, *9*, 55–72.

Carpendale, J. I. M., & Lewis, C. (2004). Constructing an understanding of mind: The development of children's social understanding within social interaction. *Behavioral and Brain Sciences*, *27*, 79–151.

Carpenter, M., Akhtar, N., & Tomasello, M. (1998). Fourteen through eighteen month old infants differentially imitate intentional and accidental actions. *Infant Behavior and Development*, *21*, 315–330.

Carpenter, M., Nagell, K., & Tomasello, M. (1998). Social cognition, joint attention and communicative competence from 9 to 15 months of age. *Monographs of the Society for Research in Child Development*, *63*(4).

Clearfield, M. W., & Thelen, E. (2001). Stability and flexibility in the acquisition of skilled movement. In C. A. Nelson & M. Luciana (Eds.), *Handbook of developmental cognitive neuroscience* (pp. 253–266). Cambridge, MA: MIT Press.

Cohen, L. B., Chaput, H. H., & Cashon, C. H. (2002). A constructivist model of infant cognition. *Cognitive Development*, *17*, 1323–1343.

Csibra, G., Biro, S., Koos, O., & Gergely, G. (2003). One-year-old infants use teleological representations of actions productively. *Cognitive Science*, *27*, 111–133.

Csibra, G., Gergely, G., Biro, S., Koos, O., & Brockbank, M. (1999). Goal attribution without agency cues: The perception of "pure reason" in infancy. *Cognition*, *72*, 237–267.

Dunn, J. (1999). Making sense of the social world: Mindreading, emotion and relationships. In P. D. Zelazo, J. W. Astington, & D. R. Olson (Eds.), *Developing theories of intention* (pp. 229–242). Mahwah, NJ: Erlbaum.

Dunphy-Lelii, S., & Wellman, H. R. (in press). Infants' understanding of occlusion of others' line-of-sight: Implications for an emerging theory of mind. *European Journal of Developmental Psychology*.

Fisher, C. (2002). The role of abstract syntactic structure in language acquisition. *Cognition*, *82*, 259–278.

Flavell, J. H., & Miller, P. H. (1998). Social cognition. In W. Damon, D. Kuhn, & R. S. Siegler (Eds.), *Handbook of child psychology II: Cognition, perception and language* (pp. 851–898). New York: Wiley.

Frith, C. (2002). Attention to action and awareness of other minds. *Consciousness and Cognition*, *11*, 481–487.

Gallese, V. (2001). The "shared manifold" hypothesis. *Journal of Consciousness Studies*, *8*, 33–50.

Gallese, V., & Goldman, A. (1998). Mirror neurons and the simulation theory of mind-reading. *Trends in Cognitive Sciences*, *2*, 493–501.

Gentner, D., & Medina, J. (1998). Similarity and the development of rules. *Cognition*, *65*, 263–297.
Gergely, G., Bekkering, H., & Kiraly, I. (2001). Rational imitation in preverbal infants. *Nature*, *415*, 755.
Gergely, G., & Csibra, G. (1998). The teleological origins of mentalistic action explanations: A developmental hypothesis. *Developmental Science*, *1*, 255–259.
Gergely, G., & Csibra, G. (2003). Teleological reasoning in infancy: The naive theory of rational action. *Trends in Cognitive Sciences*, *7*, 287–292.
Gergely, G., Nadasdy, Z., Csibra, G., & Biro, S. (1995). Taking the intentional stance at 12 months of age. *Cognition*, *56*, 165–193.
Grezes, J., & Decety, J. (2001). Functional anatomy of execution, mental simulation, observation, and verb generation of actions: A meta-analysis. *Human Brain Mapping*, *12*, 1–19.
Guajardo, J. J., & Woodward, A. L. (2004). Is agency skin-deep? Surface attributes influence infants' sensitivity to goal-directed action. *Infancy*, *6*, 361–384.
Hauf, P., Elsner, B., & Aschersleben, G. (2004). The role of action effects in infant's action control. *Psychological Research*, *68*, 115–125.
Heineman-Pieper, J., & Woodward, A. (2003). Understanding infants' understanding of intentions: Two problems of interpretation (A reply to Kiraly *et al*., 2003). *Consciousness and Cognition*, *12*, 770–772.
Hofsten, C. v. (2004). An action perspective on motor development. *Trends in Cognitive Sciences*, *8*, 266–272.
Hommel, B., Musseler, J., Aschersleben, G., & Prinz, W. (2001). The theory of event encoding. *Behavioral and Brain Sciences*, *24*, 847–937.
Iacoboni, M., Woods, R. P., Brass, M., Bekkering, H., Mazziotta, J. C., & Rizzolatti, G. (1999). Cortical mechanisms of human imitation. *Science*, *286*, 2526–2528.
Johnson, S. C. (2000). The recognition of mentalistic agents in infancy. *Trends in Cognitive Sciences*, *4*, 22–28.
Johnson, S. C., Booth, A., & O'Hearn, K. (2001). Inferring the goals of a nonhuman agent. *Cognitive Development*, *16*, 637–656.
Johnson, S., Slaughter, V., & Carey, S. (1998). Whose gaze will infants follow? The elicitation of gaze-following in 12-month-olds. *Developmental Science*, *1*, 233–238.
Jovanovic, B., Kiraly, I., Elsner, B., Gergely, G., Prinz, W., & Aschersleben, G. (2004). *The role of effects for infants' perception of action goals*. Unpublished manuscript.
Kiraly, I., Jovanovic, B., Prinz, W., Aschersleben, G., & Gergely, G. (2003). The early origins of goal attribution in infancy. *Consciousness and Cognition*, *12*, 752–769.
Kuhlmeier, V., Wynn, K., & Bloom, P. (2003). Attribution of dispositional states by 12-month-olds. *Psychological Science*, *14*, 402–408.
Leslie, A. M. (1993). ToMM, ToBY and agency: Core architecture and domain specificity. In L. A. Hirschfeld & S. A. Gelman (Eds.), *Mapping the mind: Domain specificity in cognition and culture* (pp. 119–148). Cambridge: Cambridge University Press.
Luo, Y., & Baillargeon, R., (2004). *Young infants' reasoning about a novel agent*. Paper presented at the 14th Biennial International Conference on Infant Studies, Chicago.
Mandler, J. M. (1998). Representation. In W. Damon (Series Ed.) & D. Kuhn & R. S. Siegler (Vol. Ed.), *Handbook of child psychology: Vol. 2. Cognition, perception, and language* (5th ed., pp. 255–308). New York: Wiley.
McClelland, J. L., & Siegler, R. S. (Eds.) (2001). *Mechanisms of cognitive development. Carnegie Mellon symposia on cognition*. Mahwah, NJ: Lawrence Erlbaum Associates.
Meltzoff, A. N. (1995). Understanding the intentions of others: Re-enactment of intended acts by 18-month-old children. *Developmental Psychology*, *31*, 838–850.

Meltzoff, A. N. (2002). Imitation as a mechanism of social cognition: Origins of empathy, theory of mind, and the representation of action. In U. Goswami (Ed.), *Blackwell handbook of childhood cognitive development* (pp. 6–25). Malden, MA: Blackwell Publishers.

Meltzoff, A. N., & Moore, M. K. (1977). Imitation of facial and manual gestures by human neonates. *Science*, *198*, 75–78.

Meltzoff, A. N., & Prinz, W. (Eds.) (2002). *The imitative mind: Development, evolution, and brain bases*. Cambridge: Cambridge University Press.

Molina, M., Van de Walle, G., Condry, K., & Spelke, E. S. (2004). The animate-inanimate distinction in infancy: Developing sensitivity to constraints on human action. *Journal of Cognition and Development*, *5*, 399–426.

Moore, C., & Corkum, V. (1994). Social understanding at the end of the first year of life. *Developmental Review*, *14*, 349–372.

Moses, L., Baldwin, D. A., Rosicky, J. G., & Tidball, G. (2001). Evidence for referential understanding in the emotions domain at 12 and 18 months. *Child Development*, *72*, 718–735.

Needham, A., Barrett, T., & Peterman, K. (2002). A pick-me-up for infants' exploratory skills. *Infant Behavior and Development*, *25*, 279–295.

Onishi, K. H., & Baillargeon, R. (2004). *False-belief understanding with invisible displacement in 15.5-month-old infants*, Paper presented at the 14th Biennial International Conference on Infant Studies, Chicago, IL.

Phillips, A. T., Wellman, H. M., & Spelke, E. S. (2002). Infants' ability to connect gaze and emotional expression to intentional action. *Cognition*, *85*, 53–78.

Piaget, J. (1953). *The origins of intelligence in the child*. London: Routledge & Kegan Paul.

Premack, D. (1990). The infant's theory of self-propelled objects. *Cognition*, *36*, 1–16.

Rakison, D. H., & Oakes, L. M. (2003). *Early category and concept development: Making sense of the blooming, buzzing confusion*. London: Oxford University Press.

Rakison, D. H., & Poulin-Dubois, D. (2001). Developmental origin of the animate-inanimate distinction. *Psychological Bulletin*, *127*, 209–228.

Repacholi, B., & Slaughter, V. (2003). *Individual differences in theory of mind*. New York, NY: Psychology Press.

Rizzolatti, G., & Arbib, M. A. (1998). Language within our grasp. *Trends in Neurosciences*, *21*, 188–194.

Rizzolatti, G., Fogassi, L., & Gallese, V. (2000). Cortical mechanisms subserving object grasping and action recognition: A new view on the cortical motor functions. In M. S. Gazzaniga (Ed.), *The new cognitive neurosciences* (2nd ed., pp. 539–552). Cambridge, MA: MIT Press.

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, *274*, 1926–1928.

Samuelson, L. K., & Smith, L. B. (2000). Grounding development in cognitive processes. *Child Development*, *71*, 107–118.

Scaife, M., & Bruner, J. S. (1975). The capacity for joint visual attention in the infant. *Nature*, *253*, 265–266.

Schaffer, H. R. (1984). *The child's entry into a social world*. London: Academic Press.

Schank, R. C., & Abelson, R. P. (1977). *Scripts, plans, goals and understanding*. Hillsdale, NJ: Erlbaum.

Searle, J. R. (1983). *Intentionality: An essay in the philosophy of mind*. Cambridge: Cambridge University Press.

Shimizu, Y. A., & Johnson, S. C. (2004). Infants' attribution of a goal to a morphologically unfamiliar agent. *Developmental Science*, *7*, 425–430.

Sodian, B., Schoepner, B., & Metz, U. (2004). Do infants apply the principle of rational action to human agents? *Infant Behavior and Development*, *27*, 31–41.

Sodian, B., & Thoermer, C. (2004). Infants' understanding of looking, pointing and reaching as cues to goal-directed action. *Journal of Cognition and Development*, *53*, 289–316.

Sommerville, J. A., & Woodward, A. L. (2005). Pulling out the intentional structure of action: The relation between action production and action processing in infancy. *Cognition*, *95*, 1–30.

Sommerville, J. A., Woodward, A. L., & Needham, A. (in press). Action experience alters 3-month-old infants' perception of others' actions, *Cognition*.

Sootsman, J., & Woodward, A. L. (2004). *Nine-month-old infants connect action goals to individual agents*. Poster presented at the 14th Biennial International Conference on Infant Studies, Chicago, IL.

Spaepen, E., & Spelke, E. (2004). Will any doll do? 12-month-olds' reasoning about goal objects. *Poster presented at the 14th Biennial International Conference on Infant Studies*, Chicago, IL.

Tomasello, M. (1995). Joint attention as social cognition. In C. Moore & P. J. Dunham (Eds.), *Joint attention: Its origins and role in development* (pp. 103–130). Hillsdale, NJ: Erlbaum.

Tomasello, M. (1999). *The cultural origins of human cognition*. Cambridge, MA: Harvard University Press.

Tomasello, M. (2001). Do young children have adult syntactic competence? *Cognition*, *74*, 209–253.

Tomasello, M., & Haberl, K. (2003). Understanding attention: 12- and 18-month-olds know what's new for other persons. *Developmental Psychology*, *39*, 906–912.

Trabasso, T., Stein, N. L., Rodkin, P. C., Munger, M. P., & Baugh, C. R. (1992). Knowledge of goals and plans in the on-line narration of events. *Cognitive Development*, *7*, 133–170.

Wellman, H. M. (1992). *The child's theory of mind*. Cambridge, MA: MIT Press.

Wellman, H. M., Cross, D., & Watson, J. (2001). Meta-analysis of theory of mind development: The truth about false belief. *Child Development*, *3*, 655–684.

Wellman, H. M., & Phillips, A. T. (2001). Developing intentional understandings. In B. Malle, L. Moses, & D. Baldwin (Eds.), *Intentionality: A key to human understanding* (pp. 125–148). MIT Press: Cambridge, MA.

Wellman, H. M., Phillips, A. T., Dunphy-Lelii, S., & Lalonde, N. (2004). Infant understanding of persons predicts preschool social cognition. *Developmental Science*, *7*, 283–288.

Wenner, J. A., & Bauer, P. J. (1999). Bringing order to the arbitrary: One- to two-year olds' recall of event sequences. *Infant Behavior and Development*, *22*, 585–590.

Whiten, A. (1994). Grades of mind reading. In C. Lewis & P. Mitchell (Eds.), *Children's early understanding of mind* (pp. 47–70). Hillsdale, NJ: Erlbaum.

Wilson, M. (2001). The case for sensorimotor coding in working memory. *Psychonomic Bulletin and Review*, *8*, 44–57.

Wilson, C., & Woodward, A. L. (2003). What infants know about intentional action and how they come to know it. *Behavioral and Brain Sciences*, *27*, 129.

Woodward, A. L. (1998). Infants selectively encode the goal object of an actor's reach. *Cognition*, *69*, 1–34.

Woodward, A. L. (1999). Infants' ability to distinguish between purposeful and non-purposeful behaviors. *Infant Behavior and Development*, *22*, 145–160.

Woodward, A. L. (2003). Infants' developing understanding of the link between looker and object. *Developmental Science*, *6*, 297–311.

Woodward, A. L. (2004a). Infants' use of action knowledge to get a grasp on words. In S. R. Waxman (Ed.), *Weaving a lexicon*. Cambridge, MA: MIT Press.

Woodward, A. L. (2004b). *Learning about action by acting*. Meetings of the International Society for Infant Studies, Chicago, IL.

Woodward, A. L. (in press). Infants' understanding of the actions involved in joint attention. In J. Roessler, (Ed.), *Joint attention: Communication and other minds*. Oxford: Oxford University Press.

Woodward, A. L., & Guajardo, J. J. (2002). Infants' understanding of the point gesture as an object-directed action. *Cognitive Development*, *17*, 1061–1084.

Woodward, A. L., & Sommerville, J. A. (2000). Twelve-month-old infants interpret action in context. *Psychological Science*, *11*, 73–77.

Woodward, A. L., Sommerville, J. A., & Guajardo, J. J. (2001). How infants make sense of intentional action. In B. Malle, L. Moses, & D. Baldwin (Eds.), *Intentions and intentionality: Foundations of social cognition* (pp. 149–169). Cambridge, MA: MIT Press.

Zacks, J. M., & Tversky, B. (2001). Event structure in perception and conception. *Psychological Bulletin*, *127*, 3–21.

Zacks, J., Tversky, B., & Iyer, G. (2001). Perceiving, remembering and communicating structure in events. *Journal of Experimental Psychology: General*, *130*, 29–58.

ANALYZING COMORBIDITY

Bruce F. Pennington
DEPARTMENT OF PSYCHOLOGY, UNIVERSITY OF DENVER, DENVER, CO 80208, USA

Erik Willcutt and Soo Hyun Rhee
DEPARTMENT OF PSYCHOLOGY, UNIVERSITY OF COLORADO, BOULDER, CO 80309, USA

I. Introduction

A. DEFINITION AND IMPORTANCE OF COMORBIDITY

A fundamental question for psychology is how does atypical development relate to typical development. An adequate theory of development will need to account for both human universals *and* individual differences, hopefully with the same underlying mechanisms. So every example of atypical development

Advances in Child Development and Behavior
R Kail (Editor)

poses both a challenge and an opportunity for developmental theory. In this chapter, we focus on a pervasive characteristic of atypical development, comorbidity among behavioral disorders. Comorbidity is relevant for developmental theory because it can provide insights into how behavioral disorders develop. Each disorder can be conceived of as having a particular developmental trajectory, in which some mechanisms underlying normal development are perturbed and many others are intact. The phenomenon of comorbidity often means that the developmental trajectories for different disorders intersect. These intersection points can be particularly informative about underlying mechanisms. We hope to demonstrate how analyses of comorbidity may shed some light on these intersection points and hence the mechanisms that underlie both typical and atypical development.

Comorbidity simply means the co-occurrence in a single patient of two or more diagnoses (Feinstein, 1970). Because both the scientific and clinical value of a diagnostic construct depends in part on it providing a unifying explanation of the diverse signs and symptoms presented by a patient, unexplained comorbidity is a phenomenon that potentially poses problems for the explanatory value of diagnostic constructs. Perhaps the two comorbid disorders are simply different manifestations of the same underlying disease process, in which case only one diagnostic construct is needed. These different manifestations could be present at the same time or one could precede the other. For instance, only when the infectious agent responsible for syphilis (the spirochete bacteria) was discovered, could it be appreciated that the very different signs and symptoms of the three stages of this disease were all part of the same disorder. So progress in understanding the etiological and pathogenetic mechanisms that underlie syndromal collections of signs and symptoms can explain and ultimately eliminate comorbidities, because diagnostic boundaries are redrawn. It follows that comorbidity is more likely when diagnostic constructs are more descriptive than explanatory and less is known about underlying mechanisms, which is clearly the case for psychiatric diagnoses.

In fact, unlike much of the rest of medicine, current psychiatric nosologies, such as the DSM-IV or the ICD-10, contain diagnostic constructs that are *intentionally* just descriptive, because we do not yet know enough about underlying mechanisms to use them to define psychiatric disorders. Because earlier psychiatric nosologies were based on unproven assumptions about underlying mechanisms and because diagnostic definitions were not specific enough to be reliable across diagnosticians, developing reliable descriptive diagnostic categories was a scientific step *forward* for psychopathology research. But defining psychopathologies in terms of underlying mechanisms instead of just symptoms remains a key long-term goal of research on psychiatric disorders. As we demonstrate in this chapter, analyzing comorbidity is one strategy for reaching that goal.

More specifically there are four main reasons why comorbidity is important for both research and clinical practice (Caron & Rutter, 1991; Klein & Riso, 1993; Klein, 1993, 2003). First, the presence of a comorbid disorder (e.g., anxiety) may influence the course and treatment of another disorder (e.g., depression). Second, if comorbidity is ignored, one may falsely conclude that some variable is associated with a given disorder (e.g., conduct problems in reading disability (RD)), when in fact the association is due to a comorbid condition (e.g., attention deficit hyperactivity disorder (ADHD)). Third, as discussed previously, comorbidity is a threat to the validity of diagnostic constructs. Finally, as we demonstrate in this chapter, analyses of comorbidity can also be a very useful "wedge" for prying apart underlying mechanisms, which in turn will allow us to develop more valid diagnostic constructs.

B. OVERVIEW OF CHAPTER

In what follows, we first document that comorbidity is a pervasive phenomenon in both adult and child psychiatry, then consider both artifactual and non-artifactual explanations of comorbidity, discuss genetic and cognitive methods for testing these explanations, provide examples of the application of these methods to specific, common comorbidities found in child psychopathology, explain the complex disease model and how it accounts for comorbidity, and consider implications for future research.

Extensive research documents the pervasiveness of comorbidity of psychiatric disorders, both in adults (see review by Clark, Watson, and Reynolds (1995) and children (see reviews by Caron and Rutter (1991) and by Angold, Costello, and Erkanli (1999)). For adults, more than half of individuals with one DSM diagnosis had at least one additional comorbid diagnosis in two different large national epidemiological studies: 60% in the Epidemiological Catchment Area (ECA) study (Robins, Locke, & Regier, 1991), and 56% in the National Comorbidity Survey (NCS) study (Kessler *et al.*, 1994). Moreover, those with comorbid disorders account for a large proportion of all diagnoses: 79% of all lifetime diagnoses and 82% of all 12-month diagnoses in the NCS survey. Finally, although pervasive, comorbidity is not random; some pairs of disorders co-occur much more frequently than others.

For children, somewhat similar results have been found. In an epidemiological study in Puerto Rico (Bird *et al.*, 1988), almost half of children with one diagnosis had a second diagnosis. Several other studies have documented high rates of comorbidity among childhood psychiatric disorders (e.g., Steingard *et al.*, 1992; Jensen *et al.*, 1995). Again, some pairs of disorders, such as ADHD and Conduct Disorder, CD, co-occur much more frequently than other pairs, such as CD and Anxiety disorder (Angold, Costello, & Erkanli, 1999).

II. Explanations of Comorbidity

In this section, we consider possible explanations for the phenomenon of comorbidity, dividing these into artifactual and non-artifactual explanations. Before doing that, it is useful to consider some other factors that influence the phenomenon of comorbidity. Klein and Riso (1993) made the important point that the concept of comorbidity presupposes that disorders are discrete categories, an assumption that has been hotly debated in the case of psychiatric disorders. Lilienfeld, Waldman, and Israel (1994) have argued that the term comorbidity is misleading when applied to psychiatric disorders because we do not know if they are discrete clinical entities. But even if psychiatric disorders are dimensional rather than discrete, covariation among the defining dimensions is an important phenomenon to understand. As we will see, some of the methods for analyzing comorbidity presented here do not require the assumption that psychiatric disorders are discrete categories.

Another factor that influences the phenomenon of comorbidity is whether the nosology employs hierarchical exclusion rules (Clark, Watson, & Reynolds, 1995). For instance, in DSM-IV, a diagnosis of ADHD is precluded by a diagnosis of mental retardation or autism because the latter diagnoses are more severe and pervasive. But this exclusion assumes that we know that primary autism causes comorbid ADHD, which may not be correct. In DSM-III, there were many more such hierarchical exclusion rules than in DSM-IV and consequently the rates of comorbidity observed using DSM-III are lower than when using DSM-IV. Because many of the hierarchical exclusionary rules in DSM-III lacked a theoretical or empirical rationale, they were dropped in DSM-IV. This change in DSM criteria illustrates a broader point: it is hard for diagnostic criteria to remain purely descriptive. Implicit theories of disorders inevitably creep into their definition and relations with each other.

A. ARTIFACTUAL EXPLANATIONS

Although the phenomenon of comorbidity may signal an unappreciated causal relation between two disorders, it may also simply be an artifact of some kind. So before undertaking more extensive research to discover this possible causal relation, researchers must rule out artifactual explanations. Both Caron and Rutter (1991) and Klein and Riso (1993) discuss several artifactual explanations of comorbidity: chance, sampling bias, population stratification, definitional overlap, and rater biases. We next briefly discuss each of these artifactual explanations and how they may be tested.

First of all, two disorders may co-occur simply by *chance*. The rate of such comorbidity is simply the product of their prevalences in the population. So, for two disorders A and B, each with a prevalence of 10%, the chance rate of comorbidity in the population will be 1%. If the rate of comorbidity in a population sample is significantly greater than 1%, then we can reject chance as the explanation. Both Caron and Rutter (1991) and Angold, Costello, and Erklani (1999) found that rates of comorbidity observed among several childhood disorders in epidemiological samples were significantly greater than what would be predicted by chance. Angold, Costello, and Erkanli (1999) performed a meta-analysis of epidemiological studies in which they computed the median odds ratio and 95% confidence intervals for pairs of disorders to test whether the rates of different comorbidities differed from chance and from each other. For instance, they found the median odds ratio for the comorbidity between ADHD and CD was 10.7 (confidence interval = 7.7–14.8), whereas that for Anxiety and CD was 3.1 (confidence interval = 2.2–4.6). These results tell us each comorbidity is greater than chance would predict and that the rate of comorbidity between CD and ADHD is greater than that between CD and Anxiety.

Second, apparent comorbidity might be due to *sampling bias*, the best known example of which is Berkson's bias. Berkson (1946) showed that apparent comorbidities between otherwise independent disorders will arise in referred samples if the probability for referral of either or both disorders is less than one. In this case, comorbid individuals will be over-represented because their probability of being referred is a combined function of the referral rates of each of their disorders. A simpler way of putting this is that people with more problems are more likely to seek help. Therefore, referred or clinic samples will not provide reliable estimates of comorbidity, unless we know the population rates of each disorder and the referral rates and biases affecting the clinic in question. If known, these parameters could be used to estimate the actual comorbidity from the observed comorbidity, but usually not all these parameters are known.

Third, apparent comorbidity might be due to *population stratification*. Although the risk factors for each disorder A and B are in fact independent in the population as a whole, they may co-occur in certain strata of the population. If our test of comorbidity is performed on that strata, then we will falsely conclude that comorbidity between A and B happens more frequently than by chance. For instance, non-random or assortative mating by individuals with disorders A and B will produce apparent comorbidity in their offspring, if both disorders are familial. Such non-random mating appears to explain the comorbidity between depression and alcoholism (Merikangas, 1982). As a second example of population stratification, Caron and Rutter (1991) discuss the apparent comorbidity between depression and conduct disorder in children. In this case, depression in parents acts in different ways to create an increased risk for each disorder in offspring: parental depression is both a genetic and environmental risk factor for depression in a child and parental depression increases marital discord, which is an environmental risk factor for conduct disorder in the child.

Fourth, *definitional overlap* could produce artifactual comorbidity. If some of the defining symptoms for disorders A and B are the same, then individuals with those overlapping symptoms will be more likely to be comorbid than if the defining symptoms did not overlap. If one still finds comorbidity after deleting the overlapping symptoms, then one can reject this artifactual explanation. But the converse is not necessarily the case because deleting symptoms changes the definition of each disorder.

Fifth, *rater biases* or halo effects can produce artifactual comorbidity. If the ratings for both disorders A and B are provided by the same informant, such as a teacher or parent, then the rater's concern about the child's true disorder A may lead them to endorse more symptoms of disorder B, thus producing comorbidity.

B. NON-ARTIFACTUAL EXPLANATIONS

Neale and Kendler (1995) presented the quantitative specifications of Klein and Riso's models, describing 13 comorbidity models providing the most comprehensive set of possible explanations for comorbidity. The first of these models is the Chance model described previously, and the other 12 models are non-artifactual explanations for comorbidity (see Figure 1).

In Figure 1, the latent variable "*R*" refers to the multifactorial liability for each disorder (e.g., R_A = multifactorial liability for disorder A, R_B = multifactorial liability for disorder B). The liability distributions with the thresholds in the boxes are simply another way of representing the multifactorial liability for each disorder. (Note that the path coefficient from the latent variable "*R*" and the liability distributions with the thresholds is always 1.) The individuals who cross the threshold in the liability distribution manifest disorder A or B.

All of the Neale and Kendler comorbidity models are versions of the continuous liability threshold model, which assumes that there is a continuous liability distribution of multifactorial causes (genetic and/or environmental causes) for a disorder, and that a disorder occurs if an individual crosses a particular threshold in that liability distribution. The 12 non-artifactual explanations for comorbidity can be divided into four groups of related models: alternate forms, multiformity (six models), three independent disorders, and correlated liabilities (four models).

The alternate forms model hypothesizes that comorbidity occurs because the two comorbid disorders are alternate manifestations of a single liability. For individuals who cross a particular threshold in that single liability distribution, the probability of having disorder A is p, and the probability of having disorder B is r. This means that both disorders share a single liability, and that one person manifests disorder A while another person manifests disorder B because of chance or risk factors that vary across individuals. A gene by environment interaction where the environmental risk factor is specific to an individual is

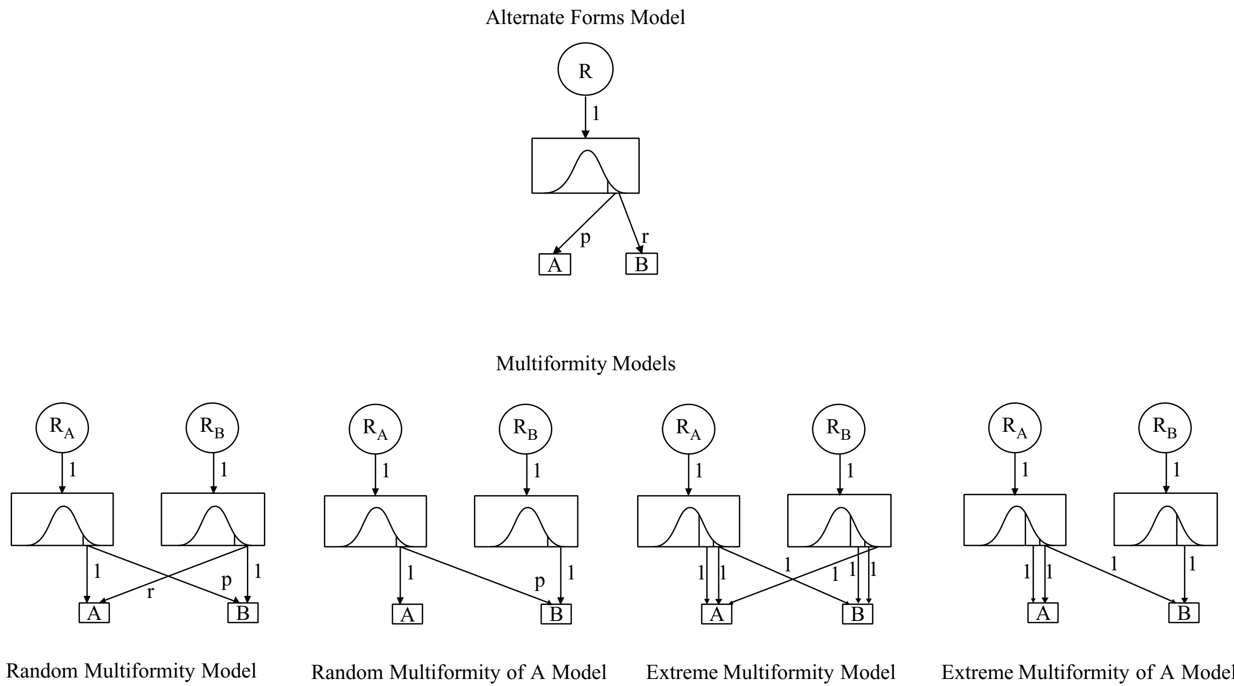

Fig. 1. Neale and Kendler models. R is a latent variable that refers to the multifactorial liability for each disorder. The latent Rs for disorder A or B are independent unless connected by a line or lines. The liability distribution is a normal curve with a threshold (or thresholds) on the extreme high end. Those falling above the threshold have the disorder A or B or both, and are represented by the boxes labeled A or B.

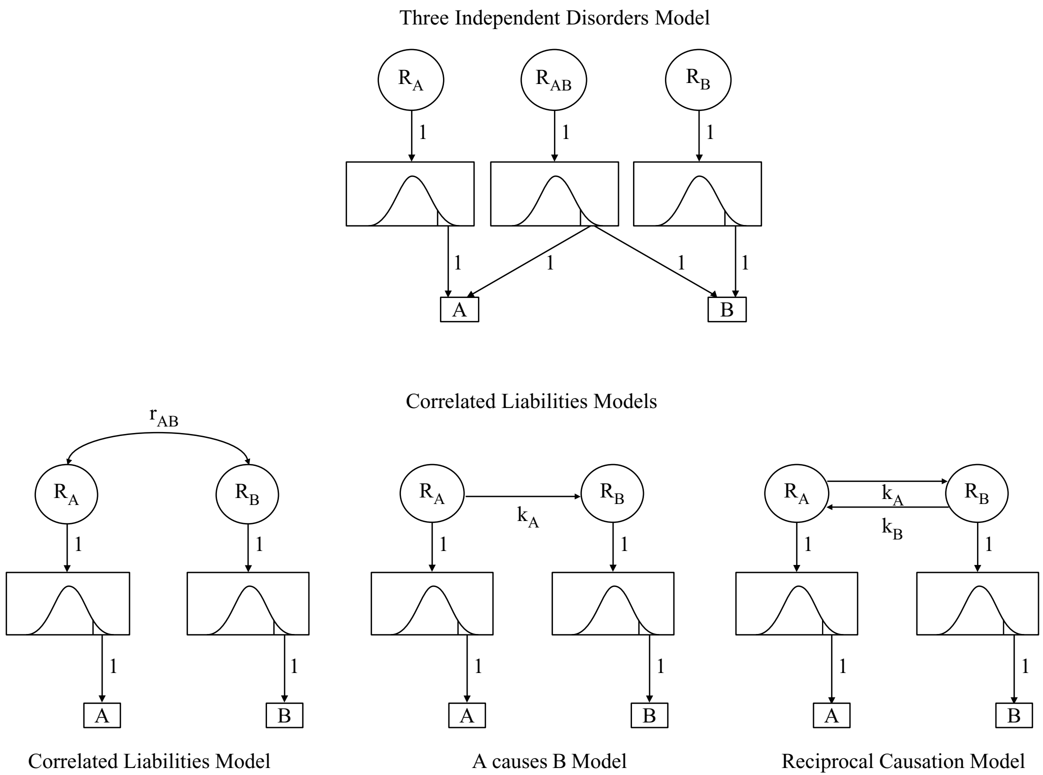

Fig. 1. Continued.

an example of how comorbidity may occur through alternate forms. If two individuals have the same overall liability but are exposed to different person-specific environmental risks, the first individual may manifest disorder A while the second individual manifests disorder B. A gene by environment interaction that replicates across individuals would not be an example of the alternate forms model.

In multiformity models, an individual who has one disorder is at an increased risk for having the second disorder, despite not having an elevated liability for the second disorder. Although having a disorder without having the liability for it may seem contradictory, it is important to remember that behavioral disorders are defined at the level of symptoms, whereas liability is defined in terms of underlying etiological mechanisms. So, sometimes the symptoms of a disorder can be produced without the usual underlying etiological mechanisms being present. Two of the multiformity models, the random multiformity of A and random multiformity of B models illustrate this possibility. They are the "phenocopy" model often discussed in the literature. The "phenocopy" model hypothesizes that the first disorder produces a copy of the second disorder; hence, the first disorder is primary whereas the second disorder is secondary. For example, Pennington, Groisser, and Welsh (1993) suggested that RD might lead to the phenotypic manifestation of ADHD in the absence of etiological influences typically associated with ADHD in isolation. One can readily imagine that a child might appear to be inattentive or hyperactive in the classroom due to the frustration elicited by difficulties with reading, rather than as a consequence of the neurocognitive difficulties that are typically associated with ADHD in the absence of RD.

According to the three independent disorders model, comorbidity occurs because the comorbid disorder is a disorder that is separate from either disorder occurring alone. It is sometimes referred to as the "subtype" hypothesis in the literature.

The four correlated liabilities models share the idea that there is a continuous relation between the liability to one disorder and the liability to the second disorder. An increase in liability for one disorder is correlated with the increase in liability for the second disorder. (In contrast, in the multiformity models, a change in liability for one disorder has absolutely no effect on the second disorder unless an individual crosses the threshold for the first disorder and is actually affected by the disorder.) The relation between the liability of the two disorders occurs via a significant correlation between the risk factors (correlated liabilities) or a direct causal relation between the manifest phenotypes of the two disorders (A causes B, B causes A, or reciprocal causation).

Although the Neale and Kendler (1995) models are a major contribution to the comorbidity literature because they are the most complete set of models yet proposed and because they are specified quantitatively, they nonetheless have some limitations. Specifically, they do not include a neural or a cognitive level, they are not explicitly developmental, they only deal with pairwise comorbidities, and some of them can be difficult to distinguish empirically, even with ideal (simulated) data sets. Most of these limitations are apparent in the sections that follow.

III. Methods for Analyzing Comorbidity

Given the importance of discriminating the correct comorbidity model among many alternatives, a series of studies (Rhee *et al.*, 2003, 2004a,b,c) examined whether various methods testing alternative comorbidity models are valid. In all studies, simulations were conducted to test the validity of the common methods used to test alternative comorbidity models.

Data were simulated for each of the 13 Neale and Kendler comorbidity models; in these simulated data, the true cause of comorbidity is known. Then, analyses commonly used to test the alternative comorbidity models were conducted on each of the 13 simulated datasets. If a particular analysis is valid, the predicted result should be found in the data simulated for the comorbidity model, and the predicted result should not be found in data simulated for other comorbidity models (i.e., the particular analysis should discriminate a particular comorbidity model from alternative hypotheses).

A. KLEIN AND RISO'S FAMILY PREVALENCE ANALYSES

For each comorbidity model, Klein and Riso (1993) presented a set of predictions regarding the prevalence of disorders in the relatives of different groups of probands. They presented a comprehensive set of predictions comparing the prevalence of disorder A-only, disorder B-only, and disorder AB (i.e., both disorders) among the relatives of probands with A-only, B-only, AB, and controls. Several studies have used these predictions to test alternative comorbidity models (e.g., Wickramaratne & Weissman, 1993; Riso *et al.*, 1996; Donaldson *et al.*, 1997).

Most of Klein and Riso's predictions were validated by the simulation results, in that most of their predicted results matched the results in the simulated datasets. However, there were several notable differences between the predicted results and results obtained in the simulated datasets. Some of Klein and Riso's predictions were not obtained in the simulated results because of lack of power in the simulated datasets. Another reason for the discrepancy between the predicted results and the results in the simulated dataset was the predictions' lack of consideration of all possible pathways for the comorbid disorder, notably the fact that there will be some individuals who have both disorders A and B due to chance.

B. FAMILY PREVALENCE ANALYSES IN THE LITERATURE

Many other researchers (e.g., Biederman *et al.*, 1992; Bierut *et al.*, 1998) have conducted a subset of the Klein and Riso analyses or analyses very similar to those presented by Klein and Riso (1993) without testing their comprehensive set of predictions. Most of the studies in the literature have focused on three comorbidity models: (a) the alternate forms model (i.e., the two comorbid disorders are alternate manifestations of the same underlying liability); (b) the correlated liabilities model (i.e., there is a significant correlation between the liabilities for the two models); and (c) the three independent disorders model (i.e., the comorbid disorder is a third disorder that is etiologically distinct from either disorder occurring alone).

The results of the study (Rhee *et al.*, 2003) examining the validity of family prevalence analyses found in the literature indicate that although some analyses validly discriminate the alternate forms model from other comorbidity models, the analyses testing the correlated liabilities model and the three independent disorders model did not discriminate them from other comorbidity models. In many cases, although the predicted results were consistent with a particular comorbidity model, they were also consistent with several alternative comorbidity models.

C. NEALE AND KENDLER MODEL FITTING APPROACH

Neale and Kendler (1995) described 13 alternative models. They illustrated the probabilities for the four combinations of disease state (neither A nor B; A but not B; B but not A; both A and B) for each comorbidity model, then illustrated the probabilities for the ten combinations of affected or unaffected status for pairs of relatives for each comorbidity model (e.g., neither A nor B in relative 1 and neither A nor B in relative 2; both A and B in relative 1, and A only in relative 2). The data that are analyzed in the Neale and Kendler model fitting approach are simply the frequency tables for the number of relative pairs in each possible combination of disease state. The observed cell frequencies are compared to the expected cell frequencies (i.e., the probabilities for the ten combinations of affected or unaffected status for pairs of relatives) in each comorbidity model. The comorbidity model with the smallest difference between the observed cell frequencies and the expected cell frequencies is chosen as the best fitting model.

In general, the Neale and Kendler model fitting approach discriminated the following classes of models reliably: the alternate forms model, the random multiformity models (i.e., random multiformity, random multiformity of A, and random multiformity of B), the extreme multiformity models (i.e., extreme multiformity, extreme multiformity of A, and extreme multiformity of B), the three independent disorders model, and the correlated liabilities models (i.e., correlated liabilities, A causes B, B causes A, and the reciprocal causation). Discrimination within these classes of models was poorer. Results from simulations varying the prevalences of the comorbid disorders indicate that the ability to discriminate between models becomes poorer as the prevalences of the disorders decreases, and suggests the importance of considering the issue of power when conducting these analyses.

D. UNDERLYING DEFICITS APPROACH

Several researchers have tested alternative comorbidities by comparing the underlying neuropsychological deficits of the two comorbid disorders in individuals with neither disorder, A only, B only, and both A and B. So, unlike the family prevalence approaches just discussed, this method examines groups of unrelated individuals. For example, Pennington, Groisser, and Welsh (1993) examined the comorbidity between reading disability and ADHD, comparing the underlying deficits associated with reading disability (i.e., phonological processes) and the underlying deficits associated with ADHD (i.e., executive functioning) in individuals with neither disorder, reading disability only, ADHD only, and both reading disability and ADHD. Most of the researchers using this approach have made predictions for 5 of the 13 Neale and Kendler comorbidity models. In addition to the three models often tested using family prevalence analyses in the literature (i.e., alternate forms, correlated liabilities, and three independent disorders), researchers have made predictions regarding the random multiformity of A or random multiformity of B models (i.e., an individual who has one disorder is at an increased risk for having the second disorder, although he or she does not have an elevated liability for the second disorder).

Given adequate power, the method of examining the underlying deficits of comorbid disorders can distinguish between all 13 Neale and Kendler comorbidity models, except the random multiformity, extreme multiformity, and three independent disorders models. As the sample sizes decreased and the magnitude of correlation between the underlying deficits and the symptom scores decreased, the ability to discriminate the correct comorbidity model from alternative hypotheses decreased. Again, the issue of power should be considered carefully.

E. CONCLUSIONS

Although most of Klein and Riso's family prevalence analyses were valid, there were notable discrepancies between their predicted results and results found in the simulated datasets. Some of the family prevalence analyses found in the literature were valid predictors of the alternate forms model, but none were valid predictors of the correlated liabilities or three independent disorders models. The Neale and Kendler model fitting approach and the method of examining the underlying deficits of comorbid disorders discriminated between several comorbidity models reliably, suggesting that these two methods may be the most useful methods found in the literature. Especially encouraging is the fact that some of the models that cannot be distinguished well using the Neale and Kendler model fitting approach can be distinguished well by examining the underlying deficits of comorbid disorders, and vice versa. The best approach may be a combination of these two methods. However, simulation results suggest that the issue of power should be considered carefully.

IV. Analyses of Specific Comorbidities

In this section, we present analyses of three specific comorbidities commonly found among childhood disorders: between speech sound disorder (SSD) and reading disability, between RD and ADHD, and between ADHD and conduct disorder (CD). These examples were chosen because each individual disorder has a high prevalence among children and each comorbidity has received enough empirical attention to make a review worthwhile. Not incidentally, these were also comorbidities that we have studied.

Of these three comorbidities, perhaps the most surprising one is that between RD and ADHD because we think of these disorders as being cognitively distinct. RD is usually conceptualized as a kind of language disorder, involving a problem in phonological development and ADHD is usually conceptualized as a kind of executive disorder, involving a problem in the development of inhibitory control (Pennington, 2002). Hence, the comorbidity between RD and ADHD qualifies as an example of what Angold, Costello, and Erkanli (1999) call “heterotypic comorbidity” in which the comorbidity is between disorders from different diagnostic groupings. In contrast, homotypic comorbidity is between disorders from the same diagnostic grouping. The frequently studied comorbidities among anxiety disorders or between depression and dysthymia are good examples of homotypic comorbidity. The other two comorbidities that we present here, between RD and SSD, and between ADHD and CD can be considered examples of homotypic comorbidities because RD and SSD are both language disorders and ADHD and CD are both externalizing disorders. Of course, as Angold, Costello, and Erkanli (1999) point out, the distinction between homotypic and heterotypic comorbidity is not completely clear-cut because it presupposes that we already have an adequate scientific understanding of these disorders.

It is also interesting to note that research on each of these three comorbidities has rejected an initially favored and seemingly intuitive explanation. For the comorbidity between SSD and RD, this favored hypothesis (severity) was an example of the alternate forms model, in which the two disorders share the same liability distribution but are different phases or expressions of that liability. On this hypothesis, the etiology shared by SSD and RD disrupts phonological development, which then manifests as speech problems in the preschool years and as reading problems in the school years. For RD and ADHD, the initially favored hypothesis for this counterintuitive comorbidity was that it was either an artifact or a phenocopy (multiformity of RD). In either of these cases, the underlying liabilities for pure RD and ADHD would be distinct. For ADHD and CD, one favored hypothesis was the three independent disorders model, in which the combination of ADHD and CD was a distinct disorder from either ADHD or CD alone. This hypothesis was favored because there were distinct correlates of comorbid ADHD + CD compared to each disorder in isolation. So research on each of these three comorbidities has produced some counterintuitive results, which have led us to new models of how disorders develop and why they are comorbid.

A. READING DISABILITY AND SPEECH SOUND DISORDER

From some perspectives, it does not make sense that SSD and RD should be comorbid. SSD (Shriberg, Tomblin, & McSweeny, 1999) involves difficulties in the preschool development of *spoken* language, specifically problems with the accurate (and therefore intelligible) production of speech sounds in spoken words (it is distinct from stuttering or mutism). RD, or dyslexia, manifests at school age with difficulty in learning *written* language, specifically printed word recognition and spelling (see IDA and NICHD working definition of dyslexia, Dickman, 2003). In the past, RD has been conceptualized as a visual disorder (Orton, 1925) and SSD has been conceptualized as an auditory or motor disorder. So, from these perspectives, each disorder would appear to require a different neurobiological origin. But as discussed earlier, if each disorder is viewed as a kind of language disorder, then their comorbidity is less surprising and could be called homotypic. However, as we will see, we have had to reject our initial favored hypothesis for their comorbidity, the severity variant of the alternate forms hypothesis.

1. Symptom Overlap Between SSD and RD

Children with early speech/language problems are at increased risk for later literacy problems (Hall & Tomblin, 1978; Snowling & Stackhouse, 1983; Aram, Ekelman, & Nation, 1984; Bishop & Adams, 1990; Magnusson & Naucler, 1990; Scarborough & Dobrich, 1990; Rutter & Mawhood, 1991; Tomblin, Freese, & Records, 1992; Snowling, Bishop, & Stothard, 2000; Catts *et al.*, 2002) and individuals with literacy problems retrospectively report increased rates of earlier speech and language problems (Hallgren, 1950; Rutter & Yule, 1975).

Moreover, the latter association is not limited to retrospective reports because young children selected for family risk for dyslexia or RD and followed prospectively also have higher rates of preschool speech and language problems than controls (Scarborough, 1990; Gilger *et al.*, 1994; Gallagher, Frith, & Snowling, 2000; Pennington & Lefly, 2001; Lyytinen *et al.*, 2002). But these previous studies have rarely distinguished SSD from specific language impairment (SLI), which is defined by deficits in semantics and syntax. So, it is less clear which subtypes (or components) of SSD *per se* presage which kinds of later literacy problems.

2. *Cognitive and Etiological Overlap Between SSD and RD*

The large majority of children with problems in printed word recognition (i.e., dyslexia or RD) have deficits on measures of phonological processing (Wagner & Torgesen, 1987), including measures of both explicit (i.e., phoneme awareness) and implicit (i.e., phonological memory and rapid serial naming) phonological processing. There is also accumulating evidence that many children with speech and language problems have phonological processing problems, such as deficits on measures of phoneme awareness and phonological memory (Leonard, 1982; Kamhi *et al.*, 1988; Bird & Bishop, 1992; Lewis & Freebairn, 1992; Bird, Bishop, & Freeman, 1995; Bishop, North, & Donlan, 1995; Clarke-Klein & Hodson, 1995; Montgomery, 1995; Edwards & Lahey, 1998).

Support for a shared etiology for SSD and RD has been provided by Lewis and colleagues (Lewis, Ekelman, & Aram, 1989; Lewis, 1990, 1992), who found that SSD and RD are co-familial. We have found that SSD and RD are coheritable as well (Tunick & Pennington, 2002).

The etiological and cognitive overlap between SSD and RD suggests a parsimonious *severity* hypothesis, namely that many cases of SSD and RD lie on a severity continuum in which shared etiological risk factors lead to a shared underlying phonological deficit. If the phonological deficit is severe enough, it first produces SSD and then later RD. So, according to the severity hypothesis, SSD and RD are alternate forms of the same underlying liability expressed at different points in development. If it is less severe, it does not produce diagnosable SSD (though it may lead to subclinical speech production problems), but it does produce later RD, because reading requires more mature phonological representations than does speech. So this hypothesis posits that RD without earlier SSD is a less severe variant of SSD and has a less extreme threshold on the same liability distribution. To account for children with SSD who do not develop later RD, the severity hypothesis must posit that they have a subtype of SSD that is not caused by an underlying phonological deficit. The already documented etiological and cognitive overlap between SSD and RD supports this severity hypothesis. But because SSD has not been clearly distinguished from SLI in previous etiological and cognitive studies, other possible hypotheses can explain the relation between SSD and RD.

3. Hypotheses to Explain SSD/RD Comorbidity

In an NIH grant application, one of us (BFP) once proposed five competing hypotheses (all but one of which were single cognitive deficit hypotheses) to account for the comorbidity of SSD and RD (see Figure 2). These hypotheses were generated without knowledge of the Klein and Risso (1993) hypotheses, yet all but one of them (cognitive phenocopy) corresponds to one of their hypotheses. These hypotheses were generated by crossing two distinctions: a common vs. distinct etiology and a common vs. distinct cognitive phenotype. These five hypotheses were: (1) severity (both etiology and cognitive phenotype are shared, but comorbid children have a more severe phonological deficit); (2) pleiotropy (a shared etiology leads to two distinct cognitive phenotypes, which co-occur in comorbid children); (3) cognitive phenocopy or genetic heterogeneity (distinct etiologies lead to a shared cognitive phenotype, thus producing comorbidity); (4) cross-assortment or non-random mating (both the etiology and cognitive phenotypes are distinct, but individuals with SSD (or RD) are more likely to select mates with RD (or SSD), thus transmitting risk alleles for both disorders to their children); and (5) synergy, in which the etiologies and cognitive phenotypes of SSD and RD are distinct, but comorbidity between SSD and *SLI* produces later RD. The severity and pleiotropy hypotheses correspond to different versions of Klein and Riso's (1993) alternate forms hypothesis; synergy is similar to the three independent disorders hypothesis; and assortment is an example of the population stratification hypothesis, as discussed earlier.

4. Tests of the Five Hypotheses

To distinguish these five hypotheses, three questions need to be addressed. (1) Do SSD and RD share a common genetic etiology? (2) Do they share an underlying cognitive phenotype? (3) Is there assortative mating between individuals with SSD and those with RD? In what follows, we present what is known about the answers to these questions.

First, there is now stronger evidence for a shared genetic etiology between RD and SSD, which rejects hypotheses 3–5, all of which posit distinct etiologies for RD and SSD. Two groups have now tested whether some of the risk loci already identified for RD are risk loci for SSD. Several replicated risk loci or QTLs for RD have been identified, on chromosomes 1p, 2p, 3p–q, 6p, 15q, and 18p (Fisher & DeFries, 2002). Stein and colleagues found that SSD is linked to the RD locus on chromosome 3 (Stein *et al*., 2004). They tested several related phenotypes, including SSD itself, phonological memory, phonological awareness, and reading. All of these phenotypes were linked to the RD risk locus on chromosome 3, indicating that this locus affects phonological development and

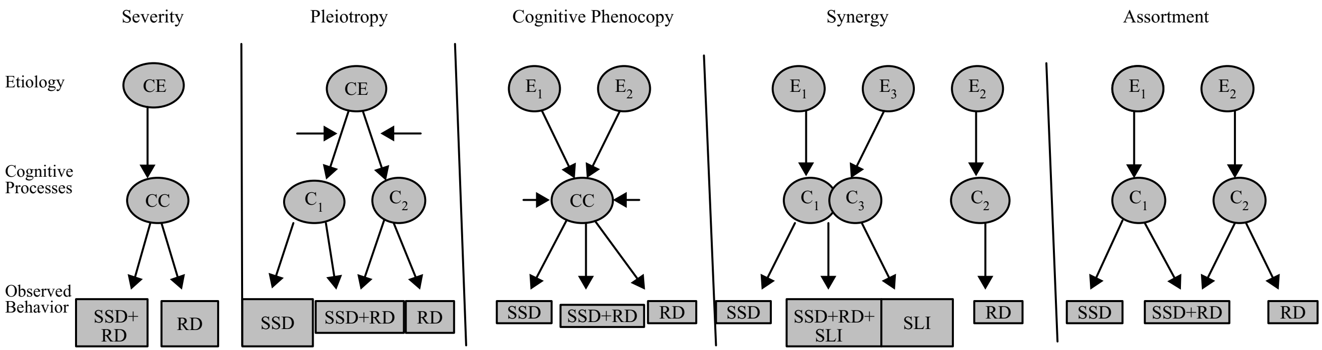

Fig. 2. Schematic representation of five hypotheses. CE = Common Etiology; CC = Common Cognitive Phenotype; E_1*,*E_2*,*E_3 *= Specific etiologies;* C_1*,*C_2*,*C_3 *= Specific cognitive phenotypes. Horizontal arrows indicate influence of additional chance factors.*

contributes to the comorbidity between SSD and RD. We have also found that SSD is linked to RD risk loci on chromosomes 6 and 15, and perhaps 1 (Smith *et al.*, 2003). We also tested multiple phenotypes, including SSD itself, phonological memory, and phonological awareness, all of which provided evidence for linkage.

The second test of these five hypotheses is whether RD and SSD share an underlying cognitive deficit. To perform this test, we examined preliteracy skills, including phoneme awareness, in a large sample of preschool children with SSD (Raitano *et al.*, 2004). Because we were also interested in whether the cognitive deficit in SSD varied by subtype, we divided the sample along two dimensions, presence vs. absence of SLI, and a persistent speech disorder vs. a speech disorder that has now normalized. We found that a phoneme awareness deficit was pervasive across the four resulting subtypes of SSD, although its severity varied in an additive fashion as a function of each subtype dimension. Those with SLI had a worse phonological awareness deficit than those without SLI; those with a persistent speech disorder had a worse phonological awareness deficit than those whose speech problems had normalized. A similar pattern of results was found for alphabet knowledge. Intriguingly, the SSD group had a less pronounced deficit in rapid serial naming. So the results of this study, along with other evidence reviewed earlier, indicate a shared underlying phonological deficit in SSD and RD, and that this shared deficit is found in all four subtypes of SSD. Thus, the results of this second test reject the phenocopy hypothesis and only partially support the severity hypothesis, which requires a fairly common subtype of SSD *without* a phonological deficit. The fact that the phonological awareness deficit is not restricted to the group with both SSD and SLI is also inconsistent with the predictions of the synergy hypothesis, which is also contradicted by the genetic results just discussed.

To address the third question regarding assortative mating, we examined the parents in our large sample of children with SSD (Tunick *et al.*, in preparation). Relative to control parents, parents of SSD probands reported higher rates of both speech and reading problems, indicating that SSD was familial in this sample and that SSD and RD were co-familial. We also found similar results in the siblings of probands; they had higher rates of speech problems and worse scores on preliteracy measures than controls. These results indicate an etiological overlap between SSD and RD, consistent with the studies discussed earlier. In contrast, we found low rates of cross-assortment in these parents. Moreover, SSD probands with comorbid preliteracy problems rarely came from cross-assorted parents. So we did not find support for the assortment hypothesis.

In sum the results of these three tests reject all but the severity hypothesis. But despite the fact that the severity hypothesis garners some support from these data and that of previous studies reviewed earlier, there are still significant challenges to how it accounts for the nature of the comorbidity between SSD and RD.

The severity hypothesis proposes that SSD and RD are comorbid because they share etiological risk factors (some of which are genetic) and these lead to a shared phonological deficit, which is more *severe* in children with comorbid SSD and RD than children with RD only. To account for SSD children who do not become RD, the severity hypothesis must postulate a subtype of SSD with a distinct etiology and a different underlying cognitive deficit. If SSD children without later RD nonetheless have an underlying phonological deficit, the severity hypothesis must be seriously questioned. But the results of Raitano *et al.* (2004) just discussed suggest there is not a common subtype of SSD without a phonological deficit. Clearer evidence on this point is provided by long-term follow-up study (Snowling, Bishop, & Stothard, 2000) of SSD children initially identified by Bishop at preschool age. These researchers found there were former SSD children with a persistent deficit in phoneme awareness in adolescence who are nonetheless normal readers. Both these results are inconsistent with the severity hypothesis.

Subsequent data from Tunick (2004) also questions the severity hypothesis. Her project involved two comparisons of SSD and RD, one between probands at age 5 and one between siblings of probands around age 8. The goal of the proband comparison was to test which deficits are shared and specific to each disorder before the onset of literacy instruction. The sibling comparison tested the familiality of these patterns and whether they persist to a later age.

Because both SSD and RD vary in the severity of the symptoms that define them diagnostically, it is important to compare SSD and RD groups that are similar in severity. Consequently, Tunick matched the SSD and RD proband groups on severity, as well as on age and gender. The 23 SSD probands were selected from the entire sample of SSD probands in our current study so as to match the 23 RD probands from our earlier longitudinal study of children at high family risk for RD (Pennington & Lefly, 2001). The RD probands were all the children in the high family risk group who were later diagnosed as RD at follow-up. For the sibling comparison, Tunick recruited a separate sample of RD siblings and matched them to a subset of our current sample of SSD siblings on (1) proband sibling's diagnostic severity, (2) their own diagnostic severity, and (3) age and gender.

The comparison of the profiles of phonological processing deficits in probands and siblings tests the severity hypothesis, which predicts similar profiles in each disorder, with greater impairment in the SSD group. We examined three phonological processing constructs: phonological awareness, phonological memory, and rapid serial naming. In the proband comparison, somewhat different measures of the same constructs had been used with each group, so their *z*-scores relative to matched controls were used to compare the SSD and RD proband groups. In the sibling comparison, the same measures were used in each group. We found that SSD and RD probands shared a deficit of similar magnitude (relative to their controls) on the phonological awareness composite, but had significantly different profiles overall, producing a significant group × domain interaction. The interaction arose because the SSD proband group performed significantly better than the RD proband group on the rapid serial naming composite and non-significantly worse score on the phonological memory composite. This interaction replicated in the sibling comparison, in which the *same* measures of these constructs were used in each group. The relative strength on rapid serial naming measures in both SSD groups is a somewhat surprising finding, given that one would expect a slower articulatory rate in SSD. So, it will be important to replicate this result in another SSD sample. But this finding could help explain why not all SSD children develop later RD, despite having a phonological awareness deficit. In sum, Tunick's (2004) results do not support the predictions of the single deficit, severity model because the phonological awareness deficit is not more severe in the SSD groups and because the profiles of phonological deficits are not parallel.

These difficulties with the severity hypothesis led us to develop an alternative *multiple* cognitive deficit model of RD and SSD, which is presented later in this chapter. In this multiple deficit model, comorbidity between these two disorders arises from partially overlapping genetic risk factors (i.e., correlated liabilities) that lead to a shared cognitive deficit (in phonological representations), which interacts with other non-shared cognitive deficits to produce the symptoms that distinguish the two disorders.

The severity and the multiple deficit hypotheses make competing predictions about the literacy outcome of children with SSD. The severity hypothesis predicts (1) that SSD children who do not develop later RD (SSD-only children) have a distinct form of SSD without an underlying phonological deficit, and (2) that SSD children who do develop later RD (comorbid children) have a more severe phonological deficit than both RD children without earlier SSD (RD-only children) and RD children in general (because only about 30% of RD children had earlier SSD). In contrast, the multiple deficit hypothesis predicts (1) that SSD-only children have a phonological deficit but compensate for it via other cognitive protective factors, and (2) that comorbid children will not necessarily have a more severe phonological deficit than RD-only children or RD children in general, but they must have an additional cognitive risk factor to explain why they have RD.

B. READING DISABILITY AND ADHD

RD and ADHD are two of the most common disorders of childhood, each occurring in approximately five percent of the population (e.g., American Psychiatric Association, 2000). ADHD and RD also co-occur significantly more frequently than expected by chance; 25–40% of individuals with ADHD also meet criteria for RD (e.g., Dykman & Ackerman, 1991; Semrud-Clikeman *et al.*, 1992), whereas 15–40% of individuals with RD meet criteria for ADHD (Gilger, Pennington, & DeFries, 1992; Shaywitz, Fletcher, & Shaywitz, 1995; Willcutt & Pennington, 2000a,b).

1. Artifactual Explanations for Comorbidity Between RD and ADHD

Most of the artifactual explanations for comorbidity described previously can be rejected for RD/ADHD comorbidity. RD and ADHD co-occur more frequently than expected by chance in both samples ascertained from clinics (e.g., Semrud-Clikeman *et al.*, 1992) and non-referred samples recruited from the community (e.g., Fergusson & Horwood, 1992; Willcutt & Pennington, 2000a,b; Willcutt *et al.*, in press a,b). Because RD is assessed by cognitive tests whereas ADHD is assessed by behavioral ratings, the relation between RD and ADHD cannot be explained by shared method variance. Similarly, the symptoms of RD and ADHD as defined in the fourth edition of the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-IV; American Psychiatric Association, 1994, 2000) do not overlap.

According to the cross-assortment hypothesis, an individual with RD is more likely to have a child with an individual with ADHD than would be expected by chance based on population base rates of RD and ADHD. In a family study of the biological relatives of children with ADHD, Faraone *et al.* (1993) found that comorbidity between learning disabilities and ADHD was best explained by cross-assortment. However, this result was not replicated in later studies (Doyle *et al.*, 2001; Friedman *et al.*, 2003), suggesting that cross-assortment is not likely to explain the majority of cases of comorbid RD and ADHD.

The rater-bias hypothesis is somewhat more difficult to test, and the possibility remains that parents or teachers may be more likely to endorse ADHD symptoms on a rating scale when they know that the child is experiencing difficulty in learning to read. However, results from our population-based twin study of RD and ADHD indicate that in addition to higher ratings of inattention symptoms by parents and teachers, children with RD report greater attentional difficulties than children without RD on self-report measures (Willcutt, Chhabildas, & Pennington, 1998). Although the rater-bias hypothesis cannot be conclusively rejected based on these results, these data suggest it is unlikely to provide a sufficient explanation for all cases of comorbidity between RD and ADHD.

2. Competing Explanations for True Comorbidity Between RD and ADHD

Three of Neale and Kendler's (1995) competing explanations have received at least some support in previous studies of comorbidity between RD and ADHD. These models include the phenocopy (multiformity of A) hypothesis (e.g., Pennington, Groisser, & Welsh, 1993), the three independent disorders hypothesis (e.g., Rucklidge & Tannock, 2002), and the correlated liabilities (common etiology) hypothesis (e.g., Willcutt, Pennington, & DeFries, 2000, Willcutt *et al.,* in press a,b). We consider each.

a. The Phenocopy Hypothesis. Pennington *et al.* (1993) described results in a small sample of children with RD and ADHD that suggested that RD might lead to the phenotypic manifestation of ADHD in the absence of the etiological influences typically associated with ADHD in isolation. They reached this conclusion because the group with ADHD without RD exhibited a significant deficit on measures of executive functions, whereas the group with ADHD and RD exhibited the phonological processing difficulties that are characteristic of RD, but did not have deficits in executive functioning. Subsequent data from larger samples, however, generally failed to support the phenocopy hypothesis. Instead, these studies suggest that the comorbid group exhibits the additive combination of the neuropsychological weaknesses associated with RD and ADHD when they occur separately (e.g., Nigg *et al.*, 1998; Seidman *et al.*, 2001; Willcutt *et al*., 2001; Rucklidge & Tannock, 2002; Willcutt, in press).

b. The Three Independent Disorders Hypothesis. This model suggests that comorbid RD + ADHD is a third disorder that is due at least in part to etiological factors that are distinct from those that increase susceptibility to RD or ADHD alone. Therefore, this hypothesis predicts that the comorbid group will exhibit a different pattern of neurocognitive deficits or other external correlates than would be predicted based on the additive combination of the deficits associated with each disorder when it occurs alone. Rucklidge and Tannock (2002) found that the comorbid group performed significantly worse than the RD-only and ADHD-only groups on measures of color naming, providing some support for this hypothesis. In contrast, other studies found that the RD + ADHD group exhibited the additive combination of the deficits associated with each individual disorder (e.g., Swanson, Mink, & Bocian, 1999; Pisecco *et al.*, 2001; Willcutt *et al*., 2001), suggesting that additional research is needed.

c. The Correlated Liabilities Hypothesis. Finally, a series of studies tested if the relation between RD and ADHD is attributable to common etiological influences that increase susceptibility to both disorders. Because this model has received the strongest support in previous studies, we describe these results in more detail in the next section (IV.B.3).

3. Behavioral Genetic Studies of RD and ADHD

a. Family Studies. Family studies provide a first step toward understanding the genetic and environmental risk factors for RD, ADHD, and their comorbidity. A family study compares the rate of a disorder in the biological relatives of individuals with and without the disorder. If a disorder occurs more often among the family members of individuals with the disorder, this suggests that familial factors play a role in the etiology of the disorder.

Biological family members of children with RD are 4–8 times more likely to meet criteria for RD than family members of children without RD (Gilger, Pennington, & DeFries, 1991). Similarly, 30–40% of the full siblings of children with ADHD also meet criteria for ADHD, a rate that is 6–8 times higher than the rate in siblings of children without ADHD (Faraone, Biederman, & Friedman, 2000). Thus, RD and ADHD are each clearly familial, and results from our laboratory suggest that the two disorders co-occur in the same families more frequently than expected by chance. Although these results should be interpreted with caution until they can be replicated in an independent sample, they are consistent with the hypothesis that RD and ADHD are attributable to common familial risk factors.

b. Twin Studies. The fact that RD and ADHD are significantly familial suggests that each disorder may be influenced by genes, but family data are not conclusive. Because members of intact biological families share both genetic and family environmental influences, other methods such as twin studies are necessary to disentangle the relative contributions of genes and environment. By comparing the similarity of identical twins, who share all of their genes, to fraternal twins, who share half of their segregating genes on average, twin studies are able to estimate the extent to which a trait is due to genetic or environmental influences. The influence of genes is quantified by estimating *heritability*, a number ranging from 0 (no genetic influences at all) to 1 (entirely due to genetic influences) that provides an index of the extent to which a trait is attributable to genes. Environmental risk factors can be subdivided into *shared* and *non-shared* environmental influence. Shared environmental influences are those that similarly influence members of a family, thereby increasing the similarity of individuals within a family in comparison to unrelated individuals in the populations. In contrast, non-shared environmental influences describe events that affect the two twins differently and lead to differences among individuals in a family.

Twin studies indicate that the heritability of RD is about 0.60, suggesting that genetic influences account for approximately 60% of the reading deficit in children and adolescents with RD (e.g., Wadsworth *et al.*, 2002). ADHD is even more highly heritable (0.75–0.80), indicating that genetic influences play an even larger role in the development of ADHD.

Based on the finding that both RD and ADHD are significantly heritable, several studies have used twin data to test if the same genetic influences contribute to both RD and ADHD. Gilger, Pennington, and DeFries (1992) conducted cross-concordance analyses in a small sample of twins selected for RD, and found that ADHD and RD were primarily attributable to independent genetic factors. However, a statistical trend suggested that children with comorbid RD and ADHD might represent an etiological subtype, providing tentative support for the three independent disorders model. The authors concluded that although most cases of RD or ADHD were not attributable to the same genetic influences, some cases of comorbid RD and ADHD might represent a separate disorder with a genetic etiology distinct from that associated with either diagnosis in isolation.

Stevenson *et al*. (1993) and Light *et al*. (1995) expanded upon the findings of Gilger, Pennington, and DeFries (1992) by conducting more powerful multiple regression analyses to estimate the bivariate heritability of ADHD and reading (Light *et al*., 1995) or spelling difficulties (Stevenson *et al*., 1993). In a sample of twins selected because at least one member of the pair met criteria for RD, Light *et al*. (1995) found significant bivariate heritability for RD and ADHD ($h^2_{g(\mathrm{RD/ADHD})} = 0.45$), suggesting that common genetic influences increase susceptibility to both disorders. In a separate community sample of twins, Stevenson *et al*. (1993) reported that the bivariate heritability of spelling deficits and ADHD was positive and similar whether probands were selected due to spelling difficulties ($h^2_{g(\mathrm{Spell/ADHD})} = 0.21$) or elevations of ADHD symptoms ($h^2_{g(\mathrm{ADHD/Spell})} = 0.15$), but these estimates of bivariate heritability were not statistically significant. Thus, these initial studies provided tentative support for the hypothesis that comorbidity between reading or spelling disability and ADHD may be attributable to common genetic influences, but the findings were somewhat inconclusive.

c. The Importance of ADHD Symptom Dimensions. The etiology of comorbidity between RD and ADHD becomes clearer when symptoms of ADHD are subdivided into dimensions of inattention symptoms and hyperactivity–impulsivity symptoms as described in the fourth edition of the *Diagnostic and Statistical Manual of Mental Disorders* (*DSM-IV*). Phenotypic analyses suggest that RD and other learning difficulties are more strongly associated with inattention symptoms than hyperactivity–impulsivity symptoms (e.g., Willcutt & Pennington, 2000a,b). Bivariate twin analyses indicate that the correlation between reading difficulties and inattention symptoms is almost entirely due to common genetic influences, whereas the lower correlation between reading and hyperactivity–impulsivity is primarily due to environmental influences (Willcutt, Pennington, & DeFries, 2000; Willcutt *et al*., 2003; Willcutt *et al.*, in press).

In summary, family and twin studies indicate that RD and ADHD are each familial and highly heritable. Bivariate twin analyses indicate that comorbidity between RD and ADHD is primarily due to common genetic influences, but suggest that these common genes are more strongly associated with inattention than hyperactivity–impulsivity. In the next section (IV.B.4) we review initial studies that have begun to search the genome to identify the specific genes that lead to the development of RD, ADHD, and their comorbidity.

4. Candidate Gene and Linkage Studies of RD and ADHD

Although an estimated 99.9% of the deoxyribonucleic acid (DNA) sequence that comprises the human genetic code is identical among all people, the genetic sequence varies at tens of thousands of locations across the remaining 0.1% of the human genome. These individual differences in the genetic code may lead to differences in protein production, which may then lead to individual differences in early brain development or adult brain functioning if the sequence difference occurs in a gene that is expressed in the central nervous system. Candidate gene analysis and linkage analysis are two primary methods that are used to identify the approximate location of genes that may contain sequence differences that influence disorders such as ADHD and RD.

a. Candidate Gene Studies. The candidate gene approach is extremely useful if previous research has identified specific physiological processes that are involved in a disorder. For example, based on evidence that ADHD is associated with dysfunction in the dopamine neurotransmitter system (e.g., Volkow *et al.*, 1998), nearly 100 candidate gene studies have tested if ADHD is associated with genes that influence dopamine or other related neurotransmitters, and significant associations have been reported for 15 different candidate genes (reviewed by Willcutt, in press). However, virtually all of these results have been replicated inconsistently or await independent replication, and each of these genes appears to account for a relatively small proportion of the total variance in ADHD symptoms in the population (e.g., Faraone *et al.*, 2001; Maher *et al.*, 2002).

Plausible candidate genes for RD have proved to be more difficult to identify, primarily because our understanding of the pathophysiology of RD is less advanced. Therefore, most molecular genetic studies of RD have conducted family-based linkage analyses, an alternative approach to identify regions of the genome that may contain genes that increase susceptibility to a disorder.

b. Linkage Studies. Linkage analysis takes advantage of the fact that alleles of genes that are close together on the same chromosome tend to be transmitted together across many generations, whereas alleles of genes that are far apart become separated over time due to recombination during meiosis (see Fisher and DeFries (2002) or Pennington (2002) for more information on linkage analysis). Linkage analysis typically does not identify the specific gene that is associated with increased risk for a disorder. Instead, this approach allows researchers to identify specific regions of the genome that may contain susceptibility loci for a disorder, and these regions can then be targeted for more extensive analysis.

Significant linkage for RD has been reported and replicated on chromosomes 1, 2, 3, 6, 15, and 18 (see review by Fisher and DeFries (2002)). The most consistent result, obtained in five independent samples, suggests that a gene on chromosome 6p21 leads to difficulties in reading, spelling, and a variety of other reading-related language measures (e.g., Cardon *et al.*, 1994, 1995; Gayan *et al.*, 1999). In addition to these six replicated linkage regions, it is likely that ongoing linkage studies will identify additional loci in the future.

Linkage studies of ADHD tell a similar story. In addition to the 15 genes identified by candidate gene studies, linkage studies have identified more than 10 additional regions of the genome that may contain genes that increase risk for ADHD (Willcutt *et al.*, 2002; Bakker *et al.*, 2003; Ogdie *et al.*, 2003). However, only a single region on chromosome 5p13 was significant in both genome scans that have been published (Bakker *et al.*, 2003; Ogdie *et al.*, 2003), and neither of these genome scans detected linkage in the regions of most of the known candidate genes for ADHD.

The results of candidate gene and linkage studies underscore two important points about the etiology of RD and ADHD. First, it is clear that multiple genetic and environmental risk factors are involved in the etiology of both disorders. Second, each of these risk factors has a relatively small effect on the final phenotype. Therefore, whereas each risk factor leads to a small increase in susceptibility to the disorder, few or none are necessary or sufficient to cause RD or ADHD. These small effect sizes and inconsistent results across studies are not unique to RD and ADHD, a similar pattern is apparent in genetic studies of other complex psychopathologies such as schizophrenia (e.g., Riley & McGuffin, 2000), addictive behaviors (e.g., Crabbe, 2002), and bipolar disorder (e.g., Craddock & Jones, 2001). In light of the complexity of these results, it is plausible that some of these genes may specifically increase risk for RD or ADHD, whereas others may have more general effects that increase risk for both disorders, sometimes resulting in comorbidity. In the next section (IV.B.5) we turn to studies that attempted to identify the genes that contribute to comorbidity by increasing risk for both RD and ADHD.

5. Linkage Studies of Comorbidity Between RD and ADHD

Linkage studies of comorbidity between RD and ADHD have begun to identify chromosomal regions that may contain a gene that increases risk for both disorders (Willcutt *et al.*, 2002; 2003, Loo *et al.*, 2004). In the first of these studies, Willcutt *et al.* (2002, 2003) reported that the well-replicated quantitative trait locus for RD on chromosome 6p21 also increases susceptibility to ADHD. In a somewhat different approach, Loo *et al.* (2004) screened the entire genome for genes that influence ADHD or RD in a sample of sibling pairs selected because both siblings met criteria for ADHD. Their results suggest that regions of chromosomes 16p and 17q may contain genes that increase susceptibility to both RD and ADHD. In contrast, their results also revealed several regions that were linked specifically to ADHD or RD.

Thus, although several of these results await independent replication, existing data provide the strongest support for the hypothesis that comorbidity between RD and ADHD is due, at least in part, to a common genetic etiology.

Although the specific functions of the genes that lead to comorbidity between RD and ADHD are unknown, one plausible model suggests that these shared genetic risk factors may cause a developmental change in a single pathophysiological substrate, and that this change then increases risk for both RD and ADHD. In this model the final phenotypic expression of this common susceptibility is then influenced by other genetic and environmental risk factors. Therefore, in some individuals this common risk factor would be expressed as RD alone, some individuals would meet criteria for ADHD alone, and some would meet criteria for both RD and ADHD. An important step in validating this hypothesis is to identify a neuropsychological deficit or other pathophysiological marker that reflects the common genetic risk for RD and ADHD (Willcutt *et al.*, in press a,b). Measures of this neurocognitive weakness may then facilitate future molecular genetic studies of RD, ADHD, and their comorbidity.

C. CONDUCT DISORDER AND ADHD

Conduct disorder (CD) and ADHD also occur together in 30–50% of the cases in both epidemiological and clinical samples (Biederman, Newcorn, & Sprich, 1991). The results of studies examining the etiology of comorbidity between ADHD and CD vary a great deal.

1. Phenotypic Tests

Many studies in the literature address whether the three independent disorders model (i.e., ADHD + CD is a third, independent disorder, or an etiological distinct subtype) explains the comorbidity between ADHD and CD. Although several researchers have noted the similarities between ADHD children with and without CD, including similarities in neurological "soft signs" and pre- and perinatal complications (e.g., August & Stewart, 1983), physical anomalies (e.g., August & Stewart, 1983; McGee, Williams, & Silva, 1984), and average intelligence (e.g., Loney & Milich, 1982; August & Stewart, 1983; McGee, Williams, & Silva, 1984), more researchers have noted the differences between ADHD children with and without CD and suggested that the two groups should be classified as two different types of ADHD. A similar idea is that CD with and without ADHD may constitute two different etiological types. In 1993, Moffitt presented her developmental taxonomy model of antisocial behavior, suggesting two categories of antisocial behavior that are distinct in etiology. The first category includes individuals who are antisocial at every stage of life (i.e., life-course-persistent), and the second category includes individuals who are antisocial only during adolescence (i.e., adolescence-limited). Moffitt noted that one of the risk characteristics in individuals with life-course-persistent antisocial behavior is hyperactivity.

Systematic reviews of studies examining ADHD only, CD only, and both ADHD and CD have reached differing conclusions. Lynam (1996) conducted a review of studies examining differences in children with hyperactivity–impulsivity–attention problems only, conduct problems only, and both hyperactivity–impulsivity–attention and conduct problems, and concluded that a "psychopathic deficit" is the underlying pathology for both kinds of symptoms in children with both sets of problems (comorbid children), but not in children with only one set of problems occurring alone. One of the main reasons for this conclusion was the finding that comorbid children have unique deficits (e.g., a distinct social information-processing pattern and qualitatively different errors on a continuous performance task) not found in children with problems only in hyperactivity–impulsivity–attention or conduct. In addition, some of these unique deficits (e.g., lowered autonomic reactivity) are also found in adult psychopathic individuals. Jensen, Martin, and Cantwell (1997) also conducted a systematic review of studies examining the differences among children with ADHD only, CD only, and ADHD + CD. Given several characteristics of children of ADHD + CD (e.g., earlier age of onset, greater male–female sex ratio, lower IQs, increased learning/reading difficulties), Jensen *et al*. concluded that there is enough evidence for a new diagnostic entity or a sub-classification of ADHD: ADHD, aggressive type.

Subsequently, Waschbusch (2002) conducted a meta-analysis of studies examining children with hyperactive-impulsive-attention problems only, conduct problems only, or both kinds of problems. Waschbusch found several differences between comorbid children and children with only one kind of problem. For example, comorbid children had more severe conduct problems, lower verbal IQ scores, more peer difficulties, and more adult offending than the children with only one kind of problem or controls. However, Waschbusch (2002) concluded that there was little evidence that comorbid children have deficits that are not also present to some degree in children with only one kind of problem. Also interesting is the fact that the general pattern of results found in studies reviewed by Waschbusch (i.e., the comorbid group was the most impaired on deficits that are also present in the other two groups) is the pattern expected when the correlated liabilities model is the correct comorbidity model (Rhee *et al.*, 2004).

Other alternative accounts of the comorbidity between ADHD and CD have been proposed based on phenotypic data. A longitudinal study examining ADHD and CD symptoms (Taylor *et al.*, 1996) reported that the outcome of the ADHD + CD group was similar to the ADHD only group and rejected the three independent disorders model. They also reported that childhood ADHD symptoms in the absence of CD symptoms predicted CD symptoms in adolescence, whereas childhood CD symptoms did not predict ADHD symptoms in adolescence. They concluded that ADHD symptoms are the major developmental risk factor and that

CD symptoms are epiphenomenal (i.e., support for the random multiformity of ADHD model).

A study examining the correlates of ADHD only, CD only, and ADHD + CD children (Schachar & Tannock, 1995) reported that ADHD only was associated with cognitive deficits, greater developmental delays, and greater reading problems, the CD only group had been exposed to significantly greater environmental adversity and had more severe problems in arithmetic, and that the ADHD + CD group had the correlates of both the ADHD only and the CD only groups. Given these results, they rejected the alternate forms model and the three independent disorders model. They asserted ADHD + CD is a hybrid of pure ADHD and pure CD and that comorbidity between ADHD and CD occurs because the risk factors for one disorder increase the probability of the risk factors for the second disorder.

In sum, phenotypic tests of the comorbidity between ADHD and CD have reached conflicting conclusions. We next examine whether behavior genetic tests can help resolve this conflict.

2. *Behavior Genetic Tests*

Faraone and his colleagues (Biederman *et al.*, 1992; Faraone *et al.*, 1991, 1997; Faraone, Biederman, & Monuteaux, 2000) took a different approach in testing the three independent disorders model in a series of family studies. They compared the risk for ADHD, CD, and ADHD + CD in the relatives of probands with ADHD only and ADHD + CD and reported two major results in all four studies. First, the risk of CD was greater in relatives of probands with ADHD + CD than in relatives of probands with ADHD only. Second, there was significant cosegregation of ADHD and CD (having one disorder increased the likelihood of having the other disorder) in the relatives of probands with ADHD + CD. Given these results, they concluded that ADHD only and ADHD + CD are etiologically distinct disorders (i.e., support for the three independent disorders model).

Given the evidence of support for the three independent disorders model in the literature, Holmes *et al.* (2002) considered the possibility of etiological heterogeneity in their examination of the association between the DRD4 gene and ADHD. Evidence of association was not found in the total sample, but significant association was found between the DRD4 gene and ADHD plus conduct problems.

Several multivariate behavior genetic studies using the twin method (Scarborough & Dobrich, 1990; Willcutt *et al.*, 1995; Silberg *et al.*, 1996; Nadder *et al.*, 1998, 2002; Thapar, Harrington, & McGuffin, 2001; Waldman *et al.*, 2001) examined whether comorbidity between ADHD and CD is due to shared genetic influences. All of these studies found a substantial overlap between the genetic influences on ADHD and the genetic influences on CD (i.e., support for the correlated liabilities model).

In a subsequent study (Rhee *et al.*, 2004c), we examined a wide range of alternative models explaining the comorbidity between ADHD and CD. As mentioned previously, a series of simulation studies (Rhee *et al.*, 2003, 2004b) showed that the Neale and Kendler model fitting approach does a better job of validly discriminating the three independent disorders model from other comorbidity models than family prevalence analyses, such as the ones used in Faraone, Biederman, & Monuteaux (2000).

All 13 alternative comorbidity models were tested in a twin sample enriched with individuals with ADHD or academic difficulties, with 110 monozygotic twin pairs and 182 dizygotic twin pairs. Of these models, several did not fit the data well and could be rejected; the three independent disorders model was one of these models. The models that fit the data and could not be rejected were random multiformity, random multiformity of B, extreme multiformity, extreme multiformity of B, correlated liabilities, A causes B, B causes A, and reciprocal causation. The best fitting model was the extreme multiformity of B model, which suggests that being affected by CD leads to increased risk for manifesting ADHD. A simulation study examining the validity of the Neale and Kendler model fitting approach found that mistakes in discrimination within and between the multiformity models and the correlated liabilities are common in small samples. Therefore, it is difficult to interpret this result as support for the extreme multiformity of B model as the "correct" hypothesis for the comorbidity between ADHD and CD. However, these results provide evidence against the three independent disorders model and support the results of the several twin studies concluding that there are significant shared genetic influences between ADHD and CD.

3. *Summary of Evidence Regarding Comorbidity of ADHD and CD*

In conclusion, the existing evidence regarding the causes of comorbidity between ADHD and CD is not consistent. Reviews of studies evaluating the three independent disorders model by examining the correlates or underlying deficits in groups of children with ADHD + CD, ADHD only, and CD only have reached different conclusions, with Jensen, Martin and Cantwell (1997), and Lynam (1996) supporting the three independent disorders model and Waschbusch (2002) and others concluding that there is little evidence for the three independent disorders model. Family studies examining the risk of ADHD and CD in relatives of probands with ADHD + CD, ADHD only, CD only, and controls provide support for the three independent disorders model, but a simulation study (Rhee *et al.*, 2003) suggests that the analyses used in theses studies are not valid tests of the three independent disorders model. Several studies have found support for other models for the comorbidity between ADHD and CD, including random multiformity of ADHD, risk factors for one disorder increasing the probability of risk factors for another disorder (a model not discussed by Neale and Kendler), and the correlated liabilities model. Our recent study, which is the only study to examine a wide range of comorbidity models using the Neale and Kendler model fitting approach (which has been validated by a simulation study), suggests that correlated liabilities is a more likely cause of the comorbidity between ADHD and CD rather than three independent disorders. Given the conflicting results in the literature, more studies examining the comorbidity between ADHD and CD using valid analytical approaches need to be conducted.

D. IMPLICATIONS FOR OTHER POSSIBLE PAIRS

We have reviewed what is known about the explanations for three of the possible six comorbidities among four common childhood disorders: RD, SSD, ADHD, and CD. At this point the reader may wonder what is known about the other three possible pairwise comorbidities among these four disorders.

We can depict the relations among these four disorders graphically (Figure 3). Each disorder is at the vertex of a rectangle and each comorbidity is a line connecting two vertices. More generally, the number of pairwise comorbidities among n disorders is $(n^2 - n/2)$. For instance, if one studied eight disorders, there would be 28 possible comorbidities. If we have only studied a subset of the possible comorbidities among a set of disorders, as is true in Figure 3, what we have already learned could place some constraints on possible solutions for the unknown comorbidities.

One possibility discussed by Angold, Costello, and Erkanli (1999) is that of "epiphenomenal" comorbidity. That is, if there are robust pairwise comorbidities between disorders A and B and between disorders B and C, the expected rate of co-occurrence of disorders A and C will be the product of these two other comorbidity rates. If this product is greater than the product of the prevalences of A and C, we will observe a comorbidity rate that appears greater than chance, but which is in fact mediated by the other two comorbidities. In other words, there is no relation between A and C independent of their relation to B. Angold, Costello, and Erkanli (1999) present evidence that the apparent comorbidity between CD and anxiety was epiphenomenal because it derived from the comorbidities

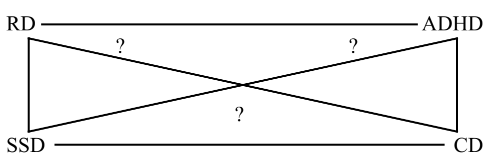

Fig. 3. Pairwise comorbidities among four disorders. RD = reading disability; SSD = speech sound disorder; ADHD = attention deficit hyperactivity disorder; CD = conduct disorder.

between CD and depression, and depression and anxiety. So some of the many non-artifactual comorbidities observed among DSM-IV diagnoses are likely to be epiphenomenal and we can winnow them from the list of comorbidities requiring a deeper explanation.

There are other ways in which what we have already learned could place constraints on less-studied comorbidities. For instance, if we had found that SSD was a risk factor for RD and that RD was a risk factor for ADHD (the causal model), then it would follow logically that SSD would be a risk factor for ADHD. Or, if we had found that the alternate forms model fit both the comorbidity between SSD and RD, and RD and ADHD, then we would expect there to be non-artifactual comorbidity between SSD and ADHD that also fits the alternate forms model.

Given what has actually been found about the three studied comorbidities in Figure 3, namely that each fits the correlated liabilities model, there are not sure predictions for the three less-studied comorbidities. Even if there is a partial etiological overlap between SSD and RD, and RD and ADHD, it does not necessarily follow that the liabilities between SSD and ADHD will be correlated.

What do we actually know empirically about the three less-studied comorbidities in Figure 3? In two studies, isolated SSD is not comorbid with ADHD, whereas SSD + LI is (Beitchman, Peterson, & Clegg, 1988; McGrath *et al.*, in preparation). Thus, isolated SSD may be etiologically and cognitively distinct from ADHD. We do not know of studies examining the relation between isolated SSD and CD, but it is well known that CD is associated with weaker language skills (e.g., Moffitt, 1993). RD and CD are comorbid (Hinshaw, 1992), but some research (Willcutt & Pennington, 2000a,b) finds this comorbidity is no longer present once ADHD is controlled, so there may not be a direct comorbidity between RD and CD. Instead, it may be epiphenomenal. A twin study (Trzesniewski, Moffitt, & Caspi, submitted) of this relation found evidence for a different possibility. While this study found there was a genetic overlap between ADHD and CD, consistent with earlier studies reviewed earlier, the relation between RD and CD was mediated environmentally.

In sum, as we learn more about comorbidities, we will be able to say more about developmental pathways from risk factors to outcomes, including where these pathways overlap and where they are distinct. This brings us to a more general model for thinking about relations between disorders.

V. Multifactorial Model

So what has research on these three comorbidities taught us about the development of disorders more generally? One lesson is that single etiology models of disorders do not seem to be adequate to account for either their development or comorbidity. We have yet to find a behaviorally defined disorder with a single necessary and sufficient etiology. The emerging etiological model for such disorders is probabilistic and multifactorial. But the prevailing cognitive model has often been deterministic and focused on a single cognitive cause, such as the phonological deficit in RD. So there is a potential contradiction in our frameworks for understanding such disorders that needs to be resolved. Another lesson is that a frequently supported explanation for comorbidity is correlated liabilities, specifically shared genetic risk factors. A third lesson is some disorders may be developmental precursors of later disorders, although usually an additional risk factor determines whether a child with the precursor disorder develops the later disorders. In this final section, we present a model that incorporates these lessons and is probabilistic and multifactorial at all levels of analysis.

Similar to the complex disease model in medicine (Sing & Reilly, 1993), this model includes six key proposals:

(1) The etiology of complex behavioral disorders is multifactorial and involves the interaction of multiple risk and protective factors, which can be either genetic or environmental.
(2) No single etiological factor is sufficient for a disorder, and few may be necessary.
(3) These risk and protective factors alter the development of psychological functions necessary for normal development, thus producing the behavioral symptoms that define these disorders.
(4) Few, if any, single cognitive risk factors are sufficient for a disorder, although some may be necessary.
(5) Consequently, comorbidity among complex behavioral disorders is to be expected because of shared etiologic and cognitive risk factors.
(6) The liability distribution for a given disease is often continuous and quantitative, rather than being discrete and categorical, so that the threshold for having the disorder is somewhat arbitrary.

So there are normally distributed individual differences in the behavioral dimensions (such as speech, reading, attention, and socially appropriate behavior) that define disorders. Those with a disorder fall beyond a somewhat arbitrary threshold on an extreme end of these distributions. The etiology of these individual differences is multifactorial, both across the whole distribution and at the extremes, and the etiologies of different behavioral dimensions partly overlap, producing cognitive overlap between dimensions and disorders.

Applying the model to the three comorbidities reviewed here, each individual disorder (SSD, RD, ADHD, and CD) has its own profile of risk factors (both etiologic and cognitive), with some of these risk factors being shared by pairs of disorders, resulting in comorbidity.

Figure 4 illustrates this multifactorial model, which is also discussed in Pennington (in press). There are four levels of analysis in this diagram: etiologic,

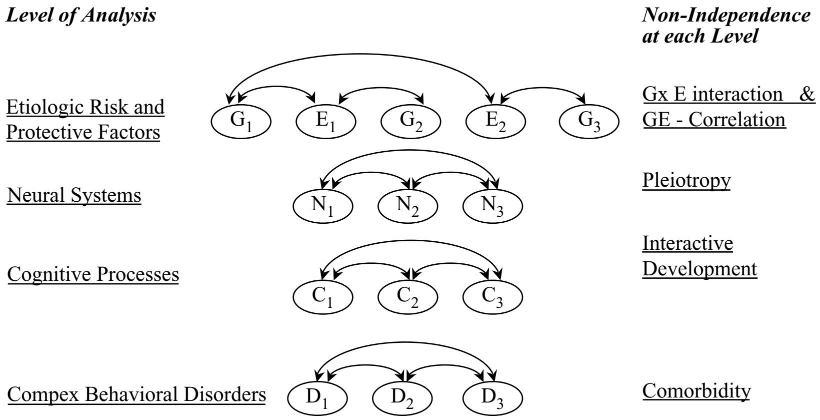

Fig. 4. Multifactorial model. G = genetic risk or protective factor, E = environmental risk or protective factor, N = neural system, C = cognitive process, D = disorder.

neural, cognitive, and symptom, where clusters of symptoms define complex behavioral disorders. For any such complex behavioral disorder, it is expected there will be more risk and protective factors than the five shown here. Bidirectional connections at each level indicate that constructs are not independent. For instance, at the etiologic level, there are likely to be gene–environment interactions and correlations. At the neural level, a single genetic or environmental risk factor will often affect more than one neural system (pleiotropy). Even if the risk factor initially only affects one neural system, this alteration will likely have downstream effects on the development of other neural systems. At the cognitive level, constructs are correlated because their developmental pathways overlap and because cognition is interactive. Overlap at the cognitive level leads to comorbidity at the symptom level. So, although a single deficit model conceptualizes the relation between disorders in terms of double dissociations, the multiple deficit model conceptualizes this relation in terms of partial overlap. At the symptom level, there is comorbidity (i.e., greater than chance co-occurrence) of complex behavioral disorders. Omitted from the diagram are the causal connections *between* levels of analyses, some of which would include feedback loops from behavior to brain or even to etiology. The existence and strength of these various causal connections must be determined empirically. The weights of the connections between levels of analysis will tell us to what extent different etiological and cognitive factors contribute to comorbidity at the symptom level.

It is also apparent that a similar but expanded model could be proposed for species-typical cognitive development, which results from the interaction of a largely shared genome (99.9% the same across unrelated humans) and species-typical environments. So in principle the same model could account for both typical and atypical development. Indeed, a complete account of any given developmental disorder will need to explain the many aspects of development that proceed typically as well as the few that go awry.

This model makes it clear that achieving a complete understanding of the development of disorders like SSD, RD, ADHD, or CD will be very difficult because of the multiple pathways and interactions involved. But this kind of model is needed because it is becoming increasingly clear that there are shared processes at the etiologic, neural, and cognitive levels across such disorders.

ACKNOWLEDGEMENTS

This work was supported by NIH grants MH38820, HD27802, HD04024, and DA13956. A portion of this chapter has been published in Rhee, S. H., Hewitt, J. K., Corley, R. P., & Stallings, M. C. (submitted). Comorbidity. In B. Everitt & D. C. Howell (Eds.), *Encyclopedia of Behavioral Statistics.* London: Wiley.

REFERENCES

American Psychiatric Association. (1994). *American Psychiatric Association Diagnostic and Statistical Manual of Mental Disorders* (3rd ed.). Washington, DC: American Psychiatric Association.

American Psychiatric Association. (2000). *American Psychiatric Association Diagnostic and Statistical Manual of Mental Disorders* (4th ed.). American Psychiatric Association: Washington, DC, Text Revision.

Angold, A., Costello, E. J., & Erkanli, A. (1999). Comorbidity. *Journal of Child Psychology and Psychiatry*, *40*, 57–87.

Aram, D. M., Ekelman, B. L., & Nation, J. E. (1984). Preschoolers with language disorders: 10 years later. *Journal of Speech and Hearing Research*, *27*, 232–244.

August, G. J., & Stewart, M. A. (1983). Familial subtypes of childhood hyperactivity. *Journal of Nervous Mental Disorders*, *171*, 362–368.

Bakker, S. C., van der Meulen, E. M., Buitelaar, J. K., Sandkuijl, L. A., Pauls, D. L., Monsuur, A. J. *et al*. (2003). A whole-genome scan in 164 Dutch sib pairs with attention-deficit/hyperactivity disorder: Suggestive evidence for linkage on chromosomes 7p and 15q. *American Journal of Human Genetics*, *72*, 1251–1260.

Beitchman, J. H., Peterson, M., & Clegg, M. (1988). Speech and language impairment and psychiatric disorder: The relevance of family demographic variables. *Child Psychiatry and Human Development*, *18*, 191–207.

Berkson, J. (1946). Limitations of the application of fourfold table analysis to hospital data. *Biometrics*, *2*, 47–51.

Biederman, J., Newcorn, J., & Sprich, S. (1991). Comorbidity of attention deficit hyperactivity disorder with conduct, depressive, anxiety, and other disorders. *American Journal of Psychiatry*, *148*, 564–577.

Biederman, J., Faraone, S. V., Keenan, K., Benjamin, J., Krifcher, B., Moore, C. *et al*. (1992). Further evidence for family-genetic risk factors in attention deficit hyperactivity disorder. Patterns of comorbidity in probands and relatives psychiatrically and pediatrically referred samples. *Archives of General Psychiatry*, *49*, 728–738.

Bierut, L. J., Dinwiddie, S. H., Begleiter, H., Crowe, R. R., Hesselbrock, V., Nurnberger, J. I. Jr. *et al*. (1998). Familial transmission of substance dependence: Alcohol, marijuana, cocaine, and habitual smoking: A report from the Collaborative Study on the Genetics of Alcoholism. *Archives of General Psychiatry*, *55*, 982–988.

Bird, J., & Bishop, D. (1992). Perception and awareness of phonemes in phonologically impaired children. *European Journal of Disorders in Communication*, *27*, 289–311.

Bird, J., Bishop, D. V., & Freeman, N. H. (1995). Phonological awareness and literacy development in children with expressive phonological impairments. *Journal of Speech and Hearing Research*, *38*, 446–462.

Bird, H., Canino, G., Rubio-Stipec, M., Gould, M., Ribera, J., Sesman, M. *et al*. (1988). Estimates of the prevalence of childhood maladjustment in a community survey in Puerto Rico: The use of combined measures. *Archives of General Psychiatry*, *45*, 1120–1126.

Bishop, D. V., & Adams, C. (1990). A prospective study of the relationship between specific language impairment, phonological disorders and reading retardation. *Journal of Child Psychology and Psychiatry*, *31*, 1027–1050.

Bishop, D. V., North, T., & Donlan, C. (1995). Genetic basis of specific language impairment: Evidence from a twin study. *Developmental and Medical Child Neurology*, *37*, 56–71.

Cardon, L. R., Smith, S. D., Fulker, D. W., Kimberling, W. J., Pennington, B. F., & DeFries, J. C. (1994). Quantitative trait locus for reading disability on chromosome 6. *Science*, *266*, 276–279.

Cardon, L. R., Smith, S. D., Fulker, D. W., Kimberling, W. J., Pennington, B. F., & DeFries, J. C. (1995). Quantitative trait locus for reading disability: Correction. *Science*, *268*, 1553.

Caron, C., & Rutter, M. (1991). Comorbidity in child psychopathology: Concepts, issues and research strategies. *Journal of Child Psychology and Psychiatry*, *32*, 1063–1080.

Catts, H. W., Fey, M. E., Tomblin, J. B., & Zhang, X. (2002). A longitudinal investigation of reading outcomes in children with language impairments. *Journal of Speech, Language, and Hearing Research*, *45*, 1142–1157.

Clark, L. A., Watson, D., & Reynolds, S. (1995). Diagnosis and classification of psychopathology: Challenges to the current system and future directions. *Annual Review of Psychology*, *46*, 121–153.

Clarke-Klein, S., & Hodson, B. W. (1995). A phonologically based analysis of misspellings by third graders with disordered-phonology histories. *Journal of Speech, and Hearing Research*, *38*, 839–849.

Crabbe, J. C. (2002). Genetic contributions to addiction. *Annual Review of Psychology*, *53*, 435–462.

Craddock, N., & Jones, I. (2001). Molecular genetics of bipolar disorder. *British Journal of Psychiatry*, *178*(Suppl. 41), S128–S133.

Dickman, G. E. (2003). The nature of learning disabilities through the lens of reading research. *Perspectives: The International Dyslexia Association*, *29*, 4–8.

Donaldson, S. K., Klein, D. N., Riso, L. P., & Schwartz, J. E. (1997). Comorbidity between dysthymic and major depressive disorders: A family study analysis. *Journal of Affective Disorders*, *42*, 103–111.

Doyle, A. E., Faraone, S. V., DuPre, E. P., & Biederman, J. (2001). Separating attention deficit hyperactivity disorder and learning disabilities in girls: A familial risk analysis. *American Journal of Psychiatry*, *158*, 1666–1672.

Dykman, R. A., & Ackerman, P. T. (1991). Attention deficit disorder and specific reading disability: Separate but often overlapping disorders. *Journal of Learning Disabilities*, *24*, 96–103.

Edwards, J., & Lahey, M. (1998). Nonword repetitions in children with specific language impairment: Exploration of some explanations for their inaccuracies. *Applied Psycholinguistics*, *19*, 279–309.

Faraone, S. V., Biederman, J., & Friedman, D. (2000). Validity of DSM-IV subtypes of attention-deficit/hyperactivity disorder: A family study perspective. *Journal of the American Academy of Child and Adolescent Psychiatry*, *39*, 300–307.

Faraone, S. V., Biederman, J., Keenan, K., & Tsuang, M. T. (1991). A family-genetic study of girls with DSM-III attention deficit disorder. *American Journal of Psychiatry*, *148*, 112–117.

Faraone, S. V., Biederman, J., Jetton, J. G., & Tsuang, M. T. (1997). Attention deficit disorder and conduct disorder: Longitudinal evidence for a familial subtype. *Psychological Medicine*, *27*, 291–300.

Faraone, S. V., Biederman, J., Lehman, B. K., Keenan, K., Norman, D., Seidman, L. J. *et al.* (1993). Evidence for the independent familial transmission of attention deficit hyperactivity disorder and learning disabilities: Results from a family genetic study. *American Journal of Psychiatry*, *150*, 891–895.

Faraone, S. V., Biederman, J., & Monuteaux, M. C. (2000). Attention-deficit disorder and conduct disorder in girls: Evidence for a familial subtype. *Biological Psychiatry*, *48*, 21–29.

Faraone, S., Doyle, A., Mick, E., & Biederman, J. (2001). Meta-analysis of the association between the 7-repeat allele of the dopamine D4 receptor gene and attention deficit hyperactivity disorder. *American Journal of Psychiatry*, *158*, 1052–1057.

Feinstein, A. (1970). The pre-therapeutic classification of co-morbidity in chronic disease. *Journal of Chronic Disease*, *23*, 57–63.

Fergusson, D. M., & Horwood, L. J. (1992). Attention deficit and reading achievement. *Journal of Child Psychology and Psychiatry*, *33*, 375–385.

Fisher, S. E., & DeFries, J. C. (2002). Developmental dyslexia: Genetic dissection of a complex cognitive trait. *Nature Reviews Neuroscience*, *3*, 767–780.

Friedman, M. C., Chhabildas, N., Budhiraja, N., Willcutt, E. G., & Pennington, B. F. (2003). Etiology of the comorbidity between RD and ADHD: Exploration of the non-random mating hypothesis. *American Journal of Medical Genetics*, *120B*, 109–115.

Gallagher, A., Frith, U., & Snowling, M. J. (2000). Precursors of literacy delay among children at genetic risk of dyslexia. *Journal of Child Psychology and Psychiatry*, *41*, 203–213.

Gayan, J., Smith, S. D., Cherny, S. S., Cardon, L. R., Fulker, D. W., Brower, A. M. *et al.* (1999). Quantitative-trait locus for specific language and reading deficits on chromosome 6p. *American Journal of Human Genetics*, *64*, 157–164.

Gilger, J. W., Borecki, I. B., DeFries, J. C., & Pennington, B. F. (1994). Comingling and segregation analysis of reading performance in families of normal reading probands. *Behavioral Genetics*, *24*, 345–355.

Gilger, J., Pennington, B., & DeFries, J. (1991). Risk for reading disability as a function of parental history in three family studies. *Reading and Writing: An Interdisciplinary Journal*, *3*, 205–217.

Gilger, J. W., Pennington, B. F., & DeFries, J. C. (1992). A twin study of the etiology of comorbidity: Attention-deficit hyperactivity disorder and dyslexia. *Journal of the American Academy Child and Adolescent Psychiatry*, *31*, 343–348.

Hall, P. K., & Tomblin, J. B. (1978). A follow-up study of children with articulation and language disorders. *Journal of Speech and Hearing Disorders*, *43*, 227–241.

Hallgren, B. (1950). Specific dyslexia: A clinical and genetic study. *Acta Psychiatrica Neurologica Scandinavia*, *65*(Suppl.), 1–287.

Hinshaw, S. P. (1992). Externalizing behavior problems and academic underachievement in childhood and adolescence: Causal relationships and underlying mechanisms. *Psychological Bulletin*, *111*, 127–155.
Holmes, J., Payton, A., Barrett, J., Harrington, R., McGuffin, P., Owen, M. *et al*. (2002). Association of DRD4 in children with ADHD and comorbid conduct problems. *American Journal of Medical Genetics*, *114*, 150–153.
Jensen, P. S., Martin, D., & Cantwell, D. P. (1997). Comorbidity in ADHD: Implications for research, practice, and DSM-V. *Journal of the Academy of Child and Adolescent Psychiatry*, *36*, 1065–1079.
Jensen, P. S., Watanabe, H. K., Richters, J. E., Cortes, R., Roper, M., & Liu, S. (1995). Prevalence of mental disorder in military children and adolescents: Findings from a two-stage community survey. *Journal of the Academy of Child and Adolescent Psychiatry*, *34*, 1514–1524.
Kamhi, A. G., Catts, H. W., Mauer, D., Apel, K., & Gentry, B. F. (1988). Phonological and spatial processing abilities in language- and reading-impaired children. *Journal of Speech and Hearing Disorders*, *53*, 316–327.
Kessler, R. C., McGonagle, K. A., Zhao, S., Nelson, C. B., Hughes, M., Eshleman, S. *et al*. (1994). Lifetime and 12-month prevalence of DSM-III-R psychiatric disorders in the United States. Results from the National Comorbidity Survey. *Archives of General Psychiatry*, *51*, 8–19.
Klein, D. F. (1993). False suffocation alarms, spontaneous panics, and related conditions. An integrative hypothesis. *Archives of General Psychiatry*, *50*, 306–317.
Klein, D. N. (2003). Patients' versus informants' reports of personality disorders in predicting $7\frac{1}{2}$-year outcome in outpatients with depressive disorders. *Psychological Assessment*, *15*, 216–222.
Klein, D. N., & Riso, L. (1993). Psychiatric disorders: Problems of boundaries and comorbidity. In C. G. Costello (Ed.), *Basic issues in psychopathology* (pp. 19–66). New York: Guilford Press.
Leonard, L. B. (1982). Phonological deficits in children with developmental language impairment. *Brain and Language*, *16*, 73–86.
Lewis, B. A. (1990). Familial phonological disorders: Four pedigrees. *Journal of Speech and Hearing Disorders*, *55*, 160–170.
Lewis, B. (1992). Genetics in speech disorders. *Clinical Communication Disorders*, *2*, 48–58.
Lewis, B. A., Ekelman, B. L., & Aram, D. M. (1989). A familial study of severe phonological disorders. *Journal of Speech and Hearing Research*, *32*, 713–724.
Lewis, B. A., & Freebairn, L. (1992). Residual effects of preschool phonology disorders in grade school, adolescence, and adulthood. *Journal of Speech and Hearing Research*, *35*, 819–831.
Light, J. G., Pennington, B. F., Gilger, J. W., & DeFries, J. C. (1995). Reading disability and hyperactivity disorder: Evidence for a common genetic etiology. *Developmental Neuropsychology*, *11*, 323–335.
Lilienfeld, S., Waldman, I., & Israel, A. (1994). Critical examination of the use of the term and concept of comorbidity in psychopathology research. *Clinical Psychology and Scientific Practice*, *1*, 71–83.
Loney, J., & Milich, R. (1982). Hyperactivity, inattention, and aggression in clinical practice. In M. Wolraich & D. Routh (Eds.), *Advances in developmental and behavioral pediatrics* (Vol. 3, pp. 113–147). Greenwich, CT: JAI Press.
Loo, S. K., Fisher, S. E., Francks, C., Ogdie, M. N., MacPhie, I. L., Yang, M. *et al*. (2004). Genome-wide scan of reading ability in affected sibling pairs with attention-deficit/hyperactivity disorder: Unique and shared genetic effects. *Molecular Psychiatry*, *9*, 485–493.
Lynam, D. R. (1996). Early identification of chronic offenders: Who is the fledgling psychopath? *Psychological Bulletin*, *120*, 209–234.

Lyytinen, H., Ahonen, T., Eklund, K., Guttorm, T. K., Laakso, M. L., Leinone, S. *et al*. (2002). Developmental pathways of children with and without familial risk for dyslexia during the first years of life. *Developmental Neuropsychology*, *20*, 535–554.

Magnusson, E., & Naucler, K. (1990). Reading and spelling in language disordered children-linguistic and metalinguistic pre-requisites: A report on a longitudinal study. *Clinical Linguistics and Phonetics*, *4*, 49–61.

Maher, B. S., Marazita, M. L., Ferrell, R. E., & Vanyukov, M. M. (2002). Dopamine system genes and attention deficit hyperactivity disorder: A meta-analysis. *Psychiatric Genetics*, *12*, 207–215.

McGee, R., Williams, S., & Silva, P. A. (1984). Behavioral and developmental characteristics of aggressive, hyperactive and aggressive-hyperactive boys. *Journal of the American Academy of Child Psychiatry*, *23*, 270–279.

McGrath, L. M., Hutaff, C., Scott, A., & Pennington, B. F. (in preparation). Attention deficit/hyperactivity disorder in speech sound disorder. A subgroup analysis.

Merikangas, K. (1982). Assortative mating for psychiatric disorders and psychological traits. *Archives of General Psychiatry*, *39*, 1173–1180.

Moffitt, T. E. (1993). Adolescence-limited and life-course-persistent antisocial behavior: A developmental taxonomy. *Psychological Review*, *100*, 674–701.

Montgomery, J. W. (1995). Sentence comprehension in children with specific language impairment: The role of phonological working memory. *Journal of Speech and Hearing Research*, *38*, 187–199.

Nadder, T. S., Rutter, M., Silberg, J. L., Maes, H. H., & Eaves, L. J. (2002). Genetic effects on the variation and covariation of attention deficit-hyperactivity disorder (ADHD) and oppositional-defiant disorder/conduct disorder (Odd/CD) symptomatologies across informant and occasion of measurement. *Psychological Medicine*, *32*, 39–53.

Nadder, T. S., Silberg, J. L., Eaves, L. J., Maes, H. H., & Meyer, J. M. (1998). Genetic effects on ADHD symptomatology in 7- to 13-year-old twins: Results from a telephone survey. *Behavior Genetics*, *28*, 83–99.

Neale, M. C., & Kendler, K. S. (1995). Models of comorbidity for multifactorial disorders. *American Journal of Human Genetics*, *57*, 935–953.

Nigg, J. T., Hinshaw, S. P., Carte, E. T., & Treuting, J. J. (1998). Neuropsychological correlates of childhood attention-deficit/hyperactivity disorder: Explainable by comorbid disruptive behavior or reading problems? *Journal of Abnormal Psychology*, *107*, 468–480.

Ogdie, M. N., Macphie, I. L., Minassian, S. L., Yang, M., Fisher, S. E., Francks, C. *et al*. (2003). A genomewide scan for attention-deficit/hyperactivity disorder in an extended sample: Suggestive linkage on 17p11. *American Journal of Human Genetics*, *72*, 1268–1279.

Orton, S. T. (1925). Word-blindness in school children. *Archives of Neurology and Psychiatry*, *14*, 582–615.

Pennington, B. F. (2002). *The development of psychopathology: Nature and nurture*. New York: Guilford.

Pennington, B. F. A multiple deficit model for understanding developmental disorders. *Cognition*. in press.

Pennington, B. F., Groisser, D., & Welsh, M. C. (1993). Contrasting cognitive deficits in attention deficit hyperactivity disorder versus reading disability. *Developmental Psychology*, *29*, 511–523.

Pennington, B. F., & Lefly, D. L. (2001). Early reading development in children at family risk for dyslexia. *Child Development*, *72*, 816–833.

Pisecco, S., Baker, D., Silva, P., & Brooke, M. (2001). Boys with reading disabilities and/or ADHD: Distinctions in early childhood. *Journal of Learning Disabilities*, *34*, 98–106.

Raitano, N. A., Pennington, B. F., Tunick, R. A., Boada, R., & Shriberg, L. D. (2004). Preliteracy skills of subgroups of children with speech sound disorders. *Journal of Child Psychology and Psychiatry*, *45*, 821–835.

Rhee, S., Hewitt, J., Corley, R., & Stallings, M. (2003). The validity of analyses testing the etiology of comorbidity between two disorders: A review of family studies. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *44*, 612–636.

Rhee, S. H., Hewitt, J. K., Corley, R. P., & Stallings, M.C. (in press). Comorbidity. In B. Everitt & D.C. Howell (Eds.), *Encyclopedia of behavioral statistics.* London: Wiley.

Rhee, S. H., Hewitt, J. K., Corley, R. P., Willcutt, E., & Pennington, B. (2004a). *Testing hypotheses regarding the causes of comorbidity: Examining the underlying deficits of comorbid disorders.* Unpublished manuscript.

Rhee, S. H., Hewitt, J. K., Lessem, J. M., Stallings, M. C., Corley, R. P., & Neale, M. C. (2004b). The validity of the Neale and Kendler model-fitting approach in examining the etiology of comorbidity. *Behavior Genetics*, *34*, 251–265.

Rhee, S., Willcutt, E., Hartman, C., Pennington, B., & DeFries, J. (2004c). *Test of alternative hypotheses explaining the comorbidity between attention-deficit/hyperactivity disorder and conduct disorders.* Paper presented at the Behavior Genetics Association, Aix-en-Provance, France.

Riley, B. P., & McGuffin, P. (2000). Linkage and associated studies of schizophrenia. *American Journal of Medical Genetics*, *97*, 23–44.

Riso, L. P., Klein, D. N., Ferro, T., Kasch, K. L., Pepper, C. M., Schwartz, J. E. *et al*. (1996). Understanding the comorbidity between early-onset dysthymia and cluster B personality disorders: A family study. *American Journal of Psychiatry*, *153*, 900–906.

Robins, L., Locke, B., & Regier, D. (1991). An overview of psychiatric disorders in America. In L. Robins & B. Locke (Eds.), *Psychiatric disorders in America* (pp. 328–366). New York: Free.

Rucklidge, J. J., & Tannock, R. (2002). Neuropsychological profiles of adolescents with ADHD: Effects of reading difficulties and gender. *Journal of Child Psychology and Psychiatry*, *43*, 988–1003.

Rutter, M., & Mawhood, L. (1991). The long-term psychosocial sequelae of specific developmental disorders of speech and language. In M. Rutter & P. Casaer (Eds.), *Biological risk factors for psychosocial disorders* (pp. 233–259). Cambridge: Cambridge University Press.

Rutter, M., & Yule, W. (1975). The concept of specific reading retardation. *Journal Child Psychology and Psychiatry*, *16*, 181–197.

Scarborough, H. S. (1990). Very early language deficits in dyslexic children. *Child Development*, *61*, 1728–1743.

Scarborough, H. S., & Dobrich, W. (1990). Development of children with early language delay. *Journal of Speech and Hearing Research*, *33*, 70–83.

Schachar, R., & Tannock, R. (1995). Test of four hypotheses for the comorbidity of attention-deficit hyperactivity disorder and conduct disorder. *Journal of the Academy of Child and Adolescent Psychiatry*, *34*, 639–648.

Seidman, L. J., Biederman, J., Monuteaux, M. C., Doyle, A. E., & Faraone, S. V. (2001). Learning disabilities and executive dysfunction in boys with attention-deficit/hyperactivity disorder. *Neuropsychology*, *15*, 544–556.

Semrud-Clikeman, M., Biederman, J., Sprich-Buckminster, S., Lehman, B. K., Faraone, S. V., & Norman, D. (1992). Comorbidity between ADDH and learning disability: A review and report in a clinically referred sample. *Journal of the American Academy of Child and Adolescent Psychiatry*, *31*, 439–448.

Shaywitz, B. A., Fletcher, J. M., & Shaywitz, S. E. (1995). Defining and classifying learning disabilities and attention-deficit/hyperactivity disorder. *Journal of Child Neurology*, *10*(Suppl. 1), S50–S57.

Shriberg, L. D., Tomblin, J. B., & McSweeny, J. L. (1999). Prevalence of speech delay in 6-year-old children and comorbidity with language impairment. *Journal of Speech, Language, and Hearing Research*, *42*, 1461–1481.

Silberg, J., Rutter, M., Meyer, J., Maes, H., Hewitt, J., Simonoff, E. *et al*. (1996). Genetic and environmental influences on the covariation between hyperactivity and conduct disturbance in juvenile twins. *Journal of Child Psychology and Psychiatry*, *37*, 803–816.

Sing, C. F., & Reilly, S. L. (1993). Genetics of common diseases that aggregate, but do not segregate in families. In C. F. Sing & C. L. Hanis (Eds.), *Genetics of cellular, individual, family and population variability* (pp. 140–161). New York: Oxford University Press.

Smith, S. D., Deffenbacher, K. E., Boada, R., Tunick, R. A., Raitano, N. A. & Pennington, B. (2003). *Speech sound disorder is linked to dyslexia risk loci on chromosome 6 and 15*. Paper Presented at the 53rd Annual Meeting of American Society for Human Genetics, Los Angeles, CA. November.

Snowling, M., Bishop, D. V., & Stothard, S. E. (2000). Is preschool language impairment a risk factor for dyslexia in adolescence? *Journal of Child Psychology and Psychiatry*, *41*, 587–600.

Snowling, M., & Stackhouse, J. (1983). Spelling performance of children with developmental verbal dyspraxia. *Developmental and Medical Child Neurology*, *25*, 430–437.

Stein, C. M., Schick, J. H., Gerry Taylor, H., Shriberg, L. D., Millard, C., Kundtz-Kludge, A. *et al*. (2004). Pleiotropic effects of a chromosome 3 locus on speech-sound disorder and reading. *American Journal of Human Genetics*, *74*, 283–297.

Steingard, R., Biederman, J., Doyle, A., & Sprich-Buckminster, S. (1992). Psychiatric comorbidity in attention deficit disorder: Impact on the interpretation of Child Behavior Checklist results. *Journal of the Academy of Child and Adolescent Psychiatry*, *31*, 449–454.

Stevenson, J., Pennington, B. F., Gilger, J. W., DeFries, J. C., & Gillis, J. J. (1993). Hyperactivity and spelling disability: Testing for shared genetic etiology. *Journal of Child Psychology and Psychiatry*, *14*, 1137–1152.

Swanson, H., Mink, J., & Bocian, K. (1999). Cognitive processing deficits in poor readers with symptoms of reading disabilities and ADHD: More alike than different? *Journal of Educational Psychology*, *91*, 321–333.

Taylor, E., Chadwick, O., Heptinstall, E., & Danckaerts, M. (1996). Hyperactivity and conduct problems as risk factors for adolescent development. *Journal of the Academy of Child and Adolescent Psychiatry*, *35*, 1213–1226.

Thapar, A., Harrington, R., & McGuffin, P. (2001). Examining the comorbidity of ADHD-related behaviours and conduct problems using a twin study design. *British Journal of Psychiatry*, *179*, 224–229.

Tomblin, J. B., Freese, P. R., & Records, N. L. (1992). Diagnosing specific language impairment in adults for the purpose of pedigree analysis. *Journal of Speech and Hearing Research*, *35*, 832–843.

Trzesniewski, K., Moffitt, T. E., & Caspi, A. Revisiting the association between reading achievement and antisocial behavior: New evidence of an environmental explanation from a twin study. *Child Development.* submitted.

Tunick, R. A. (2004). *Cognitive overlap between reading disability and speech sound disorder: A test of the severity hypothesis.* Unpublished doctoral dissertation. University of Denver, Denver, CO.

Tunick, R. A., Boada, R., Raitano, N. A., & Pennington, B. F. *Cofamiliality of speech sound disorder and reading disability.* in preparation.

Tunick, R. A., & Pennington, B. F. (2002). The etiological relationship between reading disability and phonological disorder. *Annals of Dyslexia*, *52*, 75–95.

Volkow, N. D., Wang, G. J., Fowler, J. S., Gatley, S. J., Logan, J., Ding, Y. S. *et al.* (1998). Dopamine transporter occupancies in the human brain induced by therapeutic doses of oral methylphenidate. *American Journal of Psychiatry*, *155*, 1325–1331.

Wadsworth, S. J., Davis, C., Knopik, V., Willcutt, E., & DeFries, J. (2002). Genetics of reading disabilities. *Proceedings of the 12th Postgraduate Course in Pediatric Neurology: Learning Disabilities in Reading and Writing*, 23–36.

Wagner, R. K., & Torgesen, J. K. (1987). The nature of phonological processing and its causal role in the acquisition of reading skills. *Psychological Bulletin*, *101*, 192–212.

Waldman, I., Rhee, S., Levy, F., & Hay, D. (2001). Causes of the overlap among symptoms of ADHD, oppositional defiant disorder, and conduct disorder. In F. Levy & D. Hay (Eds.), *Attention, genes, and ADHD* (pp. 115–138). East Sussex, England: Brunner-Routtledge.

Waschbusch, D. A. (2002). A meta-analytic examination of comorbid hyperactive-impulsive-attention problems and conduct problems. *Psychological Bulletin*, *128*, 118–150.

Wickramaratne, P. J., & Weissman, M. M. (1993). Using family studies to understand comorbidity. *European Archives of Psychiatry and Clinical Neuroscience*, *243*, 150–157.

Willcutt, E. (in press). Genetics of ADHD. In D. Barch (Ed.), *Cognitive and affective neuroscience psychopathology*. Oxford: Oxford University Press.

Willcutt, E., Chhabildas, N., & Pennington, B. (1998). *Psychiatric correlates of reading disability*. Paper Presented at the Kansas Conference on Clinical Child Psychology, Lawrence, KS.

Willcutt, E. G., DeFries, J. C., Pennington, B. F., Smith, S. D., Cardon, L. R., & Olson, R. K. (2003). Genetic etiology of comorbid reading difficulties and ADHD. In R. Plomin, J. C. DeFries, I. W. Craig, & P. McGuffin (Eds.), *Behavioral genetics in the postgenomic era* (pp. 227–246). Washington, DC: American Psychological Association.

Willcutt, E. G., & Pennington, B. F. (2000a). Comorbidity of reading disability and attention deficit/hyperactivity disorder: Differences by gender and subtype. *Journal of Learning Disabilities*, *33*, 179–191.

Willcutt, E. G., & Pennington, B. F. (2000b). Psychiatric comorbidity in children and adolescents with reading disability. *Journal of Child Psychology and Psychiatry*, *41*, 1039–1048.

Willcutt, E. G., Pennington, B. F., Boada, R., Ogline, J. S., Tunick, R. A., Chhabildas, N. A. *et al.* (2001). A comparison of the cognitive deficits in reading disability and attention-deficit/hyperactivity disorder. *Journal of Abnormal Psychology*, *110*, 157–172.

Willcutt, E., Pennington, B., Olson, R., Chhabildas, N., & Hulslander, J. (2005). Neuropsychological analyses of comorbidity between RD and ADHD: In search of the common deficit. *Developmental Neuropsychology*, *27*(1), 35–78.

Willcutt, E. G., Pennington, B. F., & DeFries, J. C. (2000). Twin study of the etiology of comorbidity between reading disability and attention-deficit/hyperactivity disorder. *American Journal of Medical Genetics*, *96*, 293–301.

Willcutt, E., Pennington, B., Olson, R., & DeFries, J. (in press). Understanding comorbidity: A twin study of reading disability and attention deficit/hyperactivity disorder. *American Journal of Medical Genetics (Neuropsychiatric Genetics).*

Willcutt, E., Shyu, V., Green, P., & Pennington, B. (1995). *A twin study of the comorbidity of the disruptive behavior disorders of childhood.* Paper Presented at the Society for Research in Child Development, Indianapolis, IN.

Willcutt, E. G., Pennington, B. F., Smith, S. D., Cardon, L. R., Gayan, J., Knopik, V. S., *et al.* (2002). Quantitative trait locus for reading disability on chromosome 6p is pleiotropic for attention-deficit/hyperactivity disorder. *American Journal of Medical Genetics*, *114*, 260–268.

NUMBER WORDS AND NUMBER CONCEPTS: THE INTERPLAY OF VERBAL AND NONVERBAL QUANTIFICATION IN EARLY CHILDHOOD

Kelly S. Mix

DEPARTMENT OF PSYCHOLOGY AND PROGRAM IN COGNITIVE SCIENCE,
INDIANA UNIVERSITY, BLOOMINGTON, IN 47405, USA

Catherine M. Sandhofer

DEPARTMENT OF PSYCHOLOGY, UNIVERSITY OF CALIFORNIA, LOS ANGELES, CA 90095, USA

Arthur J. Baroody

COLLEGE OF EDUCATION, UNIVERSITY OF ILLINOIS, CHAMPAIGN, IL 61820, USA

I. Introduction

A. A KEY QUESTION

The question of how language influences concepts is an old one. Do words merely map onto pre-existing concepts? Or do words actually create the concepts

Advances in Child Development and Behavior
R Kail (Editor)

to which they refer? Would children have notions of different colors, textures, or numbers if they never were exposed to the linguistic labels for these ideas? Interest in these questions can be traced deep into the philosophical roots of psychology and throughout a good deal of the empirical investigation that has been carried out since. However, despite the rhetoric surrounding these issues, we seem to know few actual details of the way language and cognition interact.

This problem is particularly acute in the area of numerical development, for which children must integrate many layers of verbal, procedural, symbolic, and conceptual meaning. To illustrate the complexity involved, consider this $3\frac{1}{2}$-year-old research participant's understanding of "five."

> Experimenter: "Can you give me five [blocks]?"
> Child: (holds up five fingers) "This is five."
> Experimenter: "Can you give me five blocks?"
> Child: (lays out 15 blocks and counts them) "One, two, three, four, eight, fiveteen. There's five!"
> Experimenter: "Okay. Can you count these for me?" (handing the child an array of 10 blocks glued on a board)
> Child: "One, two, three, four, eight."
> Experimenter: "How many is that?"
> Child: "I don't know. Dad, do you know?"

This child has learned what five fingers are and, in that limited context, could be said to understand the concept of "five." Although she confuses "five" and "fiveteen" in her count, she seems to know that counting determines cardinal number and that the last word in a count has special meaning. She uses counting to "prove" that her pile of 15 blocks equals "five." But when it comes to performing unfamiliar experimental tasks or counting larger sets, these understandings seem to evaporate. How, then, should we characterize her status? Does she understand "five" or doesn't she? And what does this tell us about the more general process by which children bring meaning to the number words?

In this chapter, we review what is known and what has been proposed regarding the interactions between number words and number concepts. We argue that both classic and current conceptualizations have obscured the rich detail of these interactions by asking, "Which comes first, language or concepts?" As the experimental transcript illustrates, number development viewed close-up is not so orderly. Indeed, we find that this polarizing framework has limited progress in two specific ways: (1) by attempting to separate empirically what cannot be separated developmentally and (2) by casting developmental change in terms of months or years instead of days, or even moments. We conclude the chapter by reviewing case study, microgenetic, and longitudinal research that reveals how fluid and tightly woven the interplay of verbal and nonverbal quantification really is.

B. WHAT DOES IT MEAN TO HAVE A NUMBER CONCEPT?

Like many concepts, number encompasses a variety of perceptual and symbolic inputs. But, it also has aspects that make it unique. For example, consider what it means to understand "five." This notion can be instantiated in groups of objects that vary widely along every other dimension (e.g., five fleas, five skyscrapers, five planets). It can include groups of non-objects, such as sounds, visual events (e.g., lights blinking), actions, ideas, or emotions. These groups may come together in space (e.g., five cookies on a plate), time (e.g., the five cookies I ate last night) or based on function (e.g., the five cookies I know how to bake). The ability to see these diverse groupings as equivalent is a large part of what might be considered nonverbal number concepts.

Number also can be represented symbolically in various ways; as a spoken word (e.g., "five"), a written word, (e.g., five), or as a written numeral (e.g., 5). These symbols for numbers can vary in their intended meaning. Sometimes they simply refer to the number of objects in a set (cardinal meaning). However, they also can refer to a set of measurement units (e.g., five inches, five years, five cups, etc.) or to less standardized measures, such as clothing size (measurement meaning). They are used to denote street addresses, room numbers, radio stations, and so forth, where only position or order matters (ordinality meaning). When they are used in fractions, they can behave differently than they do in reference to whole numbers (e.g., $1/5 < 1/3$ but $5 > 3$). And sometimes these symbols are used as names without quantitative significance, as in license plates and telephone numbers (nominal meaning). (See Fuson (1988, 1992), for an extended discussion of these and other number uses.)

In addition to providing symbols for specific numerosities, verbal numbers also are used in counting. However, counting is conceptually and developmentally distinct from labeling sets. That is, children can count "1–2–3–4–5" without realizing this is the same as determining that a set has five items. Indeed, these ideas remain disconnected for at least a year after children can produce accurate counts (Wynn, 1990, 1992). Furthermore, whereas counting leads to cardinal meaning, the relation between the numbers in a count and the specific objects to which they were applied is arbitrary. Most times, there is no reason that the second item in a count gets the label "two" except that the counter happened to tag that item second. In a subsequent count, the same item could be labeled "four" or "ten" and yet the overall count would yield the same cardinal total as before. Thus, as in other word to referent mappings, the number words are used to tag individual objects; however, in the case of counting, these local pairings are neither stable nor meaningful.

What, then, must children do to develop a concept of number? Certainly, they must come to understand each of the aspects of number outlined here. They must learn to recognize numerical symbols and apply verbal processes, such as counting, as well as sort out the various meanings and uses that these symbols take on. They must learn to recognize "threeness," "fiveness," and "twenty-fourness" in all possible instantiations. But, perhaps most importantly, they must recognize that all of these components are interrelated. Understanding "five" means knowing that all of these instantiations—five fleas, five planets, five trips, "5," "1–2–3–4–5," and "five"—are the same. To achieve this understanding, children must perform a series of mappings among many situations, skills, and inputs.

Viewed this way, number concepts should develop like other concepts (e.g., dog, blue, shiny), but with added challenges given that (a) the exemplars included in a number category can be vastly different; (b) number categories piggyback on other categories, and (c) there are unique components involved in counting and common number use. For example, to learn the concept of "dog," children have to see that different dogs are all the same because of their "dogness." However, the range of variability for a chihuahua vs. a Great Dane is much narrower than it is for five planets vs. five emotions. Like learning about numerical sets, learning about dogs involves forming an equivalence class and figuring out what to call it. However, for dogs, the equivalence class consists of individual items (i.e., dogs) whereas for number, it consists of sets (i.e., items grouped for some other purpose). Finally, for dogs, there is no analogue to learning the count word sequence. Children need not learn a sequence of animal names and, even if they did, tagging a group of animals with that sequence would not tell them whether or not they had a group of dogs. Thus, number learning likely involves similar processes but greater complexity than learning other concepts.

The present analysis illustrates the many verbal and nonverbal problems children must solve to acquire number concepts—so many, in fact, that it is unlikely they could wholly master one group (verbal or nonverbal) before learning anything about the other. The numerous layers of abstraction and meaning on both sides virtually guarantee some amount of bootstrapping. The question, then, is how much bootstrapping and at what level of detail? For individual components? Specific set sizes? Specific contexts? And at what point in development do number words and concepts begin to co-mingle? We first look to two classic positions, as well as the relevant literature specific to number development, to address these questions.

II. The Relation Between Language and Concepts

A. CLASSIC POSITIONS: A FALSE DICHOTOMY?

Discussions of language and concepts usually begin with a chicken–egg problem: which came first, the concepts or the words? On one hand is the notion that concepts emerge from some nonlinguistic origin. On the other is the claim that they are created, or at least heavily influenced, by one's native language. There are several reasons to reject this dichotomy from the outset. First, even a cursory analysis of the contents of number concepts indicates that there can be no clear leader. So many verbal and nonverbal components are acquired that one side could not completely precede the other. Second, although these extreme positions are often used to frame research, no one lays claim to either one. Despite ongoing debate about the developmental origins of number, modern investigators agree that number concepts are ultimately a mix of verbal and nonverbal components and that there is a bidirectional influence between these. Finally, presenting the issues in such gross caricature does an injustice to early theorists who, despite being strongly associated with certain ideas, also viewed these interactions as complex and bidirectional. Nonetheless, these two extreme views have framed much of the research related to number words and number concepts. For that reason, we begin with a closer examination of the distinction itself and its inherent problems.

1. Concepts Lead Language

The idea that concepts have nonlinguistic origins seems so intuitive that it can be difficult to imagine an alternative. After all, words are arbitrary symbols that derive meaning only by mapping onto a referent. The word "dog" does not mean anything until it is mapped onto an instantiation of a dog. This suggests that we must first develop some understanding of "dog" from nonlinguistic experience—perhaps guided by innate learning mechanisms or sensitivities. On this view, language is the icing on the cake—a means to communicate with others about the many ideas accumulating in our nonverbal stockpile (Fodor, 1983; Pinker, 1994).

Arguments to this effect have been made throughout the history of research on number development (Russell, 1919; Piaget, 1965; Beilin & Kagan, 1969; Beilin, 1975). For example, Piaget claimed that early number concepts emerge from the synthesis of two logical concepts, namely, class and ordinal seriation. Because early counting is initially a rote procedure, and because children fail to demonstrate logical reasoning even after they have mastered conventional counting, Piaget rejected the notion that number language contributes significantly to this development. Russell (1919) held a similar position and argued that early number instruction should be based on logical classes and not on the counting procedure. The basic idea was that because understanding number required more than counting, the origins of such concepts must be nonverbal.

This general position has been revived more recently, but ironically, it is based on evidence that numerical understandings emerge *prior* to conventional counting. For example, some have claimed that humans are endowed with a prelinguistic core of conceptual knowledge for number based on evidence that infants can detect changes in set size (Antell & Keating, 1983; Starkey & Cooper, 1980; Starkey, Spelke, & Gelman, 1990; Strauss & Curtis, 1981; Xu & Spelke, 2000) and anticipate the results of simple transformations (Simon, Hespos, & Rochat, 1995; Wynn, 1992). A weaker claim has been that certain quantitative skills can develop without mastery of conventional skills, not as part of an innate endowment, but via early experience (Huttenlocher, Jordan, & Levine, 1994; Mix, Huttenlocher, & Levine, 2002b; Saxe, 1988, 1991). Thus, the idea that numerical insight develops without mastery of conventional symbols (i.e., concepts lead language) has played a major role in theories of number development.

2. Language Leads Concepts

In contrast, other theorists have argued the opposite—that language supports the development of certain ideas and skills that might not exist otherwise. Effects of language on cognition are often discussed in terms of the Sapir–Whorf hypothesis: the idea that our concepts, and even our perceptions, can be altered by the way our native language parses the world (Whorf, 1956). Numerous studies demonstrating cross-linguistic effects on reasoning and categorization lend support to this claim (e.g., Choi & Bowerman, 1991; Levinson, 1994; Lucy, 1992). Similar evidence has been garnered in the domain of number. For example, Japanese children, whose language has an explicit base-ten structure, demonstrate a better understanding of other base-ten representations, such as place value blocks and written numerals, than their English-speaking peers (Miura & Okamoto, 1989). Here, the structure of the counting system appears to influence how children perceive other situations, such as the relations among place value blocks.

Another take on this position is that cultural tools, such as language, scaffold human thought so that new insights can be gained. For example, much of Vygotsky's work was aimed at testing whether people of different ages could use external symbols to perform cognitive tasks at increasing levels of abstraction (Vygotsky, 1962). There certainly are examples of this type of "tool use" in the literature on number concepts. From an historical perspective, the advent of improved enumeration systems has preceded new conceptual insights. For example, the shift from Roman to Arabic numerals allowed people of the middle ages to invent computations, such as long multiplication (Menninger, 1958). Similarly, some number systems better prepare children for learning computational procedures than others. Fuson and Kwon (1992) found, for example, that Korean children have an easier time learning to solve multi-digit addition and subtraction problems than English-speaking children. Here, the transparent counting system acts as a tool that children can use to gain insight into an unfamiliar computational procedure. As these examples illustrate, the claim that number words influence number concepts and mathematical thought (i.e., language leads concepts) also has a strong tradition in theories of numerical development.

3. A Case in Point

This chicken–egg dichotomy crystallized in a well-known debate on the origins of number: Principles Before vs. Principles After. The "principles" in this debate refer to the counting principles outlined by Gelman and Gallistel (1978) (see Table I). These principles were not themselves under debate—everyone agreed that they were needed for accurate and meaningful object counting (enumeration). The debate centered on when understanding of the counting principles appeared relative to acquisition of the verbal counting system.

Advocates of the principles-before view argued that children understand the counting principles before they have learned to count. This was possible because, they believed, the counting principles were embodied in both verbal and nonverbal (innate) enumeration procedures. The nonverbal counting principles were considered a skeletal structure that children fleshed out with the details of their culture-specific counting system (Gallistel & Gelman, 1992; Gelman & Meck, 1983). So, much like universal grammar is thought to aid in language acquisition, access to preverbal counting principles was thought to precede and aid learning to count.

The main evidence for this view was that children adhere to the counting principles even before they possess the skill needed to demonstrate these principles through accurate counts. For example, novice counters often use

TABLE I

Counting Principles (adapted from Gelman & Gallistel, 1978)

One-to-one principle	Every item in a display should be tagged with one and only one unique counting tag
Stable order principle	Counting tags must maintain a consistent sequence
Cardinality principle	The final tag used in a count represents the total numerosity of the set
Abstraction principle	Any combination of discrete entities can be counted
Order irrelevance principle	Items in a set may be tagged in any order as long as the other counting principles are not violated

idiosyncratic lists (e.g., 1–3–7–5) rather than the actual counting sequence. Even so, these children seem to act in accordance with the "how to count" principles. That is, those who count "1–3–7–5" use the same stable (albeit incorrect) order on every count (Gelman & Gallistel, 1978). Pre-counters also detected violations of the one-to-one and stable order principles in someone else's counting, even when the set sizes were far greater than those they could accurately count themselves (Gelman & Meck, 1983).

Proponents of the principles-after view did not believe children could access counting principles preverbally. Instead, they argued that these principles are abstracted through experience with the counting routine (Briars & Siegler, 1984; Fuson, 1988). Evidence for improvement in these skills was taken as support for the principles-after position. For example, Fuson found that early adherence to the counting principles broke down when more challenging tasks were used, such as counting large or randomly arranged sets. In contrast to Gelman and Meck's (1983) results, Briars and Siegler found an age effect for detecting counting principle violations, such that 3-year-olds were less likely to object to one-to-one and stable order errors than were 4- and 5-year-olds. Thus, children's understanding of the counting principles apparently improved as their counting skills improved. Also, other studies contradicted the principles-before findings. In particular, several investigators reported that children's own counting was actually more accurate than their error detection, indicating the reverse of the order of acquisition reported by Gelman and Meck (Baroody, 1984, 1993; Briars & Siegler, 1984; Frye *et al.*, 1989).

The Principles Before–Principles After debate is a prime example of the way language and concepts have been polarized in research on number development. But as we have seen, this dichotomy reaches far beyond this debate. Indeed, it permeates much of the research in this area. Of course, few theorists, including those cited here, have taken either position in the extreme. Still, they have been willing to argue strongly for one contribution over the other. And as with other developmental dichotomies (e.g., nature vs. nurture), the change from a categorical division to a continuum merely obscures the polarization of the two extremes. It shifts the question from "All or none?" to "Mostly this or mostly that?" In Section II.B, we consider the problems with even this type of division.

B. PROBLEMS WITH POLARIZATION

Polarizing language and concepts has limited our understanding of number development in two ways. One is a problem of definition. To show an influence of language on concepts (or vice versa), researchers must define "having language" and "having concepts." This critical task is neither straightforward nor simple. For example, which of the many verbal components of number outlined previously would qualify as the definitive test for "having number language?"

Aren't they all necessary? Yet, long before children have mastered every verbal component of number, individual verbal skills may influence conceptual growth.

Moreover, the relation between language and concepts can shift easily depending upon where investigators draw the line between having one or the other. Take, for example, Piaget's research on number conservation. Piaget defined "language" as how high a child could count, and "concepts" as the ability to judge equivalence in the face of irrelevant transformations (e.g., line length). He found that children accurately counted large sets for some time before they conserved number. In fact, he demonstrated that asking children to count the two arrays in the conservation task did not lead to improved performance.

But what if Piaget had defined "language" differently? Using counting to compare sets requires more than simply enumerating the sets accurately. One also needs to know how counting determines cardinality (i.e., knowing that a collection counted "1–2–3–4" has four items)—an understanding that is not achieved until relatively late (Wynn, 1990, 1992). There is also evidence that to conserve number, children must relate counting to ordinality (i.e., understand that $N + 1 > N$) (Baroody & White, 1983; Schaeffer, Eggleston, & Scott, 1974). These examples illustrate that by changing the definition of "having number language," a link between counting and conservation becomes more plausible.

Disagreement about how to define "counting" and "principles" also fueled the Principles Before–Principles After debate. If having principles means demonstrating any adherence to them, then young children appear to have principles. If having principles means demonstrating them under a range of complex and challenging tasks, then principles appear to develop slowly. The fact is that counting experience and counting principles cannot be completely separated—especially not using conventional counting tasks. By the time children can perform any of those tasks, they have had years of exposure to conventional counting. So, no matter how inaccurate their own counts might be, their ability to follow some procedures and detect errors may still grow out of their limited exposure to conventional counting. At the same time, conventional counting tasks do not directly test the claim that children have access to a nonverbal counting procedure. It is theoretically possible for children to quantify sets in accordance with the counting principles even before they have been exposed to conventional counting, but this possibility cannot be assessed using conventional counting tasks. In short, the effects of language appear to come and go depending on what definitions and associated measures are used.

This issue of shifting definitions is, at its core, a case of competence vs. performance—the idea that what one knows can be separated from what one does. The competence–performance distinction is what underlies the claim that one counting task is more valid than another. It is what allows us to debate which conditions reveal what children "really know" about number. And as investigators have debated the validity of different number tasks, the relation between language and concepts has been pushed back and forth. For example, researchers subsequent to Piaget argued that the number conservation task was not valid because, among other things, it used large sets. In modified tasks using smaller sets, 3- and 4-year-olds demonstrated the ability to judge the equivalence of sets despite spatial transformations (Gelman, 1972). Clearly, children at this age are less skilled counters than children who could pass Piaget's version of the conservation task; a fact that implies counting may have even less to do with conservation than previously argued. However, most preschool children have some understanding of the small number words, even though they are not proficient counters (Wynn, 1990, 1992). So if we shift the language criterion from counting sets to labeling sets, we should find another reversal—language could well lead concepts again.

If everyone could agree on the quintessential measures of both language and conceptual competence, it would be easy to test which comes first. But of course, there is no such thing as the quintessential measure of competence—there is only performance on different tasks (see Mix (2002), Sophian (1997), Thelen and Smith (1994) for discussions). Competence is inextricably connected to this performance. It cannot be separated in any meaningful sense. Thus, for the same reasons that some investigators have rejected the competence–performance distinction in general, we should question a clear unidirectional influence of either language competence or conceptual competence in number development. The appearance of such an ordering is almost certainly an artifact of the particular definitions that were used.

A second problem with polarizing these positions is that to do so requires committing to a particular time scale of analysis—a commitment that is usually not made explicitly, but yet has profound implications regarding how the interactions between language and concepts are characterized. For example, in Mix's work on numerical equivalence, a rather broad time scale was implicitly adopted. Mix and colleagues found that children could match equivalent sets (where one set was hidden) before they demonstrated the ability to count or label the same set sizes (Mix, 1999a,b, 2004a; Mix, Huttenlocher, & Levine, 1996). They then concluded that a nonverbal representation of small sets likely precedes verbal counting. The underlying assumption was that, because children could not use conventional counting to mediate their comparisons, they must use a nonverbal process instead. The time scale adopted in this research was 6-month increments—the difference between one age group tested and another. Thus, the further (implicit) assumption was that language and concepts probably did not interact until children were better counters, after which they began to recognize more abstract numerical comparisons. Although this latter claim specified a bidirectional influence between the nonverbal, high-similarity comparisons and verbal counting, this influence appeared to take place on a scale of months or even years.

However, language and concepts likely interact on a much closer time scale. For example, when a child at play hears two sets labeled with the same count word, this brief input could cause an attentional shift—one that could temporarily support a numerical comparison. Many such interactions could take place well before children can produce the labels themselves in an experimental task. Thus, what one claims about the influence of language on concepts, or lack thereof, is intimately tied to what time scale one chooses. And most existing research on number concepts uses time scales too broad to capture any subtle interplay between the two. (See Thelen and Smith (1994), for further discussion of time scales in developmental research.)

C. CURRENT CONCEPTUALIZATIONS

Given the problems with polarizing language and concepts, it is natural to ask, why can't it be both? And the answer is that, of course, it is both—a mixture of nonverbal and verbal influences that promote cognitive growth through their interactions. Current conceptualizations of the relation between number words and number concepts take this middle ground. Rather than strongly emphasizing one contribution over the other, these views describe an alternation between verbal and nonverbal influences, differing mainly in how to characterize the nonverbal component. Thus, instead of posing a chicken–egg problem, these views take more of a seesaw approach. However, because a dichotomy between verbal and nonverbal contributions remains, many of the same problems remain as well.

1. Language Transcends the Limits of Innate Knowledge Structures

In one class of current conceptualizations, investigators assume that humans are innately endowed with core knowledge systems for number. Although these systems are seen as an important foundation for mature number concepts, the argument is that they have significant limitations. Language is portrayed as the catalyst that allows young learners to transcend the limitations of their innate systems and create more powerful knowledge structures (Carey, 2001; Gelman, 1991; Spelke & Tsivkin, 2001; Wynn, 1998).

In one of the best articulated accounts, Spelke started with the assumption that human infants and many nonhuman animals possess two separate systems for representing number nonverbally (Spelke, 2003; Spelke & Tsivkin, 2001). One system represents items exactly but only works for small numbers. It uses a tracking mechanism that assigns a mental token to each object in a group. These tokens function as pointers to the objects' locations. Because there is a one-to-one relation between tokens and objects, the set of tokens can be used to represent the exact number of objects. However, only a few pointers can be active at any one time due to constraints on selective attention. Furthermore, although the representation preserves the individuality of the objects, it does not provide a representation of the whole group (i.e., in the way that a number word like "three" verbally represents a set's cardinality).

The other system represents large sets but only approximately. It is based on the accumulator mechanism, proposed by Meck and Church (1983) to explain timing and counting in rats. This mechanism works by emitting pulses of energy at a constant rate. As items are tagged, these pulses are gated into an accumulator. The resulting fullness of the accumulator, or its magnitude, provides a representation of number. However, it is inherently inexact, even for small sets, because there is not a one-to-one relation between pulses and items. Also, in contrast to the exact system, this representation does not preserve the individuality of the items, though it does represent the group as a whole. Thus, both systems have inherent limitations—the first system being limited to set sizes that the object tracking mechanism can handle (i.e., <4) and the other being limited to rough estimates. Only verbal humans, Spelke (Spelke, 2003; Spelke & Tsivkin, 2001) has asserted, can represent all set sizes exactly and they do so by counting.

An important aspect of Spelke's conceptualization is that the two core knowledge systems for number are independent of each other and highly encapsulated. That is, though they both represent an aspect of number, they do not interact so as to provide the basis for a complete number concept (i.e., the ability to represent a collection composed of individual items). This means that when children encounter a small set, they should produce two representations of it—one approximate and one exact—without seeing that these representations are related. "By our hypothesis, the child has two systems for representing arrays containing [for example] two objects... Because of the modularity of initial knowledge systems, these representations are independent. When young children hear the word *two*, therefore, they have two distinct representations to which the word could map and no expectation that the word will map to both of them" (Spelke & Tsivkin, 2001, p. 85).

The critical question in this model, then, is how children see that these systems are related and, thereby, overcome the inherent limitations of each. The answer, according to Spelke, is via acquisition of number language. Number language serves to conjoin these two representations because it provides (a) a domain general medium that allows domain specific knowledge to co-mingle and (b) a format that invites the combination of distinct concepts or systems. In Spelke's account, children make sense of counting by seeing that the same small number words map to both preverbal representations. "... because the words for small numbers map to representations in both the small-number system and the large-number system, learning these words may indicate to the child that these two sets of representations pick out a common set of entities, whose properties are the union of those picked out by each system alone" (Spelke & Tsivkin, 2001, p. 85). Having made this important inference, children are in a position to generalize to larger sets because the low end of the counting sequence (i.e., "one–two–three") is connected to the higher end that refers to larger sets.

Now, let us return to the larger question of how language interacts with concepts in terms of the "language first—concepts first" dichotomy. We see both positions represented in Spelke's account. Initially, preverbal representations precede and act as referents for the first few number words (concepts first). However, once these initial mappings have been carried out and conventional number language has been integrated with the preverbal representations, children are poised to achieve significant conceptual gains (language first). Thus, Spelke's account involves a bidirectional influence between language and concepts. Yet it achieves this in only a very rough sense.

In a similar vein, Carey (2001) distinguished between two possible relations involving number language and number concepts—a distinction that bears strong resemblance to the polarization outlined in the previous sections. The first relation, dubbed the "continuity hypothesis," holds that verbal structures are isomorphic to preverbal structures and, therefore, involve no conceptual change when they are acquired—it is a simple mapping of words to pre-existing concepts. The alternative relation, attributed to Whorf, is the idea that some preverbal concepts are incommensurable with the structure of the relevant language. These cases involve dramatic conceptual change that takes place as language is acquired. Carey argued that processes based on both relations underlie numerical development, and we can determine which situations involve which processes by evaluating whether competence is exhibited in prelinguistic babies or only in children and adults with language.

To this end, Carey (2001) outlined five aspects of number that are reflected in human language, such as singular vs. plural and the count–mass distinction, and reviewed the existing literature to determine which of the five are understood by prelinguistic infants. She concluded that even very young infants comprehend four of the five components of number (see Table II). Thus, these notions would seem to develop without language and could, therefore, serve as conceptual referents for the relevant grammatical structures as language is acquired. However, the fifth relation, integer representation, was not evident in prelinguistic infants. Carey reasoned that language input would be required for integer representation because neither of the systems that have been proposed for representing number nonverbally (i.e., the exact system for small number or the approximate system for large numbers) is structured like conventional counting. In particular, neither system can represent cardinality and thus, neither could support an easy mapping between number words and referents. In this way, Carey explained the protracted course by which children bring meaning to the number words and verbal counting.

Although these accounts vary in their treatment of certain points, they share several key assumptions: (a) number development involves a bidirectional

TABLE II

Five Aspects of Number Reflected in Natural Language (adapted from Carey, 2001)

Concept	Example	Evident in infants?
Object individuation	"I remember the toy duck that is hidden under the table"	YES
One vs. another	"I see a duck over here and a duck over there and know they are not the same duck"	YES
Count vs. mass	"Two ducks are distinct individuals, but two piles of sand are just some stuff"	YES
Sortals (nouns) vs. predicates (adjectives)	"Where the duck moves tells me it's an individual, but its color and size do not"	YES
Symbolic representation of integers (i.e., counting)	"One–two–three–four–five. Five ducks"	NO

interaction between verbal and nonverbal structures; (b) the nonverbal structures are what prelinguistic infants and nonhuman animals use; (c) these nonverbal structures play an important role, but they are limited; and (d) what helps children transcend these limitations is acquisition of number language. At some level, these seem like reasonable assumptions. However, on closer examination, we find reason to question them.

First, there is widespread agreement that number development involves an interaction between verbal and nonverbal structures. But the devil (or maybe God) is in the details—details that are left largely unspecified in these accounts. The critical turning point in each of them is when children manage to map small number words, as labels, onto their pre-existing representations for the corresponding quantities. But how, exactly, do children achieve this crucial step? Cross-sectional research indicates that this is an elusive and protracted mapping. Therefore, the specifics of how it is achieved are neither obvious nor likely to be straightforward.

This is a case where time scale may be critical. The accounts described here suggest a long period of stability, during which infants and toddlers use their innate representations, followed by an unspecified mapping process, and then another long—perhaps indefinite—period during which concepts have been transformed. Presented as such, the interactions between language and concepts resemble two large scale, unidirectional shifts that take place in sequence more than they do ongoing bidirectional, bootstrapping. Yet, as we will see, when development is studied on a different time scale, the interactions between number words and number concepts appear much more fluid and tightly linked in time than these accounts imply.

There also are problems with the way these accounts define verbal and nonverbal components. As noted previously, drawing this distinction can be complex and arbitrary. However, these accounts draw a clear line between what seemingly nonlinguistic beings (i.e., infants and animals) know and what those of us with language know. But at what point do we say that infants shift from prelinguistic to linguistic? When they have been exposed to words? When they comprehend words? When they begin to speak themselves? It seems difficult to say with certainty that any of infants' sensitivities are based on purely nonverbal information, when humans are immersed in linguistic input beginning prenatally. And what about the opposite end—the concepts of those who have acquired language? The demarcation adopted here implies that *only* prelinguistic infants possess nonverbal thoughts, even though cognition is likely a mixture of verbal and nonverbal components, even in language experts (i.e., adults). Indeed, this was one of Vygotsky's claims—that although language-mediated thought increases with development, it never fully eclipses nonverbal thought.

Finally, there is cause to question these accounts on empirical grounds. Both assume that the nonverbal foundation of early number concepts is comprised of abstract representations generated by two innate processes (i.e., the small exact number system and large approximate number system). In support of this, Spelke and Tsivkin (2001) cited evidence for dissociable enumeration systems in adults. For example, they noted that adults with acalculia, who cannot provide exact solutions to arithmetic problems, often give estimates of correct answers. This certainly suggests that humans have different systems for representing number (though this has been recognized for some time: Jensen, Reese, & Reese, 1950; Jevons, 1871; Kaufman *et al.*, 1949; Taves, 1941). And it may show that these systems are localized in different parts of the brain. However, it does not indicate that either of these systems are innate, or even early emerging.

In fact, research involving young infants has yet to clearly demonstrate any sensitivity to discrete number, with or without implicating either proposed representation (see Mix, Huttenlocher, and Levine (2002a,b), for a discussion). In brief, the evidence cited in support of numerical sensitivity in infants is undermined by confounds with non-numerical cues. For example, habituation studies showing that infants can discriminate between different set sizes (e.g., Antell & Keating, 1983; Starkey & Cooper, 1980; Starkey, Spelke, & Gelman, 1990; Xu & Spelke, 2000) are undermined because number was allowed to co-vary with contour length and/or area. When these variables have been separated, infants fail to respond to number, but continue to respond to changes in non-numerical cues (Clearfield & Mix, 1999, 2001; Feigenson, Carey, & Spelke, 2002). This casts doubt on the idea that either representation provides a nonverbal referent for the number words, because it is currently unclear when or how such representations develop.

If future research succeeds in demonstrating that these representations emerge without exposure to language, by testing very young infants under conditions that strip away or randomize every other quantitative cue, this still might not bear on what infants do in everyday situations where these other cues are available (Mix, Huttenlocher, & Levine, 2002a). This is a problem for accounts that assume children map number words onto nonverbal, abstract representations. If children use non-numerical information when it is available, then why would they map the number words to abstract representations of number and not to this information? For example, if a one-year-old hears his mother say, "two cookies," what tells him to map this phrase to a mental representation of the cookies' two locations, rather than the amount of cookie relative to the plate? For that matter, why would he map the number word to a mental representation at all, whether in terms of number or amount, when the actual objects are right in front of him?

2. Language Transcends the Limits of Experiential Knowledge Structures

An alternative to the innate knowledge models is the view that nonverbal representations of number develop in early childhood rather than comprising an innate endowment (Huttenlocher, Jordan, & Levine, 1994; Mix, Huttenlocher, & Levine, 1996, 2002a). Huttenlocher *et al.* (1994) proposed that young children develop a symbolic representation of exact number, or a mental model. In this representation, children create an array of imagined entities that stands for each of the actual entities in the real array. For example, children would represent a cookie, a brownie, and a croissant with an imagined cookie, an imagined brownie, and an imagined croissant. Huttenlocher *et al.* remained agnostic with respect to how much detail is actually preserved in these models. That is, the representations could consist of rich images of each object or they could be as sparse as a pointer. The main idea was that number is incidentally preserved because there is a one-to-one relation between the actual entities and their symbols. Huttenlocher *et al.* did not claim that the mental model was the earliest representation of quantity and allowed for the possibility that some sensitivity to quantity might be present in infancy. However, contrary to the innate knowledge views, the mental models view assumed that infants' representations are approximate—even if they are based on discrete individuals. Thus, the emergence of a mental model was considered significant because it constituted the earliest representation of *exact* number available to humans.

These claims were based on a series of experiments on young children's calculation ability. Huttenlocher *et al.* (1994) used a nonverbal task in which addition and subtraction problems were presented using concrete objects. For example, to present the problem $1 + 2$, children were shown a single block for a few seconds before it was hidden beneath a cover. Next, two more blocks were shown sliding under the cover. The child's task was to produce a set of blocks equivalent to the resulting set, even though this remained hidden from view.

Some children were able to solve such problems as early as 30 months of age. This is years before they receive instruction on the conventional algorithms for addition and subtraction, and so, not surprisingly, it is also years before they demonstrate competence on analogous problems presented in a verbal format (i.e., as word or number fact problems) (Levine, Jordan, & Huttenlocher, 1992). Thus, these authors concluded that children were using a nonverbal process to represent the entities and arrive at a solution.

These experiments also documented a shift from inexact to exact responses in the nonverbal calculation task. As noted previously, exactly correct responses were observed by 30–36 months of age. However, younger children (24- to 30-month-olds) did not perform randomly. Their responses were approximately correct. Thus, these youngsters understood that adding should result in more and subtracting should result in less, and they were able to estimate the number of items to a certain degree of accuracy, but did not reach solutions with the precision exhibited by slightly older children. Huttenlocher *et al.* (1994) argued that children who produced approximations of the correct answer did not yet possess a mental model. This shift to exact responding was considered significant because it revealed an intermediary stage between the approximate quantitative sense of infants and the advent of verbal counting. Huttenlocher *et al.* speculated that further development, including the ability to deal with larger sets, would require mastery of conventional skills.

Indirect evidence for this second shift, from limited but exact nonverbal representations to more powerful verbal representations of number, was provided in Mix's studies of numerical equivalence (Mix, 1999a,b, 2004a; Mix, Huttenlocher, & Levine, 1996). In these studies, 3- and 4-year-olds completed a forced choice matching task in which they chose a set of dots that was numerically equivalent to a standard. Across experiments, the contents of the standard sets were varied, thereby varying the degree of similarity between the two matching sets. As we discussed previously, when the standard and the choice sets were highly similar (e.g., all black dots), even children with little or no conventional counting skill matched them correctly. Although the matching sets in this condition were similar in many aspects besides number, accurate performance required numerical reasoning because the distractor sets shared the same object-based similarities as the matching sets. Also, the standard set was hidden when the choice cards were revealed; so children had to mentally represent the number of objects. That young children could do so without the counting skills needed to represent the sets verbally was further evidence that they possessed something like a mental model.

However, it appeared that this nonverbal representation was limited in that children could not perform more abstract comparisons using it alone. Only children who were proficient counters also recognized equivalence between more disparate sets (e.g., sounds and black dots). Mix and colleagues speculated that this pattern might reflect a deep relation between language and concepts—one in which the number words organize attention like other category labels (Mix, 1999a, 2004a; Mix, Huttenlocher, & Levine, 2002b; Sandhofer & Mix, 2003). For example, research outside the domain of number indicates that words alert children to possible commonalities between items and help to focus their attention on shared dimensions (Gelman & Markman, 1987; Gentner *et al.*, 1995; Sandhofer & Smith, 1999; Smith, 1993; Waxman & Markow, 1995). We have argued that number words serve the same purpose. That is, hearing two sets named with the same count word could prompt a comparison process that leads to recognition of number as a dimension of similarity. Mix, Huttenlocher, and Levine (2002b) also argued that language acts as a placeholder, or memory aid, as children acquire new skills. For example, children can use the number words to stand for the numerosity of a hidden set while they choose an equivalent set, rather than carrying out the potentially laborious process of comparing the represented (hidden) objects and visible objects one to one.

There are parallels between these accounts and the innate knowledge accounts. And these parallels lead to some of the same problems we have already outlined. The main drawback remains reaching a satisfactory separation between verbal and nonverbal contributions. In the innate knowledge models, nonverbal was defined as whatever prelinguistic infants know. In the present models, the line is drawn much closer to the acquisition of conventional skills. Therefore, what counts as verbal varies from task to task. Comparisons are made between groups of children who demonstrate mastery of the conventional skills (related to a particular task) and those who do not. When conceptual competence is revealed in children lacking these skills, it is attributed to a nonverbal process.

It is important to know which tasks require conventional skills. However, demonstrating that children do not use conventional algorithms does not necessarily mean they are using a purely nonverbal process instead. We know that by $2\frac{1}{2}$ years, the age at which children begin to solve nonverbal calculation problems exactly, they have been exposed to conventional counting for many months, have number words in their vocabularies, exhibit rudimentary counting skills, and may understand the meanings of "one" and "two" (Fuson, 1988; Gelman & Gallistel, 1978; Wynn, 1990, 1992). So, children's performance on the nonverbal calculation task could have been mediated to some extent by these words. Perhaps children represented the hidden sets with a number word or computed the solutions by counting. Huttenlocher *et al.*'s (Huttenlocher, Jordan, & Levine's, 1994) finding that children first demonstrate competence on small number problems (e.g., $1 + 1$) is consistent with the fact that children at this age may know only the meanings of "one" and "two." Maybe the shift from approximate to exact responses reflects improved number recognition or counting ability.

The same criticism could be leveled at Mix's work on numerical equivalence. The criterion for counting competence used in these studies has been the Give-a-Number task (Wynn, 1990) in which children are asked to produce sets of various numerosities (e.g., "Give me three blocks"). However, this is a high-level test of numerical understanding because it requires children to create different set sizes rather than simply recognize or label them (Benson & Baroody, 2003). This is akin to assessing children's understanding of the word "dog" by asking them to draw a picture of one. When children fail to match equivalent sets until they achieve this understanding of number words, it provides strong evidence that conventional skills are needed. However, when children match high-similarity sets *without* having passed the Give-a-Number task, it does not necessarily mean they relied on a purely nonverbal representation. They could, instead, rely on partial knowledge of the conventional number words—a level of mastery that is sufficient to support simple comparisons but not less obvious matches (i.e., those with less perceptual support).

A second problem is that the interactions between verbal and nonverbal concepts described in the experiential knowledge models are still rather broad and unidirectional. Like the innate knowledge models, these models describe a series of three, rather sweeping interactions: (a) nonverbal representations of small numbers develop; (b) number words map onto these representations; (c) children gain insights that support new concepts. These interactions are refined, somewhat, by the hypothesis that they are driven by the same well known processes that drive language acquisition and categorization in other domains. However, they are still painted in broad strokes—at a grain of detail that does not reflect the way these interactions likely unfold in real time.

3. Language and Concepts Develop Simultaneously

The models we have reviewed so far have polarized number language and number concepts by either emphasizing the importance of one over the other or describing broad, seesaw interactions between the two. We have pointed out that such polarization has inherent problems—problems that have hindered research progress on this topic. However, other models have attempted to capture the interplay of verbal and nonverbal processes at a more detailed level (Baroody, 1992; Canobi, 2004; Fuson, 1988; Rittle-Johnson & Siegler, 1998). These models provide a framework for thinking about the bootstrapping of words and concepts in real developmental time.

To illustrate, consider the iterative model (Baroody, 1992; Baroody & Ginsburg, 1986). In this view, numerical development is a gradual incremental process, in which incomplete knowledge repeatedly combines with new input to support new inferences and procedures. The model is built upon Anderson's (1984) distinction between weak and strong schemas, which held that development proceeds from disconnected, task-specific, and logically incoherent knowledge structures (weak schemas) to well-connected, highly generalizable, and logically coherent knowledge structures (strong schemas). Children are hypothesized to move along this continuum by repeatedly bootstrapping among various conventional skills and underlying concepts. The iterative model was agnostic regarding innate origins of number representation. Instead, its focus was on the developmental process that might underlie changes from infancy to school age, with or without an innate contribution.

Several key aspects of this point of view relate to the issues we have raised so far. For one, it holds that conventional skills, such as counting, can be acquired piecemeal and initially without meaning, through imitation, practice, and reinforcement. Thus, children could be exposed to number language and develop some mastery of it without completely understanding it. Importantly, however, these partial understandings were considered useful, indeed crucial, contributions to children's learning.

For example, children with partial understanding of counting might tag items with an idiosyncratic count word sequence (e.g., "five–three–two–three"). And they may not understand the implications of counting for determining cardinality, equivalence, and so forth. But as long as they know enough to say one word for each object, their counting attempts could generate important data for them. The fact that this idiosyncratic list might "fit" two different sets in terms of one–one correspondence could signal that the sets were the same. This signal could inspire further exploration of similarity between the sets. By exploring the ways in which these sets were similar, children might broaden their ideas about numerical equivalence classes. Such interactions are the essence of a genuine bidirectional influence between language and concepts—one in which number words and number concepts feed on each other at every point in development, no matter how limited and immature each side may be.

A further implication of the iterative model is that the earliest interactions between number words and concepts should be couched in specific contexts. This is because weak schemas—the starting point for development—should be task-specific and disconnected from each other. If children acquire initial skills through imitation and reinforcement, then these early contexts should follow from social routines or recurring situations in which parents have remarked on number. For example, Baroody (1992) observed that many children learn number words in the context of their age. They learn to say "two" or "three" when asked how old they are. Two- and three-year-olds certainly do not grasp the concept of years, so this is, at least initially, a meaningless mapping. But even this rote mapping could scaffold a child into deeper understanding under the right circumstances (e.g., learning to hold up the correct number of fingers at the same time, which could in turn provide a concrete referent for future mapping). These observations are significant because they demonstrate how first skills could emerge in parent-reinforced, social routines. From this, it follows that a major developmental trend will be the gradual decontextualization in which both skills and concepts become less encapsulated by these routines.

Note that this is a very different sense of decontextualization than the one discussed by Spelke (Spelke, 2003; Spelke & Tsivkin, 2001). In her view, it is the two mental representations of number that are initially encapsulated. Although these representations can operate across a variety of contexts, and in that sense already are decontextualized, they are encapsulated because they function independently, as disconnected thought processes. In simultaneous or iterative models, it is numerical competence or knowledge that is initially encapsulated because it is embedded in specific contexts and routines. Here, the process of decontextualization involves generalizing over these disparate situations.

4. Conclusions

Most current models of number development fail to capture the complex interactions between verbal and nonverbal processes. There is no definitive place to draw the line between verbal and nonverbal, so regardless of where these models have drawn it, they are probably not correct. Because they view developmental change on such long time scales, verbal and nonverbal components meet only in broad, unidirectional passes. In contrast, Baroody (1992) and others (Canobi, 2004; Rittle-Johnson & Siegler, 1998) have taken an approach that emphasizes fine-grained, bidirectional interactions. From this point of view, development involves a complex scaffolding of partial knowledge and skills—verbal and nonverbal interacting context-to-context, moment-to-moment. We do not mean to imply that other models are in opposition with these ideas. Indeed, all of the conceptualizations reviewed here are compatible with an iterative or simultaneous model. The difference is that the former, unlike the latter, have not articulated these ideas explicitly. This is not a trivial omission, because it leaves an artificially simplistic impression of the developmental process underlying these interactions.

One reason that other accounts have failed to capture the complexity of number development may be that they are based on data that tends to obscure it. Most existing research on counting, equivalence, and cardinality has used stripped-down experimental tasks presented to large groups of children in cross-sectional designs. Comparing children in 6-month increments leaves the impression that change occurs in broad strokes. When 4-year-olds succeed on a task that $3\frac{1}{2}$-year-olds fail, it seems as if there is one knowledge state at age $3\frac{1}{2}$ years and a different knowledge state 6 months later. This may be true, but children might pass through many other knowledge states along the way.

Furthermore, testing children with controlled experimental tasks virtually ensures that concepts will seem abstract because abstraction is what these decontextualized situations require. Even when tasks are designed around naturalistic situations or play activities, they still are imposed by the experimenter. Unless the experimental contexts happen to overlap completely with the contexts in which each individual child has discovered number at home, they will require a certain degree of generalization. Indeed, Mix (2002) argued that these varying degrees of overlap are at the heart of the competence–performance distinction because tasks that tap into informal knowledge at a younger age (via serendipitous overlap with contextualized knowledge) may appear more valid than those that do not. However, naturalistic tasks are not necessarily more valid than artificial tasks. In fact, artificial tasks are the best way to test whether knowledge has been abstracted. The problem is that, on its own, such evidence tells us little about how children got there. To address *this* question, we need to take a different approach—one in which we turn up the microscope, as Thelen and Smith (1994) put it, and study development at close range.

III. Number Development at Close Range

Getting close to the developmental process means moving beyond large scale, cross-sectional research to a more intense focus on the development of individual children. This requires a shift to long-term diary, case study, and microgenetic training designs. It requires one to consider performance both on standardized tasks and in children's spontaneous activity.

There is a rich history of such methods in developmental psychology, beginning with the baby biographies of the 19th century and the bedrock research of early theorists, such as Jean Piaget. More recently, these approaches have yielded fascinating insights in research on language development (e.g., Huttenlocher, Haight, Bryk, Seltzer, & Lyons, 1991; Mervis, 1985). The work of Katherine Nelson, in particular, has provided a wealth of fine-grained observations of language unfolding in the natural environment (e.g., Nelson, 1996). Her work has revealed many of the same patterns that we might expect to find for number development—early event related (contextualized) knowledge structures, idiosyncratic performance based on individual learning histories, and in-the-moment effects of language.

Studies using such methods have periodically appeared in the domain of number for many years. However, only recently has there been a more concentrated effort to track number development at close range. In this section, we review and integrate this research, with an eye toward the interaction of language and concepts in particular. These studies provide evidence for five main themes regarding numerical development: (1) it is contextualized; (2) it is piecemeal; (3) it is socially scaffolded; (4) it differs across individuals; and (5) it uses domain general processes.

A. NUMBER DEVELOPMENT IS CONTEXTUALIZED

The context-specificity of early development can be easy to miss when investigators use standardized experimental tasks, because performance on these tasks requires a certain degree of decontextualization. Children who cannot perform experimental tasks may have partial knowledge, deeply contextualized knowledge, or no knowledge at all. To determine the status of these children, we need to look at the behaviors they generate themselves.

This was the aim of a diary study that Mix (2002, 2004b) completed with her son, Spencer. Observations were recorded from the time Spencer was 12 months old until just after his third birthday. They focused on two main activities: (a) use of one–one correspondence in spontaneous play, and (b) attempts to count and use number words to label sets. In addition to the diary observations, Spencer was tested using standardized numerical equivalence and counting tasks on a monthly basis, beginning at 20 months of age. Competence in both nonverbal and verbal aspects of number was evident very early in Spencer's development. However, in both cases, this competence was highly contextualized.

This contextualization was evident in his understanding of numerical equivalence. For many months early in the study, he experienced one-to-one correspondence in his play—handing out toys to other children, touching objects one by one, or taking turns in simple games and routines. These activities did not require explicit understanding of numerical equivalence, but likely provided important data for the development of such an understanding. At 20 months, he spontaneously produced a set of objects that was equal in number to another set hidden from view. Specifically, he took exactly two dog treats for his two dogs, waiting in another room. This was not a coincidence. Over 3 weeks, he repeatedly performed the same task with almost perfect accuracy using different-sized treats. Yet, he was unable to perform the same task when the context changed. He failed Mix's (1999a,b) forced choice matching task—even for the numerosity "two." And, he failed on a slight variation on the dog treat task, one in which he was asked to give his toy trains "train treats." Spencer did not explicitly match sets in any other situations until he was 30 months old, when he went into the backyard and returned with three sticks, one for each person sitting at the dinner table (i.e., his two parents and a dinner guest). Like children in previous cross-sectional studies, Spencer began matching sets in the forced choice equivalence task starting at 34 months, with high-similarity (disks-to-dots) comparisons first. Unfortunately, the diary study concluded soon after that, so we do not know whether spontaneous matches increased around the same time that his notions of numerical equivalence became decontextualized.

During the same developmental period, Spencer's acquisition of small number words was similarly contextualized (Mix, 2004b). His earliest uses of the number words did not actually reference sets of objects. Instead, he used number words to label written numerals. This began with the numerals that appeared in several of his board books, but he eventually came to recognize numerals on signs, license plates, and addresses as well. At 23 months, he began using number words to label sets of objects. His first mappings were restricted to the number "two" and they always occurred within a particular linguistic frame: "Two ____, One. Two." For about a week, he labeled only sets of shoes using this frame (i.e., "Two shoes. One. Two."). Then he extended to other object sets, including two dogs, two spoons, and two straws, using the same frame. At 24 months, he began using the variant "____, ____, two ____." For example, for two trains, he would say, "Train. Train. Two trains." This frame appeared frequently for the next 6 weeks and, during this period, he did not label sets numerically without using it. Fuson (1988) reported that her daughter Adrienne used the same frame at age 22 months.

Throughout this period, Spencer failed all tests of conventional counting. In the Give-a-Number test, he failed to produce two objects on request and when asked how many objects were in a set of two, he responded with an idiosyncratic string of number words. Thus, although he correctly labeled different sets of two, his use of the number word "two" was far from decontextualized. In fact, it was deeply contextualized in two ways. First, it was initially restricted to specific situations—first labeling numerals, then labeling shoes. Second, these early attempts were embedded in specific linguistic frames. A similar pattern was reported in another diary study that tracked the number development of a young boy, Blake, from 18 to 49 months of age (Benson & Baroody, 2002). Blake's first number word also was "two" and he used it, initially, only when asked his age. His parents had reinforced this response in preparation for his upcoming birthday. Although this was a simple association without cardinal meaning, it is noteworthy that his first use of a number word occurred in this, and only this situation.

Both boys also overgeneralized the response "two" to the question "How many?" That is, regardless of a set's numerosity, when asked informally how many items were there, both boys tended to say, "two." This was true even though both boys spontaneously labeled sets of one and two correctly. The following excerpt from Spencer's diary (Mix, 2004b) illustrates this tension:

> (5/24/01; 26 months) While Spencer was taking a bath, I threw in one toy fish. Then I added two frogs, one by one. Spencer remarked, "Oh! Two frogs!" Then I threw in a third frog. Spencer said, "Oh! One frog and two frogs." I asked, "How many frogs all together?" He responded, "Two frogs." I replied, "No, three frogs."
> A little later, Spencer spontaneously asserted, "Two frogs." I replied, "No, three frogs." He countered, "No, one frog, one frog, one frog, and one fish." I asked, "How many is that?" His response: "Fish need water."

Although Spencer correctly labeled sets of one and two, he insisted that there were two frogs when queried. The fact that both he and Blake spontaneously generated correct labels for one and two suggests that these overgeneralization errors were specifically embedded in the "How many?" routine. Perhaps Spencer and Blake viewed the question "How many?" as part of a script where the other person always answers, "two," presumably because they had been reinforced for this response when it was correct.

B. NUMBER DEVELOPMENT IS PIECEMEAL

One of the most striking patterns to emerge from diary and longitudinal research is that acquisition of the number words has a distinct stepwise or piecemeal quality (Benson & Baroody, 2002; Fuson, 1988; Mix, 2004b; Sandhofer & Mix, 2003; Wagner & Walters, 1982; Wynn, 1992). For example, when Wynn (1992) tracked children's development using a variety of standardized tasks, she found that the number words were acquired in discrete stages. In the first stage, children can give one object on request and also identify which of two cards shows only one object (mean age = 33 months). Children at this level did not point to a single object when queried with other number words. For example, given a card with one fish vs. a card with two fish, these children never pointed to the card with one fish when asked to point to two. In fact, they consistently inferred that the "not one" option (i.e., the card with multiple items) was the referent of words other than "one." However, this did not reflect knowledge about the specific cardinal meanings of these words, because the same children performed randomly when both cards depicted multiple items (e.g., two vs. three). Thus, Wynn concluded that children first discover the difference between "one" and "more than one." In the second stage (mean age = 35–37 months), children correctly produced and identified sets of one and two, but failed to distinguish among larger numerosities. Within a few more months (mean age = 38–40 months), children correctly identified sets of three as well. Finally, children were able to produce and identify all the sets in their counting range (i.e., four and greater) at about the same time (mean age = 42 months).

Wynn's (1992) study indicates that number words are acquired in a piecemeal fashion. However, as we have already discussed, this development begins much earlier than the ages tapped by her tasks. Diary research demonstrates that children start sorting out the meanings of "one" and "two" in specific contexts around their second birthdays or even earlier (Benson & Baroody, 2002; Fuson, 1988; Mix, 2004b). This is 9 months before children demonstrated an understanding of "one" and 13 months before they demonstrated an understanding of "two" in Wynn's research. Moreover, the diary studies suggest that rather than leading with the number "one," it is the number "two" that may have special status in very early development. All three studies reported that functional use of the word "two" preceded functional use of "one." Also, Spencer labeled sets of two much more frequently (five times more) than sets of one (Mix, 2004b). Finally, whereas these children usually labeled both "one" and "two" accurately, they tended to overgeneralize the label "two" when queried by adults.

Labels for larger sets (i.e., three and four) appeared in children's spontaneous utterances several months after the appearance of "one" and "two" (Benson & Baroody, 2002; Mix, 2004b). Although some overgeneralization of the word "three" was observed, children were generally accurate from early on (29 months of age in Spencer's case). Perfect functional use of the word "three" was evident before perfect use of the word, "four," but these milestones occurred relatively close in time (within weeks for Spencer). Thus, as Wynn observed, there was a stagelike progression in the emergence of informal mappings for the small number words. However, in contrast to Wynn's findings, the order of emergence was different (i.e., two appeared before one), the separation of the stages was less apparent, especially for one and two, and the ages of acquisition were very discrepant. Specifically, by roughly the same age as the children Wynn (1992) classified as Stage I (i.e., those who demonstrated only an understanding of one vs. more than one), Blake had already demonstrated highly accurate use of both "one" and "two," and Spencer had done so for all the small number words, "one" through "three."

Although it is not entirely clear why children would repeat the same sequence in Wynn's (1992) standardized tasks that they appear to pass through in informal labeling, it seems certain that the standardized tasks did not provide the same scaffolding that children have when they label sets themselves. In other words, we might ask why certain everyday situations seem to draw correct numerical labeling out of children earlier than the Give-a-Number or Point-to-X tasks used by Wynn. The diary data suggest that labeling pairs of observable objects is what gets the ball rolling. As we noted, children label sets of two earliest, most frequently, and with great accuracy. Furthermore, the vast majority of sets they label consist of observable objects (rather than mentally represented or remembered objects) (Mix, 2004b). So, what makes this situation special?

Mix (2004b) speculated that the answer may be a coincidence between limits on children's comparison ability and an initial misinterpretation of the word "two." In the following excerpts, Spencer incorrectly extended his number frame for "two" to larger sets. Such errors were rare, so it was not the case that he routinely mislabeled larger sets "two," as if he took the word to mean "many." Instead, the following anecdotes suggest that for Spencer, the word "two" had more to do with the similarity among items in a set than it did with their cardinal number:

2/10/01 We had a carpet sample board with about 20 carpet squares. Spencer remarked, "Blue!" Then, he slapped the board 5 times, contacting 8 squares, while saying "A blue" with each slap. Then he said, "Two blues."

2/24/01 Spencer pointed at each of the three living room windows and said, "Window, window, window. Two windows."

3/8/01 Spencer saw three apples on the dining room table and said, "Apple. Apple. Apple. Two apples."

Clearly, Spencer was aware of the similarity among items in these sets and acknowledged it, both by tagging each item through touch or pointing, and by labeling each item with the same name. The fact that he summed up these comments with the number word, "two," regardless of the set's actual size, suggests that he misinterpreted "two" to mean "same." This would be a sensible error, given that number words apply to groups of things that share some commonality. Furthermore, there is little in the input to indicate that "two" or any other number word refers to number in particular, rather than something like "again," "another," or "same."

But why use only "two" this way, and not the rest of the number words? One reason might be the relative ease of comparing two items. If it is easier to determine that two objects are the same than it is to evaluate the similarity of three or more items, then a child would be more likely to comment on similarity for sets of two. Once children can make more complex comparisons (i.e., when they begin to see similarity for larger sets), they may refer to this similarity as "two," simply because "same thing" is already a major part of what "two" means to them. Exposure to the homonym "too" might further reinforce this misinterpretation. Young children have no way to know that "two" and "too" are different words. Sentences such as, "Mary wants a cookie, too," might provide additional (erroneous) evidence that "two" means something like "another."

Because young children may conflate "two" and "same," it is difficult to say whether their early uses of "two" refer to cardinal number at all. Although Spencer and Blake used the words "two" and "one" discriminately, this discrimination could have been based on the need to comment on similarity, or lack thereof. That is, there would be no reason to say "same" for a single object. Perhaps that is why children fail to demonstrate an understanding of "two" in Wynn's (1992) tasks even though they use it with great accuracy in these informal labeling situations. Clear evidence for the separation of cardinality and similarity would require discriminate use of different words for multiples, such as "two" and "three"—a development that seems to take an additional 6 months to achieve informally, yet still appears to precede Wynn's Stage II by a considerable margin.

The discrepancy between Wynn's (1992) findings and the diary results is a prime example of why verbal and nonverbal change cannot be separated developmentally. Although it could be argued that children do not really understand "two" until they can perform Wynn's tasks, and in that sense are still nonverbal with respect to number, it is equally true that they do not develop these understandings in a verbal vacuum. Children are clearly experimenting with the number words, usually with success, for many months before they appear to understand them in more structured tasks. This means that the “nonverbal” referents for specific numerosities may well be shaped or even created by exposure to the number words. These partial mappings, in turn, are likely to scaffold deeper understanding of the number words themselves.

During the same period that children acquire the pieces of verbal cardinality, notions of numerical equivalence also emerge in a stepwise manner. As noted previously, cross-sectional research has indicated that the ability to match equivalent sets begins with high-similarity matches between 30 and 36 months, and gradually extends to more disparate comparisons (Huttenlocher, Jordan, & Levine, 1994; Mix, 1999a,b, 2004a). Children cannot completely ignore object similarity in favor of numerical equivalence for another two years, until 5 years of age (Mix, 2004a). This pattern was also revealed in a longitudinal study of children’s number development (Sandhofer & Mix, 2003). Children were tested once a month from 36 to 54 months of age. They completed several versions of the forced choice number matching task that ranged from high-similarity comparisons (i.e., black disks to black dots) to low similarity comparisons (e.g., puppet jumps to black dots). As in previous cross-sectional research, children did not succeed on the full range of comparisons all at once. Instead, success in the high-similarity conditions always preceded success in the low-similarity conditions. There also were effects of set size. Children matched equivalent sets that were small (1–4) over a year before they matched larger sets (5–8). These patterns were replicated in Spencer’s diary data as well (Mix, 2004b).

Taken together, longitudinal and diary research indicates that partial number competence emerges long before reliable performance on experimental tasks (Baroody, Benson, & Lai, 2003; Benson & Baroody, 2002; Mix, 2002). During this period of growing competence, children gradually gather pieces of both verbal and nonverbal understanding. Although we can attempt to study these components separately using verbal or nonverbal tasks, aspects of the two are always present, tightly intertwined in developmental time.

But does this mean they are integrated in the child’s mind? The answer depends on what is meant by “integrated.” The complete integration of number words, counting, and all the possible instantiations of numerosity is the culmination of numerical development—an apex that is not achieved for several years. Yet, there is likely an ongoing bidirectional influence of partial knowledge across the verbal–nonverbal divide well before this achievement. Are these piecemeal interactions integrated from the child’s perspective? Perhaps, but only within specific contexts. For example, saying “Two shoes. One. Two,” implies an integration of counting and cardinality. Though not the same as having decontextualized, principled knowledge of this relation, this may reflect an explicit integration within that context, just the same. Alternatively, interactions involving partial understanding may not reflect explicit integration, even though they could be a large part of what bootstraps children into such levels of understanding.

C. NUMBER DEVELOPMENT IS SOCIALLY SCAFFOLDED

In our discussion of contextualization, we pointed out the extent to which children's early number knowledge is tied to specific situations. We speculated that one reason competence emerged in these contexts and not others was the high similarity between items in those sets. However, another quality shared by many of these situations was a high degree of social scaffolding and reinforcement. In fact, naturalistic observations have revealed time and again that numerical understandings emerge first in social games and routines (Benson & Baroody, 2002; Mix, 2002, 2004b)—a basic fact overlooked by research using standardized tasks. For example, when Spencer succeeded in matching treats to dogs, he was likely imitating a routine he had seen his parents perform every day. This matching activity also was an extension of a socially mediated one-to-one correspondence activity he had been spontaneously performing for months—namely, distributing objects (Mix, 2002). Rather than aligning objects, as in matching teacups to saucers, Spencer's most frequent one–one activity was handing out objects to people, dolls, or animals. Mix speculated that Spencer was reinforced for this activity by the attention and positive social interactions he experienced as the distributor.

Spencer's number word learning also occurred within various social routines (Mix, 2004b). For example, a series of conversations about number arose within the daily routine of taking vitamins. Spencer liked the taste of his chewable vitamins and would have eaten more, but he was only allowed to have one per day. Consider the following excerpts:

> (3/21/01; 26 months) As Spencer was feeding me toast, he said, "Just one, Mommy,"
> —the same thing I say when I give him his vitamin.
> (3/24/01; 26 months)
> KSM: Do you want a vitamin?
> S: Yeah
> KSM: Okay, here you go. Just one vitamin.
> S: No! *Two* vitamins (smiling).
> (3/30/01; 26 months) When offered a choice of two vitamins, Spencer tried to take them both. But he smiled and said, "Two. *Two* vitamins."
>
> (4/6/01; 26 months)
> KSM: Would you like a vitamin?
> S: Just one!
> KSM: Right! Just one vitamin a day.

S: Just *two* (smiling).
KSM: No, just one.
S: No, just two (smiling).

These excerpts illustrate how Spencer's early comments on cardinality were couched in familiar, recurring contexts that highlighted number. This context was not supportive simply because he encountered the same objects. It was a social situation for which talking about number served a function—it might get him another vitamin and, failing that, it might make his mother laugh.

Similar contextual support was provided in Spencer's acquisition of the counting sequence. Starting at 23 months of age, he frequently participated in a counting, turn-taking game with his father. For example, if his father said, "One," Spencer would respond "Two." His father would then say, "Three," and Spencer would reply, "Four." They would continue this way as high as Spencer could count. These episodes provided a considerable amount of the overt practice Spencer had with the counting sequence—practice he could not have had without a partner.

Social routines such as these invite observational learning and imitation—processes that could start the ball rolling in number development. Such processes were evident in the vitamin excerpts, presented previously. Spencer's first comment on cardinality in this context was an exact mimicry of what had been said to him, morning after morning, for months (i.e., "Just one."). The frames he used to comment on number (e.g., "Two ____. One. Two") also started out as imitations. His babysitter had used the counting frame for weeks to label sets around the house, including sets of shoes. So, Spencer's use of this frame to label sets of shoes was no accident. It was a direct imitation of what he had heard said in the same context. He learned that you announce "Two shoes. One. Two," in the presence of shoes just like he learned that when you see a phone, you put it to your ear and start talking.

One final aspect of social scaffolding for number bears comment. Empiricist accounts of development are sometimes criticized because they seem to require effortful instruction and reinforcement on the part of the "teacher." However, potent forms of reinforcement arise in social interactions without requiring planned rewards or punishments. Many of Spencer's comments seemed to be trial balloons that he sent up to see what ideas would be accepted. An example was when he insisted that there were two frogs in the bathtub. He did not need concrete punishment, like a rat requires a shock, to know whether his thinking was on track. The correction that followed was feedback enough. Similarly, the counting game he played with his father did not originate in an explicit attempt to teach Spencer the counting sequence. It was a mutually enjoyable social activity that the two of them invented together and rewarded each other for continuing.

Blake's diary (Benson & Baroody, 2002) provides an excellent example of yet another form of positive reinforcement—the reward that comes from making oneself understood. When he was 27 months old, his mother asked him whether he wanted to drink milk or water. He first replied, "milk," but then said, "water." Unsure how to respond, his mother asked, "Which do you want, milk or water?" Blake replied, "Two," indicating that he wanted both. His use of the number word was rewarded when he received the drinks he had requested. No further feedback or effortful instruction was required—the functionality of that word in that context was sufficient to create the behavior.

D. NUMBER DEVELOPMENT DIFFERS ACROSS INDIVIDUALS

Close range examinations of numerical development have revealed a fourth trend—different children develop the same understandings in different ways. In a longitudinal study of counting and numerical equivalence concepts, we identified two different patterns of interaction between verbal and nonverbal competence (Sandhofer & Mix, 2003). For one group of children, development was seemingly led by verbal skills (see Figure 1). These children demonstrated counting proficiency earlier than the rest of the group, with most of them accurately producing all of the small sets (2, 3, and 4) on request, at or near the beginning of the study. Even more striking was that they did so before they could match

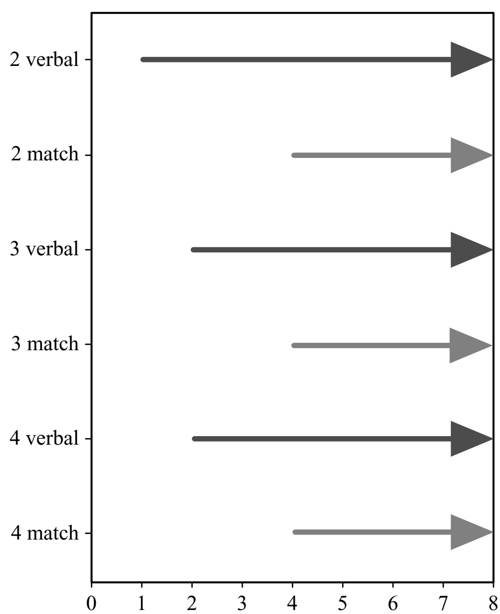

Fig. 1. Average session by which All Verbal children attained above chance performance in Sandhofer and Mix's (Sandhofer & Mix, 2003) longitudinal study of counting and numerical equivalence.

equivalent sets for *any* of the same set sizes. When they finally did match equivalent sets, an average of 2 months later, they did so for all the small set sizes at once. Thus, for these children, number development involved ascribing meaning to the small number words without forming categories for them, and then forming these categories all at once.

For a second group of children, the emergence of verbal and nonverbal skills was interleaved (see Figure 2). These children reached proficient levels for both verbal and nonverbal tasks one numerosity at a time, over a period of about 6 months. This pattern suggested that children worked out the meaning of each number word, including its corresponding equivalence class, before moving on to the next—a very different course than that obtained for children who focused on verbal skills first.

We speculated that these different patterns reflected differences in the learning histories of each child (Sandhofer & Mix, 2003). Both of our experimental tasks measured children's reasoning at a high level of abstraction. In the verbal task, children produced sets of blocks on request. In the matching task, they identified equivalent sets that were otherwise quite disparate. Verbal competence led this form of nonverbal competence in both groups, indicating that verbal skills were abstracted first and then used to abstract the children's contextualized notions of

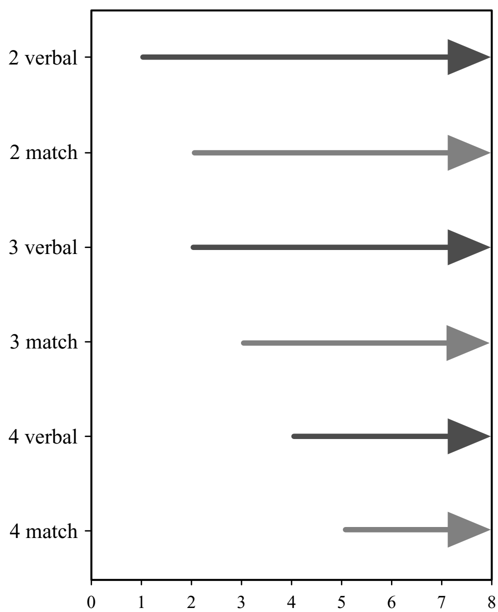

Fig. 2. Average session by which Number-by-Number children attained above chance performance in Sandhofer and Mix's (Sandhofer & Mix, 2003) longitudinal study of counting and numerical equivalence.

equivalence. The interesting difference between the two groups was that the Number-by-Number children (i.e., the second pattern) seemed to have a weak sense of number categories waiting in the wings—presumably constructed through experience with object sets and still somewhat contextualized. So, although children with this learning history were slower to abstract the verbal labels, each time they did, they were immediately able to abstract the corresponding number category as well. The All Verbal children (the first pattern) seemed to lack these weak number categories. Perhaps most of their number input had been focused on counting and number words, rather than one-to-one correspondence or play with matching sets.

E. NUMBER DEVELOPMENT IS DOMAIN GENERAL

We began this chapter by analyzing what number concepts entail, including both verbal and nonverbal components. Viewed this way, a main challenge to young learners is integrating these components—mapping one to another in a complex web of skills, situations, and ideas. Like any other mapping, numerical mappings are likely to involve noticing similarity, isolating relevant points of alignment, forming categories, and pairing words with referents. In other words, there is no reason to think that numerical development should be special, except that it may be especially difficult given the number of mappings required and the lack of obvious cues. Mix (Mix, 1999a,b, 2004a) has argued previously that domain general processes of comparison underlie the development of numerical equivalence judgments. In this section, we consider whether they also underlie verbal mappings.

The diary and longitudinal studies we have described so far provide several indications that they do. First, several trends observed in number development resemble those seen in the development of other concepts, such as color. For example, recall Wynn's (1992) finding that children realize the number words refer to numerosity before they know the specific cardinal meanings of these words. The same pattern is evident in children learning color terms. That is, they first realize that the color words as a group refer to the dimension of color (Backscheider & Shatz, 1993; Landau & Gleitman, 1985; Sandhofer & Smith, 1999). At 27 months, children asked, "what color?" respond with a color word, albeit, usually the wrong one. Within a few months, they begin to provide specific words for specific colors. For example children label red apples as "red" and blue balls as "blue," but may incorrectly label yellow balls as "purple."

Another domain general trend in number development is that local mappings precede the formation of equivalence classes. Recall that both Blake and Spencer spontaneously labeled various object sets containing the same number of items for many months before they could match equivalent sets in a forced choice task (Benson & Baroody, 2002; Mix, 2004b). The same pattern has been observed in children learning color terms. For example, Smith (1984) reported that although 2-year-olds correctly labeled objects in terms of color, they were unable to match objects by color when the objects differed on other dimensions, such as size.

This might seem counterintuitive. After all, doesn't correctly labeling a flower, a car, and a drink with the word, "red," imply that you know that these objects are of the same color? Not necessarily. Further, research using a connectionist model demonstrated that these two senses of same are quite distinct. Smith, Gasser, and Sandhofer (1997) trained a network to label three properties of a given input. For example, given a smooth red triangle, and asked, "What color is it?" the network learned to respond "red" and when asked "What shape is it?" the network learned to respond, "triangle." However, even after learning to label objects by color the network failed to represent objects that were the same on a given property as equivalent. That is, when the network was asked, "What color is it?" and was presented with a smooth red triangle, the pattern of activation on the hidden layer was different than when the network was presented with a bumpy red square and asked "What color is it?" The network apparently failed to isolate the property of color and continued to represent aspects of the shape and texture of the objects even though these were irrelevant to the task at hand.

The order of the mappings children perform for number also reveals some interesting domain general connections. First, like number categories, color equivalence classes are affected by the degree of similarity between objects. When the target and choice objects are highly similar, for example a red airplane and a similar red airplane, even children who do not comprehend color terms can match these objects by color (Soja, 1994). When the target and choice objects are less similar or when there is competing similarity from a distractor object, children fail to match objects by color until long after they have learned to comprehend and produce color terms correctly (Rice, 1980; Sandhofer & Smith, 1999; Smith, 1984). This is precisely the same pattern we described previously for numerical equivalence judgments, number words, and object similarity.

A second ordering of interest involves first uses of the number words. Both Spencer and Blake mapped the number words onto written numerals early in development. In fact, these constituted Spencer's first number word mappings. This makes sense given that children tend to interpret new words in terms of shape as their vocabularies increase (Smith *et al.*, 2002). Indeed, children with a strong shape bias can identify more letters of the alphabet than children who lack the shape bias, presumably because learning letter names requires careful attention to shape (Longfield, 2004). When children map number words to written numerals, they may be extending the shape bias to numbers. After all, numerals have a consistent shape. Numerically equivalent sets do not.

Finally, Spencer's use of number frames is reminiscent of children's use of pivot grammar more generally. Bloom (1993) noted that children often use the same simple sentence structures to incorporate new vocabulary. For example, they might learn the frame "Give ____" to request items and then use this frame repeatedly as they acquire new words (e.g., "Give milk," "Give toy," "Give cookie," etc.). Spencer's number frames have much the same quality. They provided a way for him to incorporate new sets into his category of "twoness."

Thus, we see many parallels between number development and development in other domains. In particular, number word learning looks quite a lot like other word learning—it starts out with loose associations between the words as a group and the broad dimension they describe; it involves overgeneralization (i.e., initially referring to many numerosities as "two"), an initial bias toward shape, and the transient use of pivot grammars; and it is built from local mappings without reference to a larger equivalence class.

The significance of these parallels is that they indicate common underlying processes. When children associate number words with the dimension of number, they are likely responding to patterns in linguistic input as they do when learning other words (Bloom & Wynn, 1997). When children map number words to shape, they are using the same strategy that works in other word learning situations (e.g., Landau, Smith, & Jones, 1988; Smith, Jones, & Landau, 1992). When they overgeneralize, they are struggling to reconcile their understanding of the underlying categories with the socially accepted categories to which words refer (e.g., Mervis, 1985). These parallels provide important insights into the mechanisms by which children integrate number language with conceptual understanding—mechanisms that we consider in Section IV of this chapter.

IV. Toward a Mechanistic Account

All current conceptualizations of numerical development hold that there is a bidirectional influence between number words and number concepts. Even those models that assume a considerable innate component concede that the mapping between verbal and nonverbal knowledge precipitates significant conceptual growth. These accounts contribute by speculating about possible fit (or lack thereof) between nonverbal and verbal representations. However, the claim itself—that verbal number maps onto nonverbal number and leads to conceptual change—is not an advance. It neither distinguishes current conceptualizations from those that came before nor provides insight into the details of how these interactions occur. In fact, these accounts may mislead by leaving the false impression that mapping number words onto number concepts is more discrete and unidirectional than it actually is.

When we look at number development close up, there are no clean, wholesale mappings from skill to understanding, from word to concept. Instead, we uncover a multitude of disconnected, local mappings, successfully achieved with a great deal of contextual and social support, gradually coalescing into a fully integrated conceptual structure. This process can seem messy. It may be tempting to gloss over the details for the sake of theoretical clarity. However, it is from these details that we can see traces of the learning mechanisms that underlie the achievement of mature number concepts. In order to explain how number development unfolds, it will be necessary to embrace this complexity and the mechanisms it reveals.

The mechanisms revealed so far are not new and they are not specific to number. Indeed, they take us right back to the classics. For example, it is difficult to think of a better way to characterize early number development than that provided by the Vygotskian framework. In this view, learning proceeds through successive stages of socialization. Children first imitate routines that have no meaning to the child beyond their social function. Over time, children internalize these routines, and the associated language, until words, context, and concepts become inextricably merged. As children assimilate verbal procedures into their thinking, they gain access to more powerful conceptual structures that allow them to evaluate or invent new procedures. Even so, they never completely abandon nonverbal thought. Our review of early number development, particularly the diary, case, and microgenetic studies, provided many examples of this type of learning. If we want to know specifically how children make initial mappings among the verbal and nonverbal components of number, Vygotsky's ideas provide a very good start.

There is also abundant evidence of empiricist learning processes. Children often started out imitating what they had observed in specific number-relevant situations. They initially mapped words onto object sets that they could perceive—not necessarily to representations of those sets. Their attempts to generalize beyond these situations were shaped by social approval and successful communication. Individual differences in number development suggest variations in the specific patterns of input and interactions in children's learning histories. Although this may not be all there is to number development, observational learning, associative learning, and conditioning are obviously a significant part of it. Indeed, the burden facing those who would argue for domain-specific learning is to explain why empiricist processes and social scaffolding are not sufficient on their own.

Of course, simply establishing that these processes underlie development is only a start. Much remains to be learned about the specific ways they are implemented in number learning. For example, if children hook into number words by imitating social scripts, then the next step is to analyze these scripts more closely. One basic question is why children enact some scripts and not others. By isolating and comparing the situations that are numerically meaningful to children, we can determine which situational cues direct attention toward number in particular. These might include specific spatial or temporal relations, linguistic cues, or social referencing cues. Such an analysis would also indicate whether non-numerical understandings, such as recognizing similarity, or ordering and grouping objects, help to scaffold numerical insights because different contexts will vary in the degree to which this information is either provided or required.

Another line of inquiry could focus on consistency across individuals. That is, do most children make their first number word mappings in similar contexts? Do they often name numerals, as Spencer did? Or do they first map number words to fingers, held up to represent their age, like Blake? If so, then it will be important to look especially closely at these contexts to explain their widespread use. Such universal appeal would indicate either considerable emphasis or repetition in the environment, a particularly good match to the child's cognitive capacities, or both. Perhaps there are several classes of situations different children use initially. If so, it would be possible to trace the origins of multiple pathways, similar to those we described for the case of cardinality and equivalence (Sandhofer & Mix, 2003).

Finally, researchers might ask about the structure and content of parent input. The notion that children first bring meaning to number words through imitation presupposes a major role for parents because it is they who provide social routines to imitate. If nothing else, their input sets limits on the universe of possible situations children can access. However, parent input likely makes a more profound contribution by structuring children's learning environments and directing their attention within them. For some routines, such as holding fingers up for age, parents probably teach children explicitly as part of a larger social context (i.e., occasions when new acquaintances will ask how old they are). Other routines may evolve from different social needs, such as sharing, or simply emerge as the parent comments on the environment (e.g., "Oh look! Ducks! Two ducks!"). Because these routines are pivotal in children's numerical development, it is important to know which routines parents present to children, what might lead to the creation of these routines (i.e., why these routines and not others?), and which of these routines children adopt themselves. This will tell us not only how the child's environment is structured, but also something about why it is structured that way.

It has been argued that adult intelligence is based not only on the brain but also on the environment in which the brain operates (e.g., Clark, 1997). The development of intelligence in children can be viewed the same way—as the emergence of increasingly smart, adaptive behaviors within an indivisible system of neural processes and environmental structure (Thelen & Smith, 1994). The iterative view embraces this model of conceptual change, compelling us to look beyond the boundaries of the child to explain how number concepts develop and focus instead on the close-knit interactions among numerical understanding, language, and social activity. Understanding these interactions will be time-consuming and complicated. It will require more than the typical, cross-sectional laboratory experiment, forcing us to find new and creative approaches. But the reward will be a deeper understanding of the way number words and number concepts really interact, moment-to-moment, in all their unruly complexity.

ACKNOWLEDGEMENTS

Preparation of this chapter was supported by grants from the Spencer Foundation (#200400033) and the National Science Foundation (BCS-0111829). Many thanks to Brian Bowdle for invaluable comments that helped shape this work in its earliest stages. We also are deeply indebted to Susan Jones, Linda Smith, and Esther Thelen for their inspiration and support, as well as helpful comments on the current work. Correspondence concerning this chapter should be addressed to Kelly S. Mix, Department of Psychology, Indiana University, 1101 E. 10th St, Bloomington, IN, 47405. Electronic mail may be sent to kmix@indiana.edu.

REFERENCES

Anderson, R. C. (1984). Some reflections on the acquisition of knowledge. *Educational Researcher*, *13*, 5–10.

Antell, S., & Keating, D. P. (1983). Perception of numerical invariance in neonates. *Child Development*, *54*, 695–701.

Backscheider, A. G., & Shatz, M. (1993). *Children's acquisition of the lexical domain of color.* 29th Annual Proceedings of the Chicago Linguistic Society, Chicago, IL. April.

Baroody, A. J. (1984). More precisely defining and measuring the order–irrelevance principle. *Journal of Experimental Child Psychology*, *38*, 33–41.

Baroody, A. J. (1992). The development of preschoolers' counting skills and principles. In J. Bideaud, C. Meljac, & J. P. Fischer (Eds.), *Pathways to number: Children's developing numerical abilities* (pp. 99–126). Hillsdale, NJ: Erlbaum.

Baroody, A. J. (1993). The relationship between the order–irrelevance principle and counting skill. *Journal for Research in Mathematics Education*, *24*, 415–427.

Baroody, A. J., Benson, A. P., & Lai, M. L. (2003). *Early number and arithmetic sense: A summary of three studies.* Paper Presented at the Biennial Meeting of the Society for Research in Child Development, Tampa, FL. April.

Baroody, A. J., & Ginsburg, H. P. (1986). The relationship between initial meaningful and mechanical knowledge of number. In J. Hiebert (Ed.), *Conceptual and procedural knowledge: The case of mathematics* (pp. 75–112). Hillsdale, NJ: Erlbaum.

Baroody, A. J., & White, M. S. (1983). The development of counting skills and conservation. *Child Study Journal*, *13*, 95–105.

Beilin, H. (1975). *Studies in the cognitive basis of language development*. New York: Academic Press.

Beilin, H., & Kagan, J. (1969). Pluralization rules and the conceptualization of number. *Developmental Psychology*, *1*, 697–706.

Benson, A. P., & Baroody, A. J. (2002). *The case of Blake: Number word and number development.* Paper Presented at the Annual Meeting of the American Educational Research Association, New Orleans. April.

Benson, A. P., & Baroody, A. J. (2003). *Where does nonverbal production fit in the emergence of children's mental models?* Paper Presented at the Annual Meeting of the Society for Research in Child Development, Tampa, FL. April.

Bloom, L. (1993). *The transition from infancy to language: Acquiring the power of expression*. New York, NY: Cambridge University Press.

Bloom, P., & Wynn, K. (1997). Linguistic cues in the acquisition of number words. *Journal of Child Language*, *24*, 511–533.

Briars, D. J., & Siegler, R. S. (1984). A featural analysis of preschoolers' counting knowledge. *Developmental Psychology*, *20*, 607–618.

Canobi, K. H. (2004). Individual differences in children's addition and subtraction knowledge. *Cognitive Development*, *19*, 81–94.

Carey, S. (2001). Whorf versus continuity theorists: Bringing data to bear on the debate. In M. Bowerman & S. C. Levinson (Eds.), *Language acquisition and conceptual development* (pp. 185–214). New York: Cambridge University Press.

Choi, S., & Bowerman, M. (1991). Learning to express motion events in English and Korean: The influence of language-specific lexicalization patterns. *Cognition*, *41*, 83–121.

Clark, A. (1997). *Being there: Putting brain, body, and the world together again*. Cambridge, MA: MIT Press.

Clearfield, M. W., & Mix, K. S. (1999). Number versus contour length in infants' discrimination of small visual sets. *Psychological Science*, *10*, 408–411.

Clearfield, M. W., & Mix, K. S. (2001). Amount versus number: Infants' use of area and contour length to discriminate small sets. *Journal of Cognition and Development*, *2*(3), 243–260.

Feigenson, L., Carey, S., & Spelke, E. S. (2002). Infants' discrimination of number vs. continuous extent. *Cognitive Psychology*, *44*, 33–66.

Fodor, J. (1983). *The modularity of mind: An essay on faculty psychology*. Cambridge, MA: MIT Press.

Frye, D., Braisby, N., Lowe, J., Maroudas, C., & Nichols, J. (1989). Young children's understanding of counting and cardinality. *Child Development*, *60*, 1158–1171.

Fuson, K. (1988). *Children's counting and concepts of number*. New York: Springer.

Fuson, K. (1992). Relationships between counting and cardinality from age 2 to age 8. In J. Bideaud, C. Meljac, & J. P. Fischer (Eds.), *Pathways to number: Children's developing numerical abilities* (pp. 127–149). Hillsdale, NJ: Erlbaum.

Fuson, K., & Kwon, Y. (1992). Learning addition and subtraction: Effects of number words and other cultural tools. In J. Bideaud, C. Meljac, & J. P. Fischer (Eds.), *Pathways to number: Children's developing numerical abilities* (pp. 283–306). Hillsdale, NJ: Erlbaum.

Gallistel, C. R., & Gelman, R. (1992). Preverbal and verbal counting and computation. *Cognition*, *44*, 43–74.

Gelman, R. (1972). Logical capacity in very young children. *Child Development*, *43*, 75–90.

Gelman, R. (1991). Epigenetic foundations of knowledge structures: Initial and transcendent constructions. In S. Carey & R. Gelman (Eds.), *The epigenesis of mind: Essays on biology and cognition* (pp. 293–322). Hillsdale, NJ: Erlbaum.

Gelman, R., & Gallistel, C. R. (1978). *The child's understanding of number*. New York: Harvard University Press.

Gelman, S. A., & Markman, E. M. (1987). Young children's inductions from natural kinds: The role of categories and appearances. *Child Development*, *58*, 1532–1541.

Gelman, R., & Meck, E. (1983). Preschoolers' counting: Principles before skill. *Cognition*, *13*, 343–359.

Gentner, D., Rattermann, M. J., Markman, A., & Kotovsky, L. (1995). Two forces in the development of relational similarity. In T. J. Simon & G. S. Halford (Eds.), *Developing* *cognitive competence: New approaches to process modeling* (pp. 263–313). Hillsdale, NJ: Erlbaum.

Huttenlocher, J., Haight, W., Bryk, A., Seltzer, M., & Lyons, T. (1991). Early vocabulary growth: Relation to input and gender. *Developmental Psychology*, *27*, 236–248.

Huttenlocher, J., Jordan, N. C., & Levine, S. C. (1994). A mental model for early arithmetic. *Journal of Experimental Psychology: General*, *123*, 284–296.

Jensen, E. M., Reese, E. P., & Reese, T. W. (1950). The subitizing and counting of visually presented fields of dots. *Journal of Psychology*, *30*, 363–392.

Jevons, W. (1871). The power of numerical discrimination. *Nature*, *3*, 281–282.

Kaufman, E. L., Lord, M. W., Reese, T. W., & Volkmann, J. (1949). The discrimination of visual number. *American Journal of Psychology*, *62*, 498–525.

Landau, B., & Gleitman, L. R. (1985). *Language and experience*. Cambridge, MA: Harvard University Press.

Landau, B., Smith, L. B., & Jones, S. S. (1988). The importance of shape in early lexical learning. *Cognitive Development*, *3*, 299–321.

Levine, S. C., Jordan, N. C., & Huttenlocher, J. (1992). Development of calculation abilities in young children. *Journal of Experimental Child Psychology*, *53*, 72–103.

Levinson, S. C. (1994). Vision, shape, and linguistic description: Tzeltal body part terminology and object description. *Linguistics*, *32*, 791–885.

Longfield, E. (2004). *Abstract object recognition and alphabet knowledge in preschoolers.* Unpublished Honors Thesis, Indiana University.

Lucy, J. (1992). *Language diversity and thought: A reformulation of the linguistic relativity hypothesis*. Cambridge: Cambridge University Press.

Meck, W. H., & Church, R. M. (1983). A mode control model of counting and timing processes. *Journal of Experimental Psychology: Animal Behaviour Processes*, *9*, 320–334.

Menninger, K. (1958/1992). *Number words and number symbols: A cultural history of numbers*. New York: Dover.

Mervis, C. B. (1985). On the existence of prelinguistic categories: a case study. *Infant Behavior & Development*, *8*, 293–300.

Miura, I. T., & Okamoto, Y. (1989). Comparisons of US and Japanese first graders' cognitive representation of number and understanding of place-value. *Journal of Educational Psychology*, *81*, 109–113.

Mix, K. S. (1999a). Preschoolers' recognition of numerical equivalence: Sequential sets. *Journal of Experimental Child Psychology*, *74*, 309–332.

Mix, K. S. (1999b). Similarity and numerical equivalence: Appearances count. *Cognitive Development*, *14*, 269–297.

Mix, K. S. (2002). The construction of number concepts. *Cognitive Development*, *17*, 1345–1363.

Mix, K. S. (2004a) *Are number concepts special?: Cross-mapping and complexity effects on children's numerical equivalence judgments.* Unpublished manuscript.

Mix, K. S. (2004b). *How Spencer made number: An observational study of early number word use.* Cognitive Science Technical Report #255, Indiana University.

Mix, K. S., Huttenlocher, J., & Levine, S. C. (1996). Do preschool children recognize auditory–visual numerical correspondences? *Child Development*, *67*, 1592–1608.

Mix, K. S., Huttenlocher, J., & Levine, S. C. (2002a). Multiple cues for quantification in infancy: Is number one of them? *Psychological Bulletin*, *128*, 278–294.

Mix, K. S., Huttenlocher, J., & Levine, S. C. (2002b). *Quantitative development in infancy and early childhood*. New York: Oxford University Press.

Nelson, K. (1996). *Language in cognitive development: The emergence of the mediated mind*. New York: Cambridge University Press.

Piaget, J. (1965/1941). *The child's conception of number*. New York: Norton.
Pinker, S. (1994). *The language instinct*. New York: William Morrow.
Rice, N. (1980). *Cognition to language*. Baltimore, MD: University Park Press.
Rittle-Johnson, B., & Siegler, R. S. (1998). The relationship between conceptual and procedural knowledge in learning mathematics: A review. In C. Donlan (Ed.), *The development of mathematical skills* (pp. 75–110). Hove, UK: Psychology Press.
Russell, B. (1919/1993). *Introduction to mathematical philosophy*. New York: Dover.
Sandhofer, C. M., & Mix, K. S. (2003). Number language and number concepts: Evidence from a long-range microgenetic study. In K. S. Mix & C. M. Sandhofer (Chairs), *Multiple approaches to understanding the interaction of language and cognition.* Symposium Conducted at the Biennial Meeting of the Society for Research in Child Development, Tampa, FL. April.
Sandhofer, C. M., & Smith, L. B. (1999). Learning color words involves learning a system of mappings. *Developmental Psychology*, *35*, 668–679.
Saxe, G. B. (1988). Candy selling and math learning. *Educational Researcher*, *17*, 14–21.
Saxe, G. B. (1991). *Culture and cognitive development: Studies in mathematical understanding*. Hillsdale, NJ: Erlbaum.
Schaeffer, B., Eggleston, V., & Scott, J. (1974). Number development in young children. *Cognitive Psychology*, *6*, 357–379.
Simon, T., Hespos, S. J., & Rochat, P. (1995). Do infants understand simple arithmetic? A replication of Wynn. *Cognitive Development*, *10*, 253–269.
Smith, L. B. (1984). Young children's understanding of attributes and dimensions: A comparison of conceptual and linguistic measures. *Child Development*, *55*, 363–380.
Smith, L. B. (1993). The concept of same. In H. W. Reese (Ed.), *Advances in child development and behavior* (Vol. 24, pp. 215–252). New York: Academic Press.
Smith, L. B., Gasser, M., & Sandhofer, C. M. (1997). Learning to talk about the properties of objects: A network model of the development of dimensions. In R. L. Goldstone, P. G. Schyns, & D. L. Medin (Eds.), *Psychology of learning and motivation* (Vol. 36, pp. 219–255). San Diego, CA: Academic Press.
Smith, L. B., Jones, S. S., & Landau, B. (1992). Count nouns, adjectives, and perceptual properties in children's novel word interpretations. *Developmental Psychology*, *28*, 273–286.
Smith, L. B., Jones, S. S., Landau, B., Gershkoff-Stowe, L., & Samuelson, L. (2002). Object name learning provides on-the-job training for attention. *Psychological Sciences*, *13*, 13–19.
Soja, N. N. (1994). Young children's concept of color and its relation to the acquisition of color words. *Child Development*, *65*, 918–937.
Sophian, C. (1997). Beyond competence: The significance of performance for conceptual development. *Cognitive Development*, *12*, 281–303.
Spelke, E. (2003). What makes us smart? Core knowledge and natural language. In D. Gentner & S. Goldin-Meadow (Eds.), *Language in mind* (pp. 277–311). Cambridge, MA: MIT Press.
Spelke, E. S., & Tsivkin, S. (2001). Innate knowledge and conceptual change: Space and number. In M. Bowerman & S. C. Levinson (Eds.), *Language acquisition and conceptual development* (pp. 70–97). New York: Cambridge University Press.
Starkey, P., & Cooper, R. G. Jr. (1980). Perception of numbers by human infants. *Science*, *210*, 1033–1035.
Starkey, P., Spelke, E. S., & Gelman, R. (1990). Numerical abstraction by human infants. *Cognition*, *36*, 97–127.

Strauss, M. S., & Curtis, L. E. (1981). Infant perception of numerosity. *Child Development*, *52*, 1146–1152.

Taves, E. H. (1941). Two mechanisms for the perception of visual numerousness. *Archives of Psychology*, *37*, 1–47.

Thelen, E., & Smith, L. B. (1994). *A dynamic systems approach to the development of cognition and action*. Cambridge: MIT Press.

Vygotsky, L. S. (1962). *Thought and language*. Cambridge, MA: MIT Press.

Wagner, S. H., & Walters, J. (1982). A longitudinal analysis of early number concepts: From numbers to number. In G. Forman (Ed.), *Action and thought: From sensorimotor schemes to symbolic operations* (pp. 137–161). New York: Academic Press.

Waxman, S. R., & Markow, D. B. (1995). Words as invitations to form categories: Evidence from 12- to 13-month-old infants. *Cognitive Psychology*, *29*, 257–302.

Whorf, B. L. (1956). In J. B. Carroll (Ed.), *Language, thought, and reality: Selected writings of Benjamin Lee Whorf*. Cambridge, MA: MIT Press.

Wynn, K. (1990). Children's understanding of counting. *Cognition*, *36*, 155–193.

Wynn, K. (1992). Children's acquisition of the number words and the counting system. *Cognitive Psychology*, *24*, 220–251.

Wynn, K. (1998). Psychological foundations of number: Numerical competence in human infants. *Trends in Cognitive Science*, *2*, 296–303.

Xu, F., & Spelke, E. S. (2000). Large number discrimination in 6-month-old infants. *Cognition*, *74*, 1–11.

Author Index

A

B

C

F

G

H

I

M

N

O

P

R

Subject Index

B

C

P

Q

R

S

T

U

V

W

Y

This Page Intentionally Left Blank

Contents of Previous Volumes

Volume 11

Volume 12

Volume 13

Volume 19

Volume 20

Volume 21

Volume 22

Volume 23

Volume 24

Volume 25

Volume 26

Volume 27

Volume 28

Volume 29