Jane B. Childers *Editor*

Language and Concept Acquisition from Infancy Through Childhood

Learning from Multiple Exemplars

Springer

Language and Concept Acquisition from Infancy Through Childhood

Jane B. Childers
Editor

Language and Concept Acquisition from Infancy Through Childhood

Learning from Multiple Exemplars

Springer

Editor
Jane B. Childers
Trinity University
San Antonio, TX, USA

ISBN 978-3-030-35593-7   ISBN 978-3-030-35594-4 (eBook)
https://doi.org/10.1007/978-3-030-35594-4

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

I dedicate this book to my husband, Danny, who has always supported me; my daughter, Emily, who makes my life an adventure; and my parents, Don and Edie, who taught me how to read and how to be curious about the world. Phil. 4:13.

Preface

As is true for many important things in life, this book grew out of an experience with failure. Specifically, I had just given (what I thought was) a riveting talk at a major conference detailing important new findings from my lab. In the question portion of the talk, a key question asked by an audience member was how a competing theory could explain my data, a theory I did not know much about. How could this occur? Easily. We all work within specific theoretical silos, knowing some about opposing views but (for many of us, I suspect) not knowing as much about the theories that are not within our frame of reference as we do about the theory we rely on. Thus, the idea for this book was born – a book about distinct theoretical frameworks for the same phenomenon, infants' and children's use of a set of examples for learning. The goal was simple: to bring together differing perspectives on how children accomplish this type of unsupervised learning in a single book so that we could all easily learn from this set of examples! Specifically, this book brings into juxtaposition research from a range of theories (e.g., structural alignment, statistical learning, Bayesian learning). Particular theoretical commitments drive the types of questions we ask and the kinds of studies we design – to what extent can different theories of example-based learning mutually inform each other's commitments and phenomena?

Editing this book has provided me with a chance to learn and think about the connections among the theories in this area, and it should be openly acknowledged that not all theories are represented. Yet, let's embark on this comparison of comparison theories and see what we can learn by comparing theories to each other! Based on the growing evidence of the power of learning from examples exemplified in these chapters, by the end of this book, I predict that we all will leave with a richer understanding of the nature of learning. Happy reading!

San Antonio, TX, USA Jane B. Childers

Acknowledgments

The first group to be acknowledged is the group of researchers who contributed chapters. This was my first experience editing a book, and their patience and generosity in sharing their research and ideas are greatly appreciated.

Additionally, this book would not have been completed without the guidance, support, and mentoring of Laura Namy and Susan Graham. They were extremely helpful in this process, and our periodic long-distance Skype calls were invaluable. Many thanks to both of them!

Contents

1 **Introduction** 1
Jane B. Childers

2 **Mechanisms of Statistical Learning in Infancy** 11
Scott P. Johnson

3 **How Multiple Exemplars Matter for Infant Spatial Categorization** 31
Marianella Casasola and Youjeong Park

4 **How the Demands of a Variable Environment Give Rise to Statistical Learning** 59
Erik D. Thiessen

5 **Structure-Mapping Processes Enable Infants' Learning Across Domains Including Language** 79
Susan J. Hespos, Erin Anderson, and Dedre Gentner

6 **The Emergence of Inductive Reasoning During Infancy: Learning from Single and Multiple Exemplars** 105
Susan A. Graham, Michelle S. Zepeda, and Ena Vukatana

7 **Learning Individual Verbs and the Verb System: When Are Multiple Examples Helpful?** 131
Mutsumi Imai and Jane B. Childers

8 **Multiple Examples Support Children's Word Learning: The Roles of Aggregation, Decontextualization, and Memory Dynamics** 159
Catherine Sandhofer and Christina Schonberg

9 Mechanisms for Evaluating Others' Reliability When Learning Novel Words 179
David M. Sobel, Elena Luchkina, and Kristen Tummeltshammer

10 The Search for Invariance: Repeated Positive Testing Serves the Goals of Causal Learning 197
Elizabeth Lapidow and Caren M. Walker

11 Multiple Exemplars of Relations 221
Stella Christie

12 Epilogue: Comparing Comparison Theories: What Can We Gain? 247
Jane B. Childers

Index 253

Contributors

Erin Anderson Northwestern University, Evanston, IL, USA

Marianella Casasola Cornell University, Ithaca, NY, USA

Jane B. Childers Department of Psychology, Trinity University, San Antonio, TX, USA

Stella Christie Tsinghua University, Beijing, China

Dedre Gentner Northwestern University, Evanston, IL, USA

Susan A. Graham University of Calgary, Calgary, AB, Canada

Susan J. Hespos Northwestern University, Evanston, IL, USA

Mutsumi Imai Keio University, Tokyo, Japan

Scott P. Johnson University of California, Los Angeles, Los Angeles, CA, USA

Elizabeth Lapidow Department of Psychology, University of California – San Diego, La Jolla, CA, USA

Elena Luchkina Brown University, Providence, RI, USA

Youjeong Park Seoul National University, Seoul, South Korea

Catherine Sandhofer University of California, Los Angeles, Los Angeles, CA, USA

Christina Schonberg University of Wisconsin–Madison, Madison, WI, USA

David M. Sobel Brown University, Providence, RI, USA

Erik D. Thiessen Carnegie Mellon University, Pittsburgh, PA, USA

Kristen Tummeltshammer Brown University, Providence, RI, USA

Ena Vukatana University of Calgary, Calgary, AB, Canada

Caren M. Walker Department of Psychology, University of California – San Diego, La Jolla, CA, USA

Michelle S. Zepeda University of Calgary, Calgary, AB, Canada

About the Editor

Jane B. Childers (Ph.D. 1998, University of Texas at Austin) is Professor and Chair of the Department of Psychology, and Director of the Linguistics Program, at Trinity University, San Antonio. Her main focus of research examines children's early verb learning, with an emphasis on how the comparison of events may be useful for deducing verb meaning. Her research has been funded by the Spencer Foundation and the National Institutes of Health, and she serves on the Editorial Board of the *Journal of Cognition and Development* and the *Journal of Experimental Child Psychology*.

Chapter 1
Introduction

Jane B. Childers

Abstract There are multiple theories explaining how children successfully identify and use patterns from exemplars in day-to-day life to support their learning. These influential perspectives both converge and diverge in the specific processes by which infants and children learn prior to formal schooling. This book constitutes the first systematic integration of highly influential research traditions in the domains of language and concept acquisition including statistical learning, the structure-mapping account, and other perspectives. In this chapter, a brief summary of each chapter in the book is provided, and some preliminary links between chapters are outlined. More generally, by exploring both the benefits and challenges children face as they learn from multiple examples, and the major theories guiding different areas of research, both established researchers and graduate students should be able to better understand children's early unsupervised language and concept learning.

This book examines the role of implicit, experience-based learning on children's acquisition of language and concepts. There are multiple theories explaining how children successfully identify and use patterns from exemplars in day-to-day life to support their learning. These influential perspectives both converge and diverge in the specific processes by which infants and children learn prior to formal schooling. The goal of this volume is to review, compare, and contrast accounts of how the opportunity to recognize and generalize patterns across examples influences learning. This book constitutes the first systematic integration of highly influential research traditions in the domains of language and concept acquisition including statistical learning, the structure-mapping account, and other perspectives. We invited contributors who used different theories that each seek to explain how children learn from multiple examples to consider (1) why learning in their domain requires multiple examples, (2) what mechanisms are central to the theory they use, (3) what procedures are used in their area and whether other procedures could be used to test learning in that area, (4) how other theories in the area may apply to their area, and (5) how the theory they use could be extended beyond the contexts in which it has been used.

J. B. Childers (✉)
Department of Psychology, Trinity University, San Antonio, TX, USA
e-mail: jchilder@trinity.edu

J. B. Childers (ed.), *Language and Concept Acquisition from Infancy Through Childhood*, https://doi.org/10.1007/978-3-030-35594-4_1

All contributors were charged with drawing connections between their own and the other theories with respect to the underlying assumptions and mechanisms involved. In the epilogue, we seek to synthesize the perspectives of the contributing authors with the goal of identifying integrative themes, key differences in underlying assumptions, and important future directions. By exploring both the benefits and challenges children face as they learn from multiple examples, and the major theories guiding different areas of research, we hope that both established researchers and graduate students will be able to better understand children's early unsupervised language and concept learning. In the next section, we will briefly describe each chapter that follows in this book and begin to discuss how the ideas in these chapters may fit together.

In Chap. 2, Scott Johnson provides a clear description of statistical learning as a set of processes for learning from distributional information. He then provides foundational evidence from infant speech perception studies before describing new evidence from studies of visual sequences with 2-month-olds. Infants appear to be able to show statistical learning abilities at birth (with simple stimuli) and improve over the first year, being able to process more complex stimuli with time. Other cues can also help statistical learning including social cues and prosody. Different procedures have been used to study statistical learning including using eye tracking to show that infants can anticipate items in a sequence, and using fMRI and measures of event-related potentials (ERP) to compare infants at risk for ASD with typically developing infants. There are also two possible mechanisms that underlie statistical learning: the storage of information about transitional probabilities (TPs) and the storage of chunks. Johnson provides evidence that infants may store TPs on the way to storing chunks. In sum, Johnson shows statistical learning is available very early in development, can be used in both auditory and visual tasks, can be tested using multiple procedures, and can be linked to developmental disabilities. Exploring the mechanisms that underlie statistical learning is a key task in current and future research.

In Chap. 3, Casasola and Park start with a compelling argument for exploring infant spatial cognition by linking it to several other domains including object categorization and relational learning. They then discuss prior studies that have included different types of variability that begin to show how different experiences may affect infants' learning in this area. Infants' ability to compare diverse examples improves with development, and the chapter provides a great description of how the understanding of spatial relations emerges in the first year. They remind us that abilities seen earlier in development may not extend across tasks or may be more fragile. An important set of studies described in this chapter are the studies that have tested different spatial relations (containment and support) at different ages and with varying number of examples (two vs. six). These studies show an important developmental change between 10 and 14 months, with 14-months-old possibly benefitting from seeing fewer examples (two) as compared to the 10-month-olds, but likely because they are processing the stimuli more carefully, and in more detail, than infants are at the younger ages. The chapter ends with a discussion of different mechanisms that may underlie this learning, beginning with structural alignment or structure-mapping theory and then considering statistical learning as a possible mechanism. They then discuss links between spatial tasks and language. Overall, the findings that the perceptual characteristics of stimuli, the numbers of examples, and the variety across examples are processed differently at different ages, are important findings that should extend to other domains discussed in this book. By comparing two theories in this chapter, readers can begin to see how each theory may apply to spatial cognition or to learning more generally.

In Chap. 4, Thiessen starts by reminding us that learning language requires using variable information across examples because language input is variable and children must become generative, productive speakers. The chapter focuses on statistical learning as a theoretical framework, starting with examples of learning by attending to conditional probabilities. As in other chapters, he notes that variability across examples not only increases the complexity of the learning task but also helps generalizability. Examples of acquiring phonemes and attending to voice onset time (VOT) are provided, and evidence for statistical learning in the "real world" is given. The chapter then turns from conditional probability to distributional probability or using cross-situational statistics. In this area, the frequency of exemplars influences a category, as will also be discussed in other chapters. A central thesis concerns the question of what underlies statistical learning, and an important part of this chapter is examining how aspects of human memory may explain this type of learning. Theissen briefly discusses exemplar vs. prototype models and reminds us that there is good evidence for both, and then turns to two theories of neural processing that may explain how statistical learning may be instantiated in the brain. The first is the complementary learning systems approach (McClelland, McNaughton, & O'Reilly, 1995), in which both the hippocampus and neocortex interact to produce statistical learning effects, and the second is the hippocampal dynamic theory in which different parts of the hippocampus interact in statistical learning. Finally, exploring the roles of interference in memory and the role of similarity-based activation provides other links to the memory system. In sum, statistical learning may arise from memory processes, which could be key mechanisms that underlie this ability.

Chapter 5 (Hespos, Anderson, and Gentner) applies Gentner's structure-mapping theory to infant learning, particularly focusing on infants' understanding of the same/different relation. While prior studies focusing on analogical processing have examined children over the age of 2 years, these data investigate the origins of this ability. Two key aspects of the structure-mapping view are that comparing examples of a relation helps children improve in their ability to perceive relations and that attention to specific objects can interfere with relational processing. These are tested and found to apply in studies of young infants. An initial study shows infants at 7–9 months can abstract the same/different relation from a series of examples and extend the relation to new objects. Additionally, prior experience with the specific objects involved in the test examples interferes with their relational processing. In a second study, these findings are extended to 3-month-olds. These studies also clarify the role of variability across examples: 7- and 9-month-olds learn from four examples, but not a single one, while 3-month-olds succeed with two examples, but not six. In the next part, the authors compare three different relational learning paradigms (match to sample, relational match to sample, and same/different discrimination) and review results from animal studies to put their infant results into context. Human infants are in a small group of animals that can solve this task, and they do it early (by 3 months) and without language cues. It is noteworthy that human infants can learn these relations with fewer than ten trials and it takes other species hundreds and sometimes thousands of trials to attain similar levels of performance (although in a somewhat different paradigm). In the next section of the chapter, the authors show that the structure-mapping framework extends to studies of language acquisition, including studies of word learning and statistical learning. For example, a simulation using the structure-mapping engine (SME) successfully captures infant learning in the Marcus et al. study of artificial grammar learning. The chapter ends by extending the theory to physical reasoning tasks (covering events). In sum, this chapter applies structural alignment theory to studies with young infants, providing evidence that relational abilities are present in the first year and are part of what makes humans different from other animals.

In Chap. 6, Graham et al. describe category-based inductive reasoning during the infancy and preschool years, highlighting the developmental continuity in fundamental inductive abilities across early childhood. They focus on adherence to a core induction principle, namely, *premise-conclusion similarity*, delineating the types of similarity that preschoolers and infants use to license their inductive inferences. Notably, infants as young as 13 months of age will privilege category information over perceptual information when extending nonobvious properties from one category member to another. They then review recent studies from their lab examining the developmental origins of inductive reasoning between 9 and 11 months of age, with particular attention to the role of single vs. multiple exemplars in facilitating infants' tendency to link properties with categories. Their results demonstrate that infants' ability to associate properties with object categories is in place by 9 months of age but is modulated by a number of interacting factors, including category type (familiar vs. unfamiliar) and number of exemplars (i.e., whether infants are exposed a single-category exemplar or multiple-category exemplars during familiarization). When learning about familiar categories (i.e., dogs and cats), infants will generalize a property from a single exemplar. When faced with the more demanding task of learning about an unfamiliar category, exposure to multiple exemplars facilitates infants' category-property generalizations.

In Chap. 7, Mutsumi Imai and Jane Childers focus on the problem of verb learning. The chapter starts with a description of the problem, contrasting it with noun learning and discussing differences across languages. They then remind the reader that children need to learn a verb system because a specific verb's meaning is understood in relation to all of the other verbs in the lexicon of a language. They propose that verb learning occurs in three phases: finding the core of a verb's meaning within an event, discovering dominant patterns in a language, and delineating boundaries between individual verb meanings. In the first phase, there are two types of perceptual similarity that could be helpful to young language learners—sound symbolism and the similarity between objects in the original event and in new events. Empirical evidence is given for both of these types including studies of children learning

Japanese, Chinese, and English. In sum, children could use perceptual similarity, in the form of sound symbolism or object similarity, to get to a better understanding of relational similarity. Next, several studies showing that children benefit from seeing high-similarity examples before low-similarity ones are described, as well as studies of children's use of contrast in verb learning. After describing how children may come to recognize patterns across verb meanings in a language, the authors turn to the question of how children learn a verb within an overall system by describing a study of verbs for carrying/holding in Chinese. This study used video stimuli and production data from children across multiple ages, their mothers, and undergraduates. Thirteen verbs were included in the videos and participants were asked to name the action. This study shows that mothers were not different from undergraduates in the number of verbs produced and that children produced fewer verbs than adults did, but there were no significant differences in the number of verbs produced for the videos between 3 and 7 years. The researchers examined how children were applying the verbs to the events and found that, with age, children grew closer to the adult uses. Both multidimensional scaling (MDS) and individual differences scaling (INDSCAL) analyses reveal children attended to the objects in the events and used the objects to inform their verb meanings. They also reveal three semantic islands of verb meaning. An entropy analysis shows that there is an early stage of verb learning in which input frequency is important and a later stage of verb learning in which the degree of boundary overlap with other verbs affects their ease of acquisition. In sum, the chapter provides ample evidence of children's use of multiple exemplars for verb learning, using structure mapping as a theoretical framework. It also includes an interesting developmental story that puts the whole of verb learning in perspective.

In Chap. 8, Cathy Sandhofer and Christina Schonberg discuss how statistical learning applies to the problem of learning words, specifically nouns, focusing on children who are 2 years old and older. This chapter brings in new sources not discussed in the prior infant statistical learning chapters, especially Smith and Yu's work showing that infants, children, and adults can learn co-occurrences between words and references in a laboratory task. They note that delays between examples and increased complexity of examples disrupt processing. They then describe an important study by Vlach and Sandhofer (2011) examining the role of context and timing on memory and comparison. This study shows that 2½- to 3-year-old children have difficulty learning a new word and extending it to new objects at test when the context of the training trial and test trial differs, and that a condition in which the contexts across learning trials and test trials all varied yielded the fewest numbers of extensions. They then make the important point that access to multiple correlated cues, or support, is especially important when instances are varied because these cues help the learner aggregate information across examples, as demonstrated in a study by Goldenberg and Sandhofer (2013). These correlated cues should be especially needed when there are delays between instances and/or when learners are younger or less experienced; this could explain why, in their prior findings (2011), children had the most difficulty when multiple cues were not available. Next, they tackle the question of how variation across exemplars influences category learning by describing a study by Perry, Samuelson, Malloy, and Schiffer (2010). In that study, 18-month-olds who only saw highly similar examples when learning new categories became more reliant on shape than necessary for those categories, while toddlers experiencing more variation were able to discover when object shape was important and when it was not. A final section of the chapter links statistical learning with memory processes (also linked in Chap. 4, Thiessen). They describe a study of word learning (Vlach & Sandhofer, 2011) showing that most forgetting of a new word over time happens between the immediate test and 1-week delay, with less forgetting between 1-week and 1-month delays. They describe how forgetting can help learning from a study-phase retrieval theory perspective and give evidence (Vlach, Ankowski, & Sandhofer, 2012) showing that spaced practice led to better performance in a memory task presented 15 minutes after learning than did simultaneous or massed practice conditions. Overall, several interesting new ideas are discussed in this chapter including the role of context cues, the tension between similarity and variability of examples, and learning over delays.

In Chap. 9, David Sobel, Elena Luchkina, and Kristen Tummeltshammer describe how toddlers and young children learn words from reliable speakers, and avoid unreliable speakers, using statistical learning processes and real-world knowledge. The chapter begins by describing a developmental change between 16 months, when toddlers expect all speakers to be accurate labelers, and 4 years, when children can distrust speakers. It includes descriptions of many relevant prior studies in this area that provide the foundation for their studies. Interestingly, earlier understanding of distrust can be seen in studies that include contrast information between speakers. They then discuss a new study with 18-month-old children using the intermodal preferential looking paradigm (IPLP) procedure and reliable and unreliable speakers. This study shows children only looked longer at the correct object + label match when the speaker was accurate, which is at an earlier age than has been shown. A second study examines the origins of children's ability to use social cues to assess others' reliability. Eight-month-old infants were shown a face that gazed at an interesting video consistently over trials vs. one that was inconsistent. This study shows that these young infants could follow the gaze of a reliable face and generalize their response to new locations and did not do so with unreliable one. Statistical learning could account for these results if infants are forming a unit across multiple co-occurrences of a word + object + speaker and then use those co-occurrences to judge the valence of the speaker. The chapter then returns to developmental change noting that in categorization, 14-month-old infants attend to any regularity but 18-month-olds are more selective about the types of regularities they consider (Madole & Cohen, 1995). To test this in the word-learning domain, the authors conducted a new study with dynamic arrows in addition to faces which showed that at 5 months, infants attended to regularities between objects and locations for faces and arrows (and could not generalize for either type); at 8 months, they generalized from faces but not arrows; and at 12–13 months, they followed the gaze and generalized from both accurate and inaccurate faces. The authors give reasons for this pattern of results, linking it back to prior research. In a final study, speakers asked a question with a new word (instead of giving a label in a statement) so that the co-occurrence of the label, object, and speaker was preserved without accuracy information present. Eighteen-month-olds showed learning of familiar object labels but were at chance for novel object (+novel speaker at test) labels. Three- and four-year-old children remembered labels from both accurate and inaccurate speakers equally, and exceeded chance, but were slower to respond when speakers were inaccurate. In sum, these researchers propose that infants and toddlers are using statistical learning processes across examples to form the initial associations between speakers, labels, and objects, and to generalize this information to new instances when judging whether speakers are reliable speakers or not. Children continue to use statistical learning processes with development but add existing knowledge (e.g., about pragmatics). This model of statistical learning with a higher-level rational mechanism could extend to all selective social learning tasks or even to all of learning.

In Chap. 10, Elizabeth Lapidow and Caren Walker tackle a problem not yet explored in this book, which is how children create multiple examples that can be compared in the context of hypothesis testing. A key issue in this area is that both adults and children typically like to perform tests or actions that support their current hypotheses—or positive tests—instead of conducting tests that could falsify a hypothesis. Such repeated demonstrations of the current hypothesis initially seem less helpful for learning. However, in the chapter, Lapidow and Walker show that positive examples can be useful because by demonstrating a hypothesis across time and/or contexts, the learner can discern the range of events to which that hypothesis can apply. In the first section of the chapter, the researchers define the positive testing strategy (PTS) and distinguish it from "confirmation bias," providing an overview of past research in the area of scientific reasoning and rule learning. They then discuss previous accounts of why children and adults engage in PTS, grouping these theories into two categories. According to one prior account, learners are assumed to be motivated to produce desirable outcomes or an "engineering goal." According to others, learners fail to recognize the value of disconfirming alternative hypotheses. A main point is that PTS should not always be seen as an error in thinking; there can be good reasons for it. Finally, the authors introduce a novel theory—the search for invariance (SI) hypothesis—drawing on accounts of early exploratory causal learning, which emphasizes the importance of invariance for the acquisition of causal knowledge. According to this account, learners produce multiple positive tests in order to assess the degree to which a current hypothesis holds across time and contexts. This fits a common theme throughout the chapters which is information across examples informs transfer or generalizability. The authors then go back to three prominent studies of positive testing and reinterpret their results from the perspective of invariance. Overall, the chapter extends the question of how children form multiple examples to a new domain (hypothesis testing) and asks why learners create opportunities to gather examples in the ways that they do, and how these examples influence learning and transfer.

In the final chapter (Chap. 11), Christie adds one more area of research in which children likely learn from multiple examples—the social domain. She begins with descriptions of studies that show that relations can be hard to perceive when learners are not aware of the particular relation, and when object or perceptual similarity competes with attention to relations. She gives an overview of structure-mapping theory (also see Hespos, Chap. 5; Imai and Childers, Chap. 7) with evidence from a study showing 4-year-olds needed to compare across examples to learn a novel relation. She then discusses two issues in comparison: the level of surface similarity needed across comparisons and the number of comparisons needed. The answers to these questions depend on the domain and the prior experience of the learner. Two adult studies (negotiation strategies, geoscience learning) are described. A study of adults' learning nonadjacent dependencies in grammar suggests many examples can be useful and, at times, explicit instructions to compare are important. This leads into four factors that promote comparison across examples: high object similarity (including progressive alignment examples), specific comparative words (e.g., "more," "taller") within a domain, systematic description of a problem (a study with number words provides evidence), and labels. The final section extends these ideas into the social domain. Young children can perceive social relations in the first 18 months but also can show some difficulties learning them. This could be due to a lack of specific social knowledge that they need and issues when surface similarity competes with relational reasoning. A new study applying structure-mapping theory to children's false belief understanding shows that children who can compare across examples show better false belief reasoning at test (vs. children without comparisons). Comparing highly similar examples usually helps comparison but could also be harmful in the social domain as it could create narrower social categories. On the other hand, alignment and comparison can highlight deeper relational information, which could be helpful in the social domain. New evidence shows children reason differently about people vs. animals, thinking more broadly about people. In sum, the chapter brings in a new domain—children's learning of social roles and categories—and describes how structure-mapping theory could apply in this new promising area of research.

In sum, this book begins with five chapters focusing on infant development and their use of multiple exemplars in the understanding of objects in visual sequences, spatial relations, speech sounds, and relational concepts and their ability to make inductive generalizations in a category. In the next five chapters, the authors focus on development in preschoolers and older children taking on the problems of how children could use information across multiple exemplars to learn verbs (and the verb system) and nouns, consider whether words are heard from reliable speakers, construct multiple examples in scientific reasoning, and learn social categories. Across these chapters, researchers largely frame their work in terms of statistical learning or structure-mapping theories, though at times the particular theory is not articulated and the focus is on development. Key themes across chapters are that comparison skills develop such that children can process more varied exemplars with age and can bring in real-world knowledge to add to their processing with age. In the epilogue, I will return to these key themes and add new ones to more fully consider how this body of work fits together.

References

Goldenberg, E. R., & Sandhofer, C. M. (2013). Same, varied, or both? Contextual support aids young children in generalizing category labels. *Journal of Experimental Child Psychology, 115*(1), 150–162. https://doi.org/10.1016/j.jecp.2012.11.011.

Madole, K. L., & Cohen, L. B. (1995). The role of object parts in infants' attention to form-function correlations. *Developmental Psychology, 31*(4), 637–648.

McClelland, J. L., McNaughton, B. L., & O'Reilly, R. C. (1995). Why there are complementary learning systems in the hippocampus and neocortex: Insights from the successes and failures of connectionist models of learning and memory. *Psychological Review, 102*(3), 419–457.

Perry, L. K., Samuelson, L. K., Malloy, L. M., & Schiffer, R. N. (2010). Learn locally, think globally: Exemplar variability supports higher-order generalization and word learning. *Psychological Science, 21*(12), 1894–1902.

Vlach, H. A., Ankowski, A. A., & Sandhofer, C. M. (2012). At the same time or apart in time? The role of presentation timing and retrieval dynamics in generalization. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 38*(1), 246–254. https://doi.org/10.1037/a0025260.

Vlach, H. A., & Sandhofer, C. M. (2011). Developmental differences in children's context-dependent word learning. *Journal of Experimental Child Psychology, 108*(2), 394–401. https://doi.org/10.1016/j.jecp.2010.09.011.

Chapter 2
Mechanisms of Statistical Learning in Infancy

Scott P. Johnson

Abstract Statistical learning is the process of identifying patterns of probabilistic co-occurrence among stimulus features, essential to our ability to perceive the world as predictable and stable. Research on auditory statistical learning has revealed that infants use statistical properties of linguistic input to discover structure that may facilitate language acquisition. More broadly, statistical learning operates across sensory modalities and across species. Research on infants' visual statistical learning has revealed that statistical learning develops over time, yet the mechanisms (including developmental mechanisms) underlying infant performance remain unclear. This chapter examines competing models of statistical learning and how learning might be constrained by limits in infants' attention, perception, and memory.

The means by which humans acquire and represent knowledge is fundamental to cognitive science, and a central question asked by developmental psychologists concerns how infants learn so much in so little time without explicit instruction. For example, the rapidity and apparent ease with which infants and young children understand and produce speech, recognize faces, interpret others' mental states, detect violations of physical laws governing object properties, and discriminate different numbers of items have led some theorists to suggest that innate cognitive mechanisms—independent of learning and experience—provide the infant with some knowledge in each of these domains (Chomsky, 1965; Johnson & Morton, 1991; Leslie, 1997; Spelke, 1990; Wynn, 1992). Yet such views may neglect the role of environmental structure in guiding development, and studies of infant *statistical learning* (SL), the focus of this chapter, can help shed light on this issue. Statistical learning (SL) is a set of processes for learning stimulus features, objects, and events from *distributional* information over space and time: simple associations, probabilistic correspondence, frequencies, spatial positions, and order in sequence. SL contributes to segmentation of continuous information (such as speech) and the

S. P. Johnson (✉)
University of California, Los Angeles, Los Angeles, CA, USA
e-mail: spjohnson@g.ucla.edu

J. B. Childers (ed.), *Language and Concept Acquisition from Infancy Through Childhood*, https://doi.org/10.1007/978-3-030-35594-4_2

formation of representations of units in time and space, thus helping to shape cognitive development (Frost, Armstrong, Siegelman, & Christiansen, 2015; Siegelman & Frost, 2015), and it is an important part of language acquisition (see Chap. 4).

In this chapter, I discuss research efforts to discover the nature of SL in infancy, the kinds of statistical structure that infants are able to learn, the impact of different testing methods on infant learning, implications of infant SL for cognitive development and developmental disabilities, and, finally, mechanisms underlying statistical learning in infancy. As I will try to make clear, the importance of statistical learning for understanding cognitive development, language acquisition in particular, has become increasingly evident in the 20+ years since publication of the first paper describing SL in infants (Saffran, Aslin, & Newport, 1996). Yet much remains unknown about the foundational processes and mechanisms of SL in infancy.

Statistical Learning in Infancy

Research on detection of structure in complex input sequences has a considerable history (e.g., Gibson & Gibson, 1955; Reber, 1967). It has long been known that adult learners can detect patterns in the absence of explicit (articulable) knowledge (Reber, 1989), raising questions of learnability of complex sequences by nonverbal populations. SL in infants was first reported by Saffran et al. (1996) with a head-turn procedure. Eight-month-old infants listened to a continuous stream of computer-generated speech for 2 minutes, followed by a test phase during which segments of the familiarized speech stream, now separated by brief pauses, alternated with segments whose order was scrambled or whose parts had co-occurred relatively infrequently in the training set. One study, for example, familiarized infants with the pseudo-words *tupiro*, *golabu*, *padoti*, and *bidaku* in random order and with no pauses or immediate repetitions (e.g., *tupirogolabupadotibidakugolabutupirobidakupadotitupiro*…). The test phase involved two of the four original "words" (e.g., *tupiro*, *golabu*) and two "nonwords" (e.g., *dapiku*, *tilado*) formed from a random assembly of syllables; words and nonwords were separated by a 500 ms gap. Infants in a second experiment heard words alternating with "part-words" formed from the last syllable of a word combined with the first two syllables of a different word (e.g., *bupado*, *kugola*). Discrimination of words from nonwords and part-words was evaluated during the test phase by recording look durations toward a flashing light that accompanied repeated presentation of test stimuli, on the right or left side of a testing chamber, on the assumption that interest in the sound sequences could be operationalized as attention in the direction of the light. Infants in both experiments exhibited increased interest in the novel items (nonwords and part-words).

How were infants able to parse the speech stream into coherent words, recognize them when heard in isolation, and discriminate them from the part- and nonwords? One possibility is that infants learned words from differences in *transitional probabilities* (TPs) between adjacent syllables, because there were no other cues to segmentation, such as pauses and prosody, that typically mark word and phrase boundaries in real-world speech (Fougeron & Keating, 1997; Wightman, Shattuck-Hufnagel, Ostendorf, & Price, 1992). TP is a statistical measure that describes the predictability of adjacent items in an array or sequence (Miller & Selfridge, 1950; the TP of successive element XY is defined as probability of Y|X = frequency of XY/frequency of X). In the Saffran et al. (1996) study, TPs within words such as *tupiro* were always 1.0, meaning *tu* perfectly predicted *pi*; in turn, *pi* perfectly predicted *ro* (see Fig. 2.1). TPs between words, however, were lower, averaging 0.33. This is because *ro* (in tupiro) was sometimes followed by *go* (in *golabu*), sometimes by *pa* (in *padoti*), and sometimes by *bi* (in *bidaku*). Thus nonwords and part-words heard during the test phase such as *dapiku* and *bupado* had lower TPs between syllables than words such as *tupiro* and *padoti*. The Saffran et al. results imply that infants detected the TP differences in the test stimuli and preferred to listen to the low-TP stimulus owing to its violation of word boundaries.

But there is an alternative explanation: Words in the familiarization stimulus were heard 3× more often than nonwords, and part-words were never heard, and so it is possible that infants preferred nonwords and part-words simply because they were unfamiliar, not due to lower TPs. To address this possibility, Aslin, Saffran, and Newport (1998) tested 8-month-olds with a "frequency-balanced" design in which the word and part-word heard at test were presented the same number of times during familiarization. TP differences, however, were the same as those in the Saffran et al. (1996) study. Infants showed increased interest in the part-word at test relative to the word, replicating the Saffran et al. results and providing evidence that segmentation and learning were based on TPs, not simple frequencies of syllables or words. TP differences between syllables, therefore, seem to facilitate the learning of sequence structure by signaling boundaries and units in an otherwise uninterrupted stream of items.

More broadly, SL operates across sensory modalities and across species. In human adults, SL participates in fundamental perceptual and cognitive functions including visual search, object perception, motor planning, and event prediction (Baker, Olson, & Behrmann, 2004; Fiser & Aslin, 2002a; Hunt & Aslin, 2001; Turk-Browne, Scholl, Johnson, & Chun, 2010). Animal species learn statistically

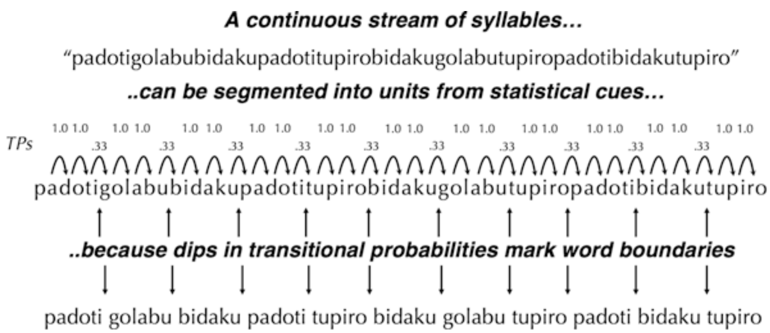


Fig. 2.1 Schematic description of how transitional probabilities between syllables mark word boundaries

structured speech streams (e.g., Hauser, Newport, & Aslin, 2001; Toro & Trobalón, 2005), and human infants parse streams of musical tones based on statistical probabilities (Saffran, Johnson, Aslin, & Newport, 1999).

Experiments in my lab (Kirkham, Slemmer, & Johnson, 2002) provided the first demonstration of infants' SL in visual sequences with an experiment in which infants were habituated to a stream of looming colored shapes organized in pairs defined by TPs. (Habituation is defined as a decrement in looking across trials according to a predetermined criterion, e.g., a decline of 50% or more during four successive trials relative to the first four trials.) TPs within pairs were 1.0, and TPs across pairs were 0.33 (see Fig. 2.2). Each shape had a unique color and loomed from about 4 to 24 cm across in 1 s, with no pauses to mark pairs. Following habituation, infants viewed two test sequences with the same shapes: a "structured" sequence, defined by the same TPs as those in the habituation stimulus, and a pseudorandom sequence (no shape repetitions). Infants at 2, 5, and 8 months looked reliably longer at the random sequence, interpreted by Kirkham et al. as showing sensitivity to the statistical properties of the input—the TPs defining shape pairs in the habituation sequence—and noted when these statistics were violated. Infants at all three ages showed a reliable novelty preference for the random pattern, with no significant age differences aside from longer looking in general by the youngest group.

My colleagues and I then asked if visual SL may be available at birth with a replication of the Kirkham et al. (2002) methods, modified to suit newborns' limited color vision with the use of monochromatic stimuli (Bulf, Johnson, & Valenza, 2011). Newborn infants provided no evidence of discriminating random from struc-

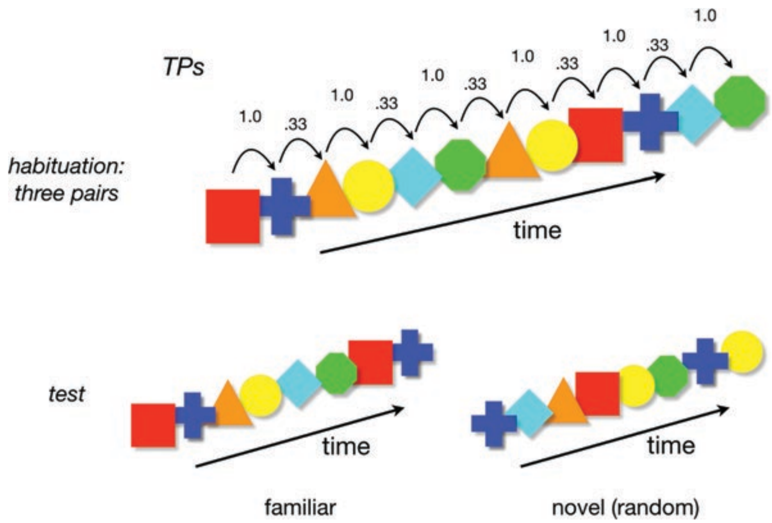


Fig. 2.2 Schematic depiction of habituation and test displays testing for visual statistical learning in infants (Kirkham et al., 2002)

tured six-item sequences. We reasoned that three pairs of shapes (a *high-demand* condition; e.g., ABCDEFCDABEFABCDABEF…) might overwhelm newborns' ability to track probabilities. This hypothesis was addressed with a two-pair, *low-demand* condition (e.g., ABCDCDABCDABABCDAB…). We observed a novelty preference for the random sequence in the low-demand condition, as did the older infants observed by Kirkham et al. who were tested with a high-demand condition. Thus the Bulf et al. study provides evidence that visual SL may be constrained by newborns' limited cognitive resources, perhaps preventing identification of relevant visual information necessary to detect statistical structure.

The Bulf et al. (2011) and Kirkham et al. studies (2002) documented young infants' ability to detect statistical information in sequences of discrete, looming shapes. By 8 months, infants detect probabilistic patterns in spatiotemporal visual sequences in which shapes appeared in locations defined by TPs of 1.0 or 0.33 (Kirkham, Slemmer, Richardson, & Johnson, 2007), and by 9 months, infants encode the underlying spatial statistical structure of multiple-element scenes in which shapes were arranged in groups defined by conditional probabilities among individual items (Fiser & Aslin, 2002b). These results led to claims of a domain-general SL device that is available early and operates across modalities, across time and space, and across species, suggesting that SL might be a predisposed, general associative mechanism (Kirkham et al., 2002). This hypothesis is supported by reports of SL and discrimination of visual and linguistic sequences in newborns (Bulf et al., 2011; Teinonen, Fellman, Näätänen, Alku, & Huotilainen, 2009), constituting evidence for sensitivity to statistical information at birth in at least two modalities (vision and audition). SL is now a central feature of recent theories of human perception, cognition, and development (e.g., Aslin & Newport, 2012, 2014; Hasson, 2016; Krogh, Vlach, & Johnson, 2013; Thiessen, 2016; Turk-Browne, 2012).

Kinds of Statistical Structure Infants Are Able to Learn

As noted previously, early studies of SL were aimed largely at questions of (a) whether infants might use SL to segment continuous speech into discrete units (Aslin et al., 1998; Saffran et al., 1996) and (b) the possibility that infants' SL might be a domain-general learning mechanism (Kirkham et al., 2002; Saffran et al., 1999). These studies examined SL with methods involving a learning (familiarization or habituation) phase with streams of unsegmented auditory or visual sequences, followed by a test phase probing for recognition of clusters of items that were either high or low in TPs. Results were taken to indicate that SL in infancy was domain-general and innate: that is, SL was proposed to operate across multiple kinds of sensory inputs and available from birth (Kirkham et al., 2002). Yet only the Aslin et al. study was designed to rule out other kinds of statistical information, such as frequency, in favor of TPs. The Kirkham et al. study did not test for infants' TP learning or segmentation: Instead, infants were tested for simple discriminations of

TP-structured sequences vs. pseudorandom sequences. This kind of discrimination was later discovered to occur even without a learning phase: Five-month-olds were tested with two six-shaped visual sequences, seen in alternation, that either were ordered randomly or followed the statistical structure described previously. Interestingly, the infants looked longer at random vs. structured sequences of visual shapes, even without prior familiarization or habituation (Addyman & Mareschal, 2013), thus demonstrating a spontaneous preference for random sequences that does not require prior experience or learning. This implies that infants in the Kirkham et al. and Bulf et al. (2011) experiments did not necessarily learn anything during the experiment, calling into question the likelihood that SL operates from birth and undergoes little developmental change in infancy.

The Addyman and Mareschal (2013) results also imply that young infants can discriminate sequences solely from *ordinal* information—the orderings of items. Ordinal information, like TP information, is a kind of statistic, but recognition of ordinal violations may be less demanding than recognition of TP violations, especially when infants are also required to segment an input stream into units. Consistent with this possibility, infants as young as 3 months were reported to identify violations of serial order in audiovisual sequences (Lewkowicz, 2008); in contrast, 4.5- and 5-month-olds, but not younger infants, segmented visual sequences from TP differences (Marcovitch & Lewkowicz, 2009; Slone & Johnson, 2015). These studies highlight an important distinction between *discrimination* of different sequences based on statistical information and *learning* statistical information to segment sequences of items into clusters or units. The studies also highlight the distinction between different statistics that might be identified and/or learned. Furthermore, the Addyman and Mareschal results are important in demonstrating that infants' preferences for items in sequence might stem from differences in complexity (cf. Kidd, Piantadosi, & Aslin, 2012, 2014; Tummeltshammer & Kirkham, 2013).

Other kinds of inputs have been examined in infant SL tasks. For example, by 11 months, infants can learn probabilistic sequences of items appearing in predictable spatial locations, and 8-month-olds can learn spatiotemporal sequences when item location combines with color and shape cues (Kirkham et al., 2007; cf. Sobel & Kirkham, 2006; Tummeltshammer & Kirkham, 2013); 5-month-olds tested under identical conditions did not appear sensitive to spatial information for the sequence. Infants at 8 months also were reported to learn predictable co-occurrences of items in visual arrays, akin to TPs between items in sequence (Fiser & Aslin, 2002b), and at 9 months, infants' SL of object features in visual arrays was facilitated by a social cue: a woman seen to be looking in the location where a coherent configuration was displayed (Wu, Gopnik, Richardson, & Kirkham, 2011). Also, other cues to segmentation that are present in real-world speech, such as prosody (Thiessen & Saffran, 2003) and word length (Lew-Williams, Pelucchi, & Saffran, 2011; Lew-Williams & Saffran, 2012), interact with, and constrain, SL of speech sounds.

Finally, there have been claims that infant SL has an important role in development of abstract "rule learning," a kind of pattern learning involving identification of simple reduplicative patterns and generalization of the pattern to new items (e.g., Gerken, 2006; Marcus, Vijayan, Rao, & Vishton, 1999), and an important foundation for analogical reasoning (see Chap. 5). Infants' learning and generalization of simple abstract rules in sequential patterns were first investigated by Marcus et al. (1999), who exposed 7-month-old infants to strings consisting of computer-generated speech. In their first experiment, strings followed either an "ABA" pattern (e.g., *gah tee gah*, *nee lah nee*) or an "ABB" pattern (e.g., *gah tee tee*, *nee lah lah*). A and B items were separated by 250 ms of silence, strings by 1 s of silence. The speech stream was accompanied by a flashing light, mounted centrally in the testing chamber. After 2 minutes of continuous repetitions of one of these two familiarization patterns, the infants received trials of the same (familiar) pattern instantiated by different phonemes (e.g., *woh fei woh*, *dee koh dee*) and the second (novel) pattern on alternating trial, from a speaker located either to the left or right of the infant. Each kind of test stimulus was also accompanied by a flashing light, located either left or right, and learning was operationalized in terms of differences in looking time toward the light when the word or part-word was heard. The infants exhibited a reliable preference for the novel pattern, a result that extended to a test of ABB vs. AAB. The balance of phonetic features across familiarization and test stimuli ruled out the possibility that performance was based on learning sequences of low-level cues (such as voiced vs. unvoiced consonants). Importantly, the positive outcome of the ABB/AAB comparison obviated an account based on learning a simple reduplication pattern (i.e., adjacent repetition) without respect to its place in sequence (i.e., initial/final edge position).

The Marcus et al. (1999) task bears superficial similarities to the Saffran et al. (1996) task: Infants listened to a structured speech stream for 2 minutes, and they were tested for recognition of the underlying pattern using a head-turn method to generate preferences for a flashing light on one vs. the other side of a testing chamber. Yet there is a vital difference in what is tested in these two paradigms. In SL tasks such as the Saffran et al. study, infants are asked to segment a speech stream into units that are bounded by dips in TPs: that is, the words heard at test (now segmented) had higher internal TPs than nonwords or part-words. In abstract rule-learning tasks such as the Marcus et al. study, in contrast, infants are not required to segment the input (it is already segmented into units) nor are they required to recognize correspondences among items, learned during familiarization, to the same items at test. This is because no items from familiarization were heard at test. Instead, infants were required to learn an *abstract* pattern that, as noted previously, was *independent of surface features* (such as vowels and consonants).

Nevertheless, there have been proposals for a common mechanism supporting infant SL and abstract rule learning (see Chap. 5 for additional discussion), perhaps because (a) language experience facilitates both SL (e.g., Saffran & Wilson, 2003) and abstract rule learning (Marcus, Fernandes, & Johnson, 2007), (b) simple connectionist models can explain both sets of results (e.g., Christiansen & Curtin, 1999), (c) simple reduplications may comprise a "perceptual primitive" as a basis for pattern extraction (e.g., Gerken, Dawson, Chatila, & Tenenbaum, 2015; Gómez & Gerken, 2000), or (d) abstract categories can arise from purely statistical input (Aslin & Newport, 2012, 2014; see Reeder, Newport, & Aslin, 2013 for evidence from adults). However, to my knowledge, there are no reports of any direct demonstrations in infants that SL and abstract rule learning stem from a single learning mechanism. Indeed, experiments in my lab designed to test SL and abstract rule learning from identical four-item audiovisual sequences found that 11-month-olds could learn about specific items and their positions in sequence—that is, statistical information, in this case order of items in a series. In contrast, the infants did not appear to learn a simple reduplication—that is, an abstract rule that was independent of surface features (Schonberg, Marcus, & Johnson, 2018; see Fig. 2.3).

In summary, studies of SL in infancy have tended to focus on infants' learning of TPs in segmentation tasks. Other kinds of statistical information are also available (ordinal information, frequency, repetition, linguistic cues), but their roles in segmenting and learning, and their interactions with TPs between stimulus features as contributions to learning, are not well understood at present.

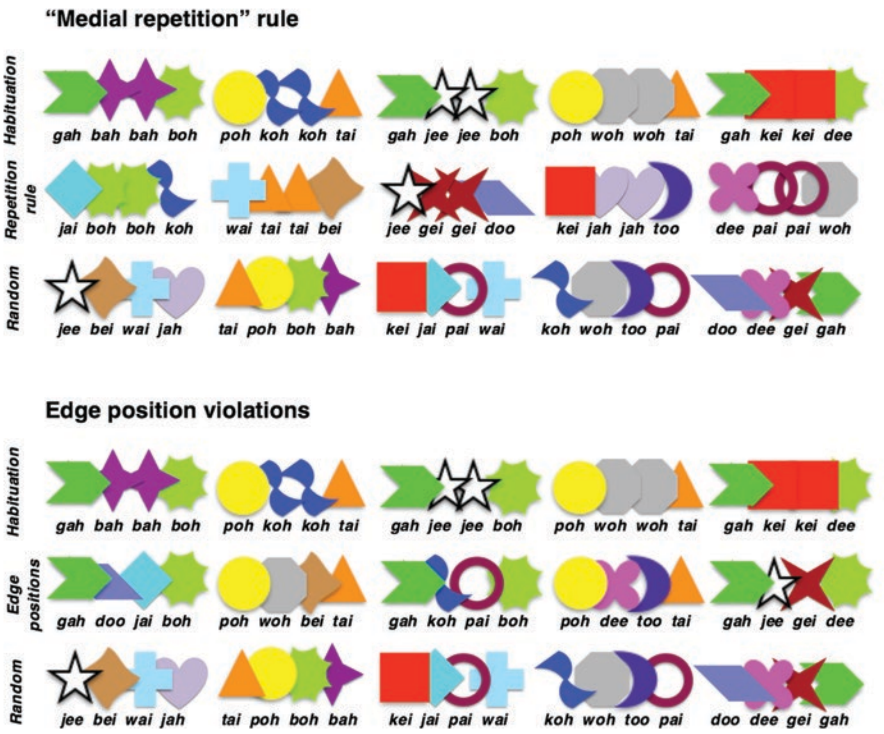


Fig. 2.3 Schematic depiction of stimuli used to test for infants' abstract rule learning, a "medial repetition rule" (top panel), and statistical learning, the specific positions of items in their ordinal positions (bottom panel). Each condition used identical habituation stimuli but tested for learning of either an abstract pattern or one based on items in sequence. Eleven-month-olds appeared to learn edge position violations, but not the abstract repetition rule. (Adapted from Schonberg et al. (2018))

Testing Methods

The majority of published infant SL studies have employed a learning phase (familiarization or habituation) followed by a test phase in which infants are observed for evidence of segmentation of continuous input, undifferentiated except by virtue of TP differences among adjacent items, and recognition of parsed units vs. foil stimuli consisting of a reordering of individual items (see Saffran & Kirkham, 2018, for review). Effects of variations in testing methods, such as the use of different stimuli in the same paradigm, are not well understood (see Chap. 4 for further discussion), but there is some evidence that their investigation can be fruitful (Kirkham et al., 2007). For example, Lewkowicz (2004) examined infants' detection of violations of serial order of items in sequence and found that ordinal information was more readily identified in sequences of linearly moving objects than looming objects presented in a single location (as in the Kirkham et al., 2002, method).

Eye-tracking and brain-based methods have provided complementary and, in some cases, unique insights into infants' SL. Eye-tracking methods involve records of infants' point of gaze as they view displays on a monitor (Gredebäck, Johnson, & von Hofsten, 2010). SL studies have examined eye movement (saccadic) latencies to items in sequence, the prediction being that spatial locations of more predictable items, by virtue of high TPs between items, will be fixated more quickly. As noted previously, evidence in support of this prediction was provided by the Kirkham et al. (2007) experiment in which infants were found to look toward locations in which a predictable item appeared vs. one of the other five locations on the display. More recently, Tummeltshammer and Kirkham (2013) examined 8-month-olds' saccadic latencies when viewing six-location visual arrays with sequences of spatiotemporal events. Arrays resembled a house or storefront with windows in which shapes appeared one at a time in a probabilistic sequence comprising three shape pairs. Each shape appeared in a particular window, disappeared, and subsequently reappeared in a different window according to its assigned probability. Items appeared in sequence with TPs of 1.0, 0.75, or 0.5, and one group of infants viewed arrays with additional competing visual distracters. Items with higher TPs were attended more often and with fewer errors (i.e., predictive looks) overall, and this effect interacted with the presence of distracters: With no distraction, latencies were fastest to high-probability (0.75) TP events, but with distracters, latencies were fastest to "deterministic" events with TPs of 1.0. These findings suggest that infants' SL guides predictive behavior and that predictions are influenced by distributional properties of the entire scene, even events (distracters) unrelated to the predictable items.

Brain-based methods have been used to examine cortical loci of SL with functional MRI under various testing conditions in adults (e.g., Karuza et al., 2013; Lieberman, Chang, Chiao, Bookheimer, & Knowlton, 2004; Turk-Browne, Scholl, Chun, & Johnson, 2009) and children (McNealy, Mazziotta, & Dapretto, 2011). Electrophysiological methods, in particular event-related potentials (ERPs), have yielded evidence concerning the time course of "online" learning in adults from changes in the timing and strength of electrical cortical potentials (viz., ERP components) recorded at the scalp (e.g., Abla, Katahira, & Okanoya, 2008; Abla & Okanoya, 2009). ERPs have been used as an index of differences in visual SL between children with autism spectrum disorder (ASD) and typically developing children and have revealed impairments in some children with ASD (Jeste et al., 2015). ERP methods are more feasible for use with young populations relative to fMRI, and they have been used to examine SL in infants. For example, Teinonen et al. (2009) observed ERP differences to statistically structured vs. unstructured speech sequences in sleeping neonates, and Marin et al. (2019) observed ERP differences during a visual SL task between 3-month-old infants at elevated risk for ASD (due to high genetic load) and low-risk infants. The Jeste et al. and Marin et al. studies are discussed in more detail in the next section.

In sum, eye-tracking and brain-based methods, in particular electrophysiological methods, require specialized designs and equipment but can provide particularly sensitive measures of SL. This can be especially important for infant studies. Infants' control of eye movements is well-established even at birth (Gredebäck et al., 2010), and clever research designs can exploit infants' tendency to explore novel scenes and learn contingencies among events, including probabilistic events. ERPs, likewise, can be used in infants at all ages (de Haan, 2007) and can reveal cortical activity in response to probabilistic events that more overt behaviors cannot necessarily reveal.

Implications of Infant SL for Cognitive Development and Developmental Disabilities

There is extensive evidence that SL is related to and perhaps facilitates language acquisition (see Romberg & Saffran, 2010, for review). In 8-month-olds, for example, nonsense words acquired via SL are treated as "candidate" words when embedded in new linguistic contexts (Saffran, 2001); moreover, SL provides candidate words that can become associated with novel objects at 17 months (Graf Estes, Evans, Alibali, & Saffran, 2007) and with novel object categories at 22 months (Lany & Saffran, 2010). In addition, 8.5-month-olds' performance on a visual SL task was correlated with the infants' vocabulary size, assessed by parental report (Shafto, Conway, Field, & Houston, 2011). Six-month-old infants' oculomotor responses to events in a visual pattern-learning task predicted vocabulary size 16 months later (Ellis, Gonzalez, & Deák, 2014), and 6- to 8-year-olds' visual SL performance predicted their comprehension of native-language syntax (Kidd & Arciuli, 2016). However, measures of cognitive development more broadly (i.e., independent of language), such as performance on the Bayley Scales (Bayley, 2005), general IQ, and gesture comprehension, were not related to SL performance (Ellis et al., 2014; Kidd & Arciuli, 2016; Shafto et al., 2011).

Evidence for how SL might affect developmental disabilities is consistent with these findings: SL is related to language acquisition and performance but may have somewhat less impact on cognitive function. For example, the possibility that SL is impaired in ASD has received mixed support. Some studies report impaired SL (e.g., Jeste et al., 2015; Scott-Van Zeeland et al., 2010), but others report little or no impairment (Mayo & Eigsti, 2012) or even *enhanced* SL in ASD (Roser, Aslin, McKenzie, Zahra, & Fiser, 2015). ASD, however, is a heterogeneous disorder that remains poorly understood at the level of individual differences (Jeste et al., 2015), and notably, these studies of SL in ASD used varying methods and tested different populations (e.g., children with unknown symptom severity vs. high-functioning adults), making direct comparisons of results difficult. In infants with Williams syndrome, a developmental disorder characterized by strong language skills but impaired intellectual capacity, SL seems to be intact (Cashon, Ha, Graf Estes, Saffran, & Mervis, 2016). A recent meta-analysis found strong and consistent evidence for reduced SL in individuals with specific language impairment but mixed evidence for reduced SL in individuals with ASD (Obeid, Brooks, Powers, Gillespie-Lynch, & Lum, 2016).

Recently, Jeste et al. (2015) investigated ERP correlates of SL in children with ASD vs. typically developing controls. ERP was recorded as children watched streams of looming shapes, similar to methods described previously with infants (Kirkham et al., 2002), and after a learning phase, they introduced a violation of the expected sequence by showing an unexpected shape. This study revealed two important findings. First, the ASD group showed attenuated evidence of SL in two ERP components: a reduced "N1" component, which was theorized to signify early visual recognition, akin to the N100 in adults (Coull, 1998) and a reduced P300 component, which represents attention to salient information and probabilities of a target stimulus (Picton, 1992). Second, analyses of individual differences in the ASD group revealed a positive correlation between N1 amplitude difference and nonverbal IQ and a positive correlation between P300 amplitude difference and adaptive social function. This study demonstrates, therefore, that ASD is highly variable among individuals, and variability in learning capacity may help explain deficits in social, and perhaps cognitive, function.

In infants, my colleagues and I recently recorded ERPs in 3-month-old infants at elevated or low risk for ASD, due to the presence (or not) of one or more close family members having received a diagnosis of ASD (Marin et al., 2019). We asked whether visual SL at 3 months, recorded as described previously for the Jeste et al. (2015) experiment, might predict cognitive function and ASD symptoms at 18 months. Interestingly, higher-risk infants demonstrated increased neural responses (late slow wave and N700 components) to the probabilistic event, whereas low-risk infants demonstrated increased neural responses to the deterministic (expected) event. Moreover, individual differences in these ERP components at 3 months predicted visual reception ability and ASD symptoms at 18 months of age. The reasons for these differences so early in infancy are not yet clear, but the potential predictive value for emerging ASD symptoms from such observations may be an important finding.

Mechanisms Underlying Statistical Learning in Infancy

As noted in prior sections, SL is a powerful means by which infants learn about a structured environment, and studies of SL can be particularly informative about learning in children with developmental disabilities. Yet the specific processes underlying SL remain unclear. Recently, research in my lab (Slone & Johnson, 2018) investigated two types of models underlying statistical learning: "statistical" (or "transition-finding") and "chunking" (or "clustering") models that have been proposed to account for SL in adults (Thiessen, Kronstein, & Hufnagle, 2013).

The goal of both statistical and chunking models is to account for sensitivity to sequential structure and the units that are learned, but they differ in the proposed representations stored in memory. Statistical or TP-learning models can be instantiated in computational models known as simple recurrent networks (e.g., Elman, 1990) that compute and represent statistical relations between items, such as TPs, in memory. Representing TPs informs the model of the likelihood of two items occurring together and allows the model to predict individual items based on previous items in a sequence. In the syllable stream used by Saffran et al. (1996), for example (Fig. 2.1), the model would learn that the probability of *pi* after *tu* and the probability of *ro* after *pi* are high, because items *tu*, *pi*, and *ro* always appear in order (in the word *tupiro*). The probability of *pa* after *ro*, in contrast, will be lower because *padoti* follows *tupiro* only 1/3 of the time in the familiarization sequence. In this way, statistical models can distinguish statistically coherent units of information contained within a sequence (e.g., *tupiro*) from less coherent units like part- words (e.g., *ropado*). Importantly, statistical models do not explicitly represent statistically coherent units (e.g., words); rather, they represent statistical relations between items (e.g., syllables) as TPs.

Chunking models, in contrast, represent statistically coherent units of information in memory. One such computational model, the "truncated recursive autoassociative chunk extractor" (TRACX), forms groupings simply by joining items that tend to co-occur (French, Addyman, & Mareschal, 2011; Mareschal & French, 2017). Groupings, or chunks, become single units that can be stored in memory. Representations of units whose component items co-occur regularly are progressively strengthened in memory, whereas representations of units whose component items do not co-occur regularly are forgotten. In the Saffran et al. (1996) sequence, for example, the model could initially capture the sequence *tupiropadoti* in three separate chunks: *tupi*, *ropa*, and *doti*. Over time, chunks *tupi* and *doti* will be reinforced in memory because their component items always co-occur. In contrast, chunk *ropa* will only be weakly represented because its component items co-occur less frequently. Moreover, once the sequences *tupi* and *piro* become represented as single chunks, it becomes possible for *tupiro* to be captured as an even larger chunk (i.e., as the aggregate of *tupi* and *piro*). Thus, with sufficient exposure, the model will form strong representations of statistically coherent units of information (e.g., *tupiro*) and distinguish them from weakly represented part-words (e.g., *ropado*). Statistical relations among items are not retained in memory over time—only the chunks.

Several studies have suggested that adults' SL is best accounted for by chunking models (Fiser & Aslin, 2002a; Giroux & Rey, 2009; Orbán, Fiser, Aslin, & Lengyel, 2008; Perruchet & Poulin-Charronnat, 2012), but others have provided evidence that statistical TP-learning models may often provide a better fit for adult performance in SL tasks (Endress & Langus, 2017; Endress & Mehler, 2009). Moreover, it remains unknown which type of model best accounts for infants' SL performance.

We addressed the question of whether statistical or chunking was the best account of infant sequence learning in three experiments with 8-month-olds (Slone & Johnson, 2018). In the first experiment, infants were familiarized with five-item sequences for 5 minutes. Sequences were constructed such that certain items were shared across units (see Fig. 2.4a). Following habituation, infants were tested for recognition of a familiar triplet (tantamount to a word in the Saffran et al., 1996,

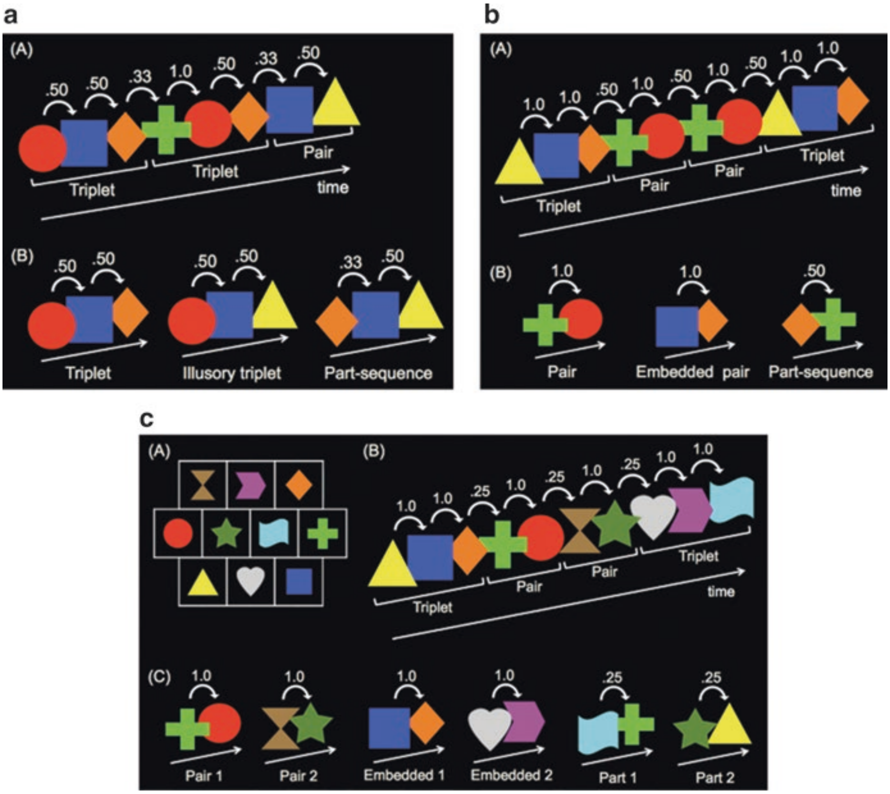


Fig. 2.4 Schematic depiction of familiarization and test sequences in experiments testing statistical vs. chunking models (see text for details). Numbers above adjacent shapes represent TPs during familiarization. Familiarization sequences are seen at the top in each panel and test sequences at the bottom. Brackets below shapes indicate the unit structure of the familiarization sequences. (**a**) Illusory triplet, (**b**) embedded pair, (**c**) embedded pair with increased exposure. (Adapted from Slone and Johnson (2018))

study), a part-sequence (triplet), and an "illusory" triplet, composed of two pairs of items that had high TPs but had not been seen together. We reasoned that if infants had learned a chunk (the triplet) during familiarization, they would recognize the triplet when seen in isolation at test, but not the illusory triplet or the part-sequence. If infants recognized the illusory triplet, however, this would support statistical models, because the TPs of the familiar and illusory triplets were identical. The first prediction was supported, in line with chunking models.

In the second experiment, infants were familiarized with five-item sequences composed of one unique triplet and one unique pair (no shared items; see Fig. 2.4b). At test, infants viewed a familiar pair, a part-sequence (pair), and an "embedded pair," composed of items that were part of the triplet. We reasoned that infant looking at test would reveal whether they formed a triplet chunk that excluded the embedded pair, consistent with chunking models: recognition of the familiar pair but not the part-sequence or embedded pair. This prediction was also supported, again in line with chunking models.

Finally, in a third experiment, we asked if we might capture a point in time during familiarization when infants had learned TPs among adjacent items but not yet formed full chunks. We did this with a condition testing for recognition of embedded pairs, as in the previous experiment, but now employing twice the numbers of items and units: two unique triplets and two unique pairs, comprising 10 items in total (see Fig. 2.4c). Exposure time was kept the same, however, requiring infants to track more relations among items and thus perhaps impairing chunk formation. In support of this prediction and in contrast to the second experiment, infants in the third study appeared to recognize both familiar and embedded pairs, evidence that infants learned TPs between adjoining items, but exposure time had been insufficient for learning chunks of triplets. Taken together, these results inform the nature of infants' SL: As a first step in sequence learning, TPs between items are acquired, and then chunks are learned from the accumulation of TP-linked pairs. But whether TPs are immediately discarded may depend on the learning requirements in context (cf. Endress & Langus, 2017).

Conclusions and Broader Implications

Results of statistical learning studies can provide important constraints for theories of cognitive development, in particular computational models of associative learning in developmental disorders (Tovar, Westermann, & Torres, 2018), cross-situational/multimodal computational models of language acquisition (Monaghan, 2017), and Bayesian computational models of category learning (Tenenbaum, Kemp, Griffiths, & Goodman, 2011). Yet many models of infant cognition do not take account of possible effects of stimulus modality on learning or possible constraints in infant attention, memory, and learning capacity (e.g., Franz & Triesch, 2010; Rogers, Rakison, & McClelland, 2004; Tenenbaum et al., 2011).

In sum, much remains to be discovered with respect to infants' SL, despite important progress in our understanding of SL as a vital part of language acquisition and as a window into the nature of some developmental disabilities. For example, links between infants' SL and abstract rule learning remain unexplored but may involve comparison processes between items and relations (see Chap. 5). In addition, neural processes that give rise to statistical learning are becoming understood as interactions between the declarative and nondeclarative memory systems of the brain (Batterink, Paller, & Reber, 2019), but little is known about how these interactions develop early in life. Nor is the developmental time course of SL in individuals well understood (Siegelman & Frost, 2015). Finally, an important question concerns the role of SL in infants' learning of real-world events. For example, when children begin to learn relations between objects, these may become chunked into a unit of "causal action" and associated with a label (e.g., a verb). Evidence for this and other possible contributions of SL to cognitive development await future study.

References

Abla, D., Katahira, K., & Okanoya, K. (2008). On-line assessment of statistical learning by event-related potentials. *Journal of Cognitive Neuroscience, 20*, 952–964. https://doi.org/10.1162/jocn.2008.20058

Abla, D., & Okanoya, K. (2009). Visual statistical learning of shape sequences: An ERP study. *Neuroscience Research, 64*, 185–190. https://doi.org/10.1016/j.neures.2009.02.013

Addyman, C., & Mareschal, D. (2013). Local redundancy governs infants' spontaneous orienting to visual-temporal sequences. *Child Development, 84*, 1137–1144. https://doi.org/10.1111/cdev.1206

Aslin, R. N., & Newport, E. L. (2012). Statistical learning: From acquiring specific items to forming general rules. *Current Directions in Psychological Science, 21*, 170–176. https://doi.org/10.1177/0963721412436806

Aslin, R. N., & Newport, E. L. (2014). Distributional language learning: Mechanisms and models of category formation. *Language Learning, 64: Cognitive Neuroscience Supplement, 2*, 86–105. https://doi.org/10.1111/lang.12074

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. https://doi.org/10.1111/1467-9280.00063

Baker, C. I., Olson, C. R., & Behrmann, M. (2004). Role of attention and perceptual grouping in visual statistical learning. *Psychological Science, 15*, 460–466. https://doi.org/10.1111/j.0956-7976.2004.00702.x

Batterink, L. J., Paller, K. A., & Reber, P. J. (2019). Understanding the neural bases of implicit and statistical learning. *Topics in Cognitive Science, 11*, 482–503. https://doi.org/10.1111/tops/12420

Bayley, N. (2005). *Bayley scales of infant and toddler development-III*. San Antonio, TX: Pearson.

Bulf, H., Johnson, S. P., & Valenza, E. (2011). Visual statistical learning in the newborn infant. *Cognition, 121*, 127–132. https://doi.org/10.1016/j.cognition.2011.06.010

Cashon, C. H., Ha, O. R., Graf Estes, K., Saffran, J. R., & Mervis, C. B. (2016). Infants with Williams syndrome detect statistical regularities in continuous speech. *Cognition, 154*, 165–168. https://doi.org/10.1016/j.cognition.2016.05.009

Chomsky, N. A. (1965). *Aspects of a theory of syntax*. Cambridge, MA: MIT Press.

Christiansen, M. H., & Curtin, S. (1999). Transfer of learning: Rule acquisition or statistical learning? *Trends in Cognitive Sciences, 3*, 289–290. https://doi.org/10.1016/S1364-6613(99)01356-X

Coull, J. T. (1998). Neural correlates of attention and arousal: Insights from electrophysiology, functional neuroimaging and psychopharmacology. *Progress in Neurobiology, 55*, 343–361. https://doi.org/10.1016/S0301-0082(98)00011-2

De Haan, M. (2007). *Infant EEG and event-related potentials*. London, UK: Psychology Press.

Ellis, E. M., Gonzalez, M. R., & Deák, G. O. (2014). Visual prediction in infancy: What is the association with later vocabulary? *Language Learning and Development, 10*, 36–50. https://doi.org/10.1080/15475441.2013.799988

Elman, J. L. (1990). Finding structure in time. *Cognitive Science, 14*, 179–211. https://doi.org/10.1207/s15516709cog1402_1

Endress, A., & Mehler, J. (2009). The surprising power of statistical learning: When fragment knowledge leads to false memories of unheard words. *Journal of Memory and Language, 60*, 351–367. https://doi.org/10.1016/j.jml.2008.10.003

Endress, A. D., & Langus, A. (2017). Transitional probabilities count more than frequency, but might not be used for memorization. *Cognitive Psychology, 92*, 37–64. https://doi.org/10.1016/j.cogpsych.2016.11.004

Fiser, J., & Aslin, R. N. (2002a). Statistical learning of higher-order temporal structure from visual shape sequences. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28*, 458–467. https://doi.org/10.1037//0278-7393.28.3.458

Fiser, J., & Aslin, R. N. (2002b). Statistical learning of new visual feature combinations by infants. *Proceedings of the National Academy of Sciences (USA), 99*, 15822–15826. https://doi.org/10.1073/pnas.232472899

Fougeron, C., & Keating, P. A. (1997). Articulatory strengthening at edges of prosodic domains. *Journal of the Acoustical Society of America, 101*, 3728–3740. https://doi.org/10.1121/1.418332

Franz, A., & Triesch, J. (2010). A unified computational model of the development of object unity, object permanence, and occluded object trajectory perception. *Infant Behavior and Development, 33*, 635–653. https://doi.org/10.1016/j.infbeh.2010.07.018

French, R. M., Addyman, C., & Mareschal, D. (2011). TRACX: A recognition-based connectionist framework for sequence segmentation and chunk extraction. *Psychological Review, 118*, 614–636. https://doi.org/10.1037/a0025255

Frost, R., Armstrong, B. C., Siegelman, N., & Christiansen, M. H. (2015). Domain generality versus modality specificity: The paradox of statistical learning. *Trends in Cognitive Sciences, 19*, 117–125.

Gerken, L. (2006). Decisions, decisions: Infant language learning when multiple generalizations are possible. *Cognition, 98*, B67–B74. https://doi.org/10.1016/j.cognition.2005.03.003

Gerken, L., Dawson, C., Chatila, R., & Tenenbaum, J. (2015). Surprise! Infants consider possible bases of generalization for a single input example. *Developmental Science, 18*, 80–89. https://doi.org/10.1111/desc.12183

Gibson, E. J., & Gibson, J. J. (1955). Perceptual learning: Differentiation or enrichment? *Psychological Review, 62*, 32–41. https://doi.org/10.1037/h0048826

Giroux, I., & Rey, A. (2009). Lexical and sublexical units in speech perception. *Cognitive Science, 33*, 260–272. https://doi.org/10.1111/j.1551-6709.2009.01012.x

Gómez, R. L., & Gerken, L. (2000). Infant artificial language learning and language acquisition. *Trends in Cognitive Sciences, 4*, 178–186. https://doi.org/10.1016/S1364-6613(00)01467-4

Graf Estes, K., Evans, J. L., Alibali, M. W., & Saffran, J. R. (2007). Can infants map meaning to newly segmented words? Statistical segmentation and word learning. *Psychological Science, 18*, 254–260. https://doi.org/10.1111/j.1467-9280.2007.01885.x

Gredebäck, G., Johnson, S. P., & von Hofsten, C. (2010). Eye tracking in infancy research. *Developmental Neuropsychology, 35*, 1–19. https://doi.org/10.1080/87565640903325758

Hasson, U. (2016). The neurobiology of uncertainty: Implications for statistical learning. *Philosophical Transactions of the Royal Society B, 372*, 20160048. https://doi.org/10.1098/rstb.2016.0048

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. https://doi.org/10.1016/S0010-0277(00)00132-3

Hunt, R. H., & Aslin, R. N. (2001). Statistical learning in a serial reaction time task: Access to separable statistical cues by individual learners. *Journal of Experimental Psychology: General, 130*, 658–680. https://doi.org/10.1037//0096-3445.130.4.658

Jeste, S. S., Kirkham, N., Senturk, D., Hasenstab, K., Sugar, C., Kupelian, C., et al. (2015). Electrophysiological evidence of heterogeneity in visual statistical learning in young children with ASD. *Developmental Science, 18*, 90–105. https://doi.org/10.1111/desc.12188

Johnson, M. J., & Morton, J. (1991). *Biology and cognitive development: The case of face recognition*. Oxford: Blackwell.

Karuza, E. A., Newport, E. L., Aslin, R. N., Starling, S. J., Tivarus, M. E., & Bavelier, D. (2013). The neural correlates of statistical learning in a word segmentation task: An fMRI study. *Brain and Language, 127*, 46–54. https://doi.org/10.1016/j.bandl.2012.11.007

Kidd, C., Piantadosi, S. T., & Aslin, R. N. (2012). The Goldilocks effect: Human infants allocate attention to visual sequences that are neither too simple nor too complex. *PLoS One, 7*, e36399. https://doi.org/10.1371/journal.pone.0036399

Kidd, C., Piantadosi, S. T., & Aslin, R. N. (2014). The Goldilocks effect in infant auditory attention. *Child Development, 85*, 1795–1804. https://doi.org/10.1111/cdev.12263

Kidd, E., & Arciuli, J. (2016). Individual differences in statistical learning predict children's comprehension of syntax. *Child Development, 87*, 184–193. https://doi.org/10.1111/cdev.12461

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. https://doi.org/10.1016/S0010-0277(02)00004-5

Kirkham, N. Z., Slemmer, J. A., Richardson, D. C., & Johnson, S. P. (2007). Location, location, location: Development of spatiotemporal sequence learning in infancy. *Child Development, 78*, 1559–1571. https://doi.org/10.1111/j.1467-8624.2007.01083.x

Krogh, L., Vlach, H. A., & Johnson, S. P. (2013). Statistical learning across development: Flexible yet constrained. *Frontiers in Psychology, 3*, 598. https://doi.org/10.3389/fpsyg.2012.00598

Lany, J., & Saffran, J. R. (2010). From statistics to meaning: Infants' acquisition of lexical categories. *Psychological Science, 21*, 284–291. https://doi.org/10.1177/0956797609358570

Leslie, A. M. (1997). Pretense and representation: The origins of "theory of mind". *Psychological Review, 94*, 412–426. https://doi.org/10.1037/0033-295X.94.4.412

Lewkowicz, D. J. (2004). Perception of serial order in infants. *Developmental Science, 7*, 175–184. https://doi.org/10.1111/j.1467-7687.2004.00336.x

Lewkowicz, D. J. (2008). Perception of dynamic and static audiovisual sequences. *Child Development, 79*, 1538–1554. https://doi.org/10.1111/j.1467-8624.2008.01204.x

Lew-Williams, C., Pelucchi, B., & Saffran, J. R. (2011). Isolated words enhance statistical language learning in infancy. *Developmental Science, 14*, 1323–1329. https://doi.org/10.1111/j.1467-7687.2011.01079.x

Lew-Williams, C., & Saffran, J. R. (2012). All words are not created equal: Expectations about word length guide infant statistical learning. *Cognition, 122*, 241–246. https://doi.org/10.1016/j.cognition.2011.10.007

Lieberman, M. D., Chang, G. Y., Chiao, J., Bookheimer, S. Y., & Knowlton, B. J. (2004). An event-related fMRI study of artificial grammar learning with a balanced chunk strength design. *Journal of Cognitive Neuroscience, 16*, 427–438. https://doi.org/10.1162/089892904322926764

Marcovitch, S., & Lewkowicz, D. J. (2009). Sequence learning in infancy: The independent contributions of conditional probability and pair frequency information. *Developmental Science, 12*, 1020–1025. https://doi.org/10.1111/j.1467-7687.2009.00838.x

Marcus, G., Fernandes, K., & Johnson, S. (2007). Infant rule learning facilitated by speech. *Psychological Science, 18*, 387–391. https://doi.org/10.1111/j.1467-9280.2007.01910.x

Marcus, G. F., Vijayan, S., Rao, S. B., & Vishton, P. M. (1999). Rule learning by seven-month- old infants. *Science, 283*, 77–80. https://doi.org/10.1126/science.283.5398.77

Mareschal, D., & French, R. M. (2017). TRACX2: A connectionist autoencoder using graded chunks to model infant visual statistical learning. *Philosophical Transactions of the Royal Society B, 372*, 20160057. https://doi.org/10.1098/rstb.2016.0057
Marin, A., Hutman, T., Ponting, C., McDonald, N. M., Dickinson, A., Dapretto, M., et al. (2019). *Electrophysiological signatures of visual statistical learning in three-month old infants at familial and low risk for autism spectrum disorder*. Manuscript submitted for publication.
Mayo, J., & Eigsti, I. M. (2012). A comparison of statistical learning in school-aged children with high functioning autism and typically developing peers. *Journal of Autism and Developmental Disorders, 42*, 2476–2485. https://doi.org/10.1007/s10803-012-1493-0
McNealy, K., Mazziotta, J. C., & Dapretto, M. (2011). Age and experience shape developmental changes in the neural basis of language-related learning. *Developmental Science, 14*, 1261–1282. https://doi.org/10.1111/j.1467-7687.2011.01075.x
Miller, G. A., & Selfridge, J. A. (1950). Verbal context and the recall of meaningful material. *American Journal of Psychology, 63*, 176–185. https://doi.org/10.2307/1418920
Monaghan, P. (2017). Canalization of language structure from environmental constraints: A computational model of word learning from multiple cues. *Topics in Cognitive Science, 9*, 21–34. https://doi.org/10.1111/tops.12239
Obeid, R., Brooks, P. J., Powers, K. L., Gillespie-Lynch, K., & Lum, J. A. G. (2016). Statistical learning in specific language impairment and autism spectrum disorder: A meta-analysis. *Frontiers in Psychology, 7*, 1245. https://doi.org/10.3389/fpsyg.2016.01245
Orbán, G., Fiser, J., Aslin, R. N., & Lengyel, M. (2008). Bayesian learning of visual chunks by human observers. *Proceedings of the National Academy of Sciences (USA), 105*, 2745–2750. https://doi.org/10.1073/pnas.0708424105
Perruchet, P., & Poulin-Charronnat, B. (2012). Beyond transitional probability computations: Extracting word-like units when only statistical information is available. *Journal of Memory and Language, 66*, 807–818. https://doi.org/10.1016/j.jml.2012.02.010
Picton, T. W. (1992). The P300 wave of the human event-related potential. *Journal of Clinical Neurophysiology, 9*, 456–479.
Reber, A. S. (1967). Implicit learning of artificial grammars. *Journal of Verbal Learning and Verbal Behavior, 6*, 855–863. https://doi.org/10.1016/S0022-5371(67)80149-X
Reber, A. S. (1989). Implicit learning and tacit knowledge. *Journal of Experimental Psychology: General, 118*, 219–235. https://doi.org/10.1037/0096-3445.118.3.219
Reeder, P. A., Newport, E. L., & Aslin, R. N. (2013). From shared contexts to syntactic categories: The role of distributional information in learning linguistic form-classes. *Cognitive Psychology, 66*, 30–54. https://doi.org/10.1016/j.cogpsych.2012.09.001
Rogers, T. T., Rakison, D. H., & McClelland, J. L. (2004). U-shaped curves in development: A PDP approach. *Journal of Cognition and Development, 1*, 137–145. https://doi.org/10.1207/s15327647jcd0501_14
Romberg, A. R., & Saffran, J. R. (2010). Statistical learning and language acquisition. *Wiley Interdisciplinary Reviews: Cognitive Science, 1*, 906–914. https://doi.org/10.1002/wcs.78
Roser, M. E., Aslin, R. N., McKenzie, R., Zahra, D., & Fiser, J. (2015). Enhanced visual statistical learning in adults with autism. *Neuropsychology, 29*, 163–172. https://doi.org/10.1037/neu0000137
Saffran, J. R. (2001). Words in a sea of sounds: The output of statistical learning. *Cognition, 81*, 149–169. https://doi.org/10.1016/S0010-0277(01)00132-9
Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science, 274*, 1926–1928. https://doi.org/10.1126/science.274.5294.1926
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. https://doi.org/10.1016/S0010-0277(98)00075-4
Saffran, J. R., & Kirkham, N. Z. (2018). Infant statistical learning. *Annual Review of Psychology, 69*, 181–203. https://doi.org/10.1146/annurev-psych-122216-011805
Saffran, J. R., & Wilson, D. P. (2003). From syllables to syntax: Multilevel statistical learning by 12-month-old infants. *Infancy, 4*, 273–284. https://doi.org/10.1207/S15327078IN0402_07

Schonberg, C., Marcus, G. F., & Johnson, S. P. (2018). The roles of item repetition and position in infants' abstract rule learning. *Infant Behavior and Development, 53*, 64–80. https://doi.org/10.1016/j.infbeh.2018.08.003

Scott-Van Zeeland, A. A., McNealy, K., Wang, A. T., Sigman, M., Bookheimer, S. Y., & Dapretto, M. (2010). No neural evidence of statistical learning during exposure to artificial languages in children with autism spectrum disorders. *Biological Psychiatry, 88*, 345–351. https://doi.org/10.1016/j.biopsych.2010.01.011

Shafto, C. L., Conway, C. M., Field, S. L., & Houston, D. M. (2011). Visual sequence learning in infancy: Domain-general and domain-specific associations with language. *Infancy, 17*, 247–271. https://doi.org/10.1111/j.1532-7078.2011.00085.x

Siegelman, N., & Frost, R. (2015). Statistical learning as an individual ability: Theoretical perspectives and empirical evidence. *Journal of Memory and Language, 81*, 105–120. https://doi.org/10.1016/j.jml.2015.02.001

Slone, L. K., & Johnson, S. P. (2015). Infants' statistical learning: 2- and 5-month-olds' segmentation of continuous visual sequences. *Journal of Experimental Child Psychology, 133*, 47–56. https://doi.org/10.1016/j.jecp.2015.01.007

Slone, L. K., & Johnson, S. P. (2018). When learning goes beyond statistics: Infants represent visual sequences in terms of chunks. *Cognition, 178*, 92–102. https://doi.org/10.1016/j.cognition.2018.05.016

Sobel, D. M., & Kirkham, N. Z. (2006). Blickets and babies: The development of causal reasoning in toddlers and infants. *Developmental Psychology, 42*, 1103–1115. https://doi.org/10.1037/0012-1649.42.6.1103

Spelke, E. S. (1990). Principles of object perception. *Cognitive Science, 14*, 29–56. https://doi.org/10.1016/0364-0213(90)90025-R

Teinonen, T., Fellman, V., Näätänen, R., Alku, P., & Huotilainen, M. (2009). Statistical language learning in neonates revealed by event-related brain potentials. *BMC Neuroscience, 10*, 21. https://doi.org/10.1186/1471-2202-10-21

Tenenbaum, J. B., Kemp, C., Griffiths, T. L., & Goodman, N. D. (2011). How to grow a mind: Statistics, structure, and abstraction. *Science, 331*, 1279–1285. https://doi.org/10.1126/science.1192788

Thiessen, E. D. (2016). What's statistical about learning? Insights from modelling statistical learning as a set of memory processes. *Philosophical Transactions of the Royal Society B, 372*, 20160048. https://doi.org/10.1098/rstb.2016.0056

Thiessen, E. D., Kronstein, A. T., & Hufnagle, D. G. (2013). The extraction and integration framework: A two-process account of statistical learning. *Psychological Bulletin, 139*, 792–814. https://doi.org/10.1037/a0030801

Thiessen, E. D., & Saffran, J. R. (2003). When cues collide: Use of stress and statistical cues to word boundaries by 7- to 9-month-old infants. *Developmental Psychology, 39*, 706–716. https://doi.org/10.1037/0012-1649.39.4.706

Toro, J. M., & Trobalón, J. B. (2005). Statistical computations over a speech stream in a rodent. *Perception & Psychophysics, 67*, 867–875. https://doi.org/10.3758/BF03193539

Tovar, A. E., Westermann, G., & Torres, A. (2018). From altered synaptic plasticity to atypical learning: A computational model of Down syndrome. *Cognition, 171*, 15–24. https://doi.org/10.1016/j.cognition.2017.10.021

Tummeltshammer, K. S., & Kirkham, N. Z. (2013). Learning to look: Probabilistic variation and noise guide infants' eye movements. *Developmental Science, 16*, 760–771. https://doi.org/10.1111/desc.12064

Turk-Browne, N. B. (2012). Statistical learning and its consequences. In M. D. Dodd & J. H. Flowers (Eds.), *The influence of attention, learning, and motivation on visual search (Nebraska Symposium on Motivation)* (pp. 117–146). New York, NY: Springer.

Turk-Browne, N. B., Scholl, B. J., Chun, M. M., & Johnson, M. K. (2009). Neural evidence of statistical learning: Efficient detection of visual regularities without awareness. *Journal of Cognitive Neuroscience, 21*, 1934–1945. https://doi.org/10.1162/jocn.2009.21131

Turk-Browne, N. B., Scholl, B. J., Johnson, M. K., & Chun, M. M. (2010). Implicit perceptual anticipation triggered by statistical learning. *Journal of Neuroscience, 30*, 11177–11187. https://doi.org/10.1523/JNEUROSCI.0858-10.2010

Wightman, C. W., Shattuck-Hufnagel, S., Ostendorf, M., & Price, P. J. (1992). Segmental durations in the vicinity of prosodic phrase boundaries. *Journal of the Acoustical Society of America, 91*, 1707–1717. https://doi.org/10.1121/1.402450

Wu, R., Gopnik, A., Richardson, D. C., & Kirkham, N. Z. (2011). Infants learn about objects from statistics and people. *Developmental Psychology, 47*, 1220–1229. https://doi.org/10.1037/a0024023

Wynn, K. (1992). Addition and subtraction by human infants. *Nature, 358*, 749–750. https://doi.org/10.1038/358749a0

Chapter 3
How Multiple Exemplars Matter for Infant Spatial Categorization

Marianella Casasola and Youjeong Park

Abstract The goal of the present chapter is to outline how infants' experience with multiple exemplars contributes to their ability to form representations of the small-scale spatial relations, such as above, below, between, inside, or on top. Although infants discriminate between changes in the spatial arrangement of objects early in development, this skill undergoes significant advances throughout infancy. In particular, the ways in which infants benefit from multiple exemplars evolve with the development of their spatial skills and their categorization skills. We outline theoretical views that inform these developmental changes and point to the possible mechanisms that may underlie how infants' experience with multiple exemplars contributes to their ability to form abstract representations of spatial relations. We also consider how manipulating the number and type of exemplars is an important tool for understanding the development of infant spatial categorization.

To quote a familiar adage, variety is the spice of life. Diverse experiences promote engagement, boost enjoyment, and maintain interest. This same mantra can be applied to learning. Learning contexts that include multiple examples facilitate generalization to new contexts more effectively than those that include only a single example (e.g., Gentner, Loewenstein, & Thompson, 2003). The goal of the present chapter is to evaluate how experience with multiple exemplars contributes to infants' ability to form representations of the spatial relations between and among objects, relations such as above, below, between, inside, or on top. From their first days, infants show that they can discriminate the spatial layout of the objects in their environment (Antell, Caron, & Myers, 1985). For example, neonates can generalize across changes in the absolute location of one object relative to the left versus right of another, provided the left-right spatial arrangement between the two objects is

M. Casasola (✉)
Cornell University, Ithaca, NY, USA
e-mail: mc272@cornell.edu

Y. Park
Seoul National University, Seoul, South Korea

J. B. Childers (ed.), *Language and Concept Acquisition from Infancy Through Childhood*, https://doi.org/10.1007/978-3-030-35594-4_3

maintained (Gava, Valenza, & Turati, 2009). By 6 months, infants form abstract categorical representations of particular spatial relations, generalizing the spatial relation from a familiarized set of objects to novel ones (Casasola, Cohen, & Chiarello, 2003; Hespos & Spelke, 2004; Quinn, Cummins, Kase, Martin, & Weissman, 1996). Nonetheless, infants' ability to form categorical representations of spatial relations continues to develop throughout infancy (Casasola, 2008, 2017; Quinn, 2003, 2005, 2007), and studies manipulating the number and type of exemplars presented to infants in the spatial categorization task have been pivotal in documenting the nature and scope of this development (Casasola & Park, 2013; Hespos & Piccin, 2009; Quinn, Adams, Kennedy, Shettler, & Wasnik, 2003; Quinn, Polly, Furer, & Dobson, 2002).

Prior studies of infant spatial categorization show that infants' experience with different exemplars plays a central role in the development of this skill in at least two ways. First, experimental tasks of infant spatial categorization often rely on presenting infants with multiple exemplars of a spatial relation. The prediction is that if infants can attend to the commonality of the spatial relation across the distinct examples of the relation, they will generalize the experience with the familiarization of the spatial relations to a novel instance of the spatial relation and show longer looking to a novel than a familiarized spatial relation. Importantly, the exemplars used across studies of infant spatial categorization can differ drastically in the type and number of exemplars they include. Whereas some studies depict the same objects during familiarization with the only changes across exemplars in the location of a figure object in relation to a referent, such as a bar (e.g., Quinn, 1994), others present exemplars in which the target relation is depicted by a diverse array of objects depicting that relation (e.g., McDonough, Choi, & Mandler, 2003). This difference in the number and type of exemplars across studies has yielded important insights into not only *which* spatial categories infants form at a given age but also *how* they form these categories; it has made it possible to trace developmental changes in this skill (Casasola & Cohen, 2002; Quinn et al., 1996, 2003). Infants' experience with multiple exemplars also has been important to infant spatial categorization because, under some circumstances, varying the number or type of familiarization exemplars can promote infants' ability to form a spatial category that they may not otherwise form (Casasola, 2005; Casasola & Park, 2013; Hespos & Piccin, 2009; Park & Casasola, 2015; Park, Casasola, & Kim, 2012; Quinn et al., 2002). Specifically, in some cases, increasing the number of exemplars facilitates spatial categorization (e.g., Casasola & Park, 2013), but in other cases, limiting the number of exemplars does so (Casasola, 2005). Infants' ability to benefit from, or be hindered by, the number and type of familiarization exemplars provides a window into the developmental processes that contribute to advances in infants' categorization of spatial relations.

The present chapter discusses how infants' experience with multiple exemplars contributes to their ability to form abstract representations of spatial relations. We consider multiple exemplars, both as a tool for testing infant spatial categorization and as the spice that can promote generalization, when infants have the skills to benefit from increased variability in the exemplars. In the first section of the chapter, we argue that infants' ability to form a categorical representation of a spatial relation offers an ideal venue for understanding how infants encode and generalize relational information. We also outline how infant categorization of spatial relations intersects with other cognitive skills, such as their categorization of objects, analogical reasoning, and their acquisition of spatial language, and propose that understanding advances in infant spatial categorization may serve to inform our understanding of these other skills. We next discuss the use of multiple examples in tasks of infant spatial categorization and note how differences in the number of exemplars as well as the perceptual similarity of the exemplars used across studies have served to highlight the conceptual hurdles that infants must overcome to form abstract categorical representations of spatial relations. We outline possible mechanisms of facilitation and link these mechanisms to theoretical accounts of infant spatial categorization, specifically, and relational learning, more broadly.

As we conclude, we will address possible directions for future study, including identifying other methodological approaches that may be useful in understanding how infants' spatial categorization and other types of relational learning develop over the first year. We also will consider the degree to which our understanding of how multiple exemplars shape spatial categorization can be extended to bolster our understanding of other types of cognitive skills, including those that are both spatial and those that are nonspatial.

Why Spatial Relations?

Studying how infants learn to form categorical representations of spatial relations informs our understanding of their ability to look beyond the objects in their environment. Infant spatial categorization has often been situated within the larger domain of spatial cognition (Newcombe & Huttenlocher, 2000). Infants' discrimination and categorization of the spatial configuration between or among objects early in life contribute to a number of later-emerging spatial skills, such as map reading, navigation, and reorientation. It also falls under the umbrella of how infants parse and categorize dynamic events and overlaps with studies that examine other aspects of motion events, such as infants' discrimination and categorization of the manner and path of motion (e.g., a girl skipping across the street) or the figure and ground objects in these events (Göksun, Hirsh-Pasek, & Golinkoff, 2010; Konishi, Pruden, Golinkoff, & Hirsh-Pasek, 2016; Pruden, Roseberry, Göksun, Hirsh-Pasek, & Golinkoff, 2013; Pulverman, Song, Hirsh-Pasek, Pruden, & Golinkoff, 2013; Song, Pruden, Golinkoff, & Hirsh-Pasek, 2016). Given the focus of infants' generalization of spatial relations, this skill easily connects to the rich literature on infants' categorization of objects. Developmental changes in infants' categorization of spatial relation do parallel their ability to do so with objects, suggesting overlap in the processes that infants recruit in forming each type of category (Quinn, 2003; Rakison & Oakes, 2003) and suggesting shared mechanisms across spatial and nonspatial domains.

Infants' ability to form an abstract representation of spatial relations has been linked to other cognitive skills. In particular, infants' manipulation of objects has been related to their representations of spatial relations in studies that examine developmental changes in infants' skill in fitting objects through openings (Örnkloo & von Hofsten, 2007; Oudgenoeg-Paz, Boom, Volman, & Leseman, 2016; Oudgenoeg-Paz, Leseman, & Volman, 2015). Infants' mastery in successfully inserting an object into the appropriate opening has been attributed to gains in their ability to accurately match an object's shape to its corresponding negative space in a shape sorter and appreciate the relation between shape and opening (Örnkloo & von Hofsten, 2007; Shutts et al., 2009). Thus, infants' ability to attend to spatial relations has been relevant to work that intersects with motor development and infants' play with objects. In addition, infant spatial categorization has been discussed in relation to young infants' ability to discriminate between physically possible and impossible dynamic events (e.g., Baillargeon, 2004). Theoretical accounts of infants' physical reasoning have outlined a progression dictated by the type of spatial events that depict the violation in the physical interaction between two objects. Infants display distinct developmental timelines for responding to a particular physical violation across containment, support, and occlusion events (Baillargeon & Wang, 2002; Hespos & Baillargeon, 2001; Wang & Baillargeon, 2008), intersecting with work on infant categorization of spatial relation that also notes distinct developmental timelines for when they form a spatial category of particular spatial relations (to be discussed in more detail below).

Furthermore, which types of spatial relations infants can organize into categories has been the relevant to studies of early verb learning as well as infants' acquisition of locative terms, such as "in," and "on." A number of studies have documented infants' sensitivity to manner, and path of motion predicts their later acquisition of verbs, supporting arguments that infants' representations of these events contribute to a conceptual foundation for relational language (Göksun et al., 2010; Pulverman et al., 2013; Song et al., 2016). Similarly, a number of studies have linked infants' stacking or nesting of objects to their exposure to spatial language (Casasola, Bhagwat, Doan & Love, 2017) or their later acquisition of spatial language, particularly locative terms (e.g., Marcinowski & Campbell, 2017), linking infants' motor experience, particularly their manipulation of objects into specific spatial configurations, and their spatial vocabulary and spatial skills (see also Oudgenoeg-Paz et al., 2015).

Finally, infants' ability to discriminate and categorize spatial relations has been argued to be related to their analogical reasoning (Ferry, Hespos, & Gentner, 2015; Park & Casasola, 2017). In particular, infants' ability to generalize across different instantiations of a spatial relation can be argued to be a type of relational learning. To form a *perceptual* category of a spatial relation, for example, infants must attend to the consistency of the spatial arrangement between or among specific objects, despite changes in absolute location. To form an *abstract* representation, infants must also do so across changes in the objects depicting the spatial relation, documenting that they recognize the spatial relation independent of specific objects.

Infant spatial categorization offers a number of advantages for studying the origins and development of relational learning. We will first outline some of these advantages and then later provide more details of the studies that support these claims. First, this categorization is evident, in some form, even in neonates, and may possibly be one of the first types of relational learning documented in infants. Its early emergence provides a unique opportunity to explore how very young infants attend to these types of relations in their environment. Second, this skill undergoes significant and rapid development in the first year. In particular, by about 6 months of age, infants become significantly more adept at recognizing the equivalence of a spatial relation across different examples of the relation, although this ability continues to develop into the second year. This developmental progression in how easily infants can generalize a spatial relation across changes in objects makes this skill especially well-suited for appreciating how multiple exemplars can play a role in relational learning, particularly in infancy.

In sum, infant spatial categorization intersects with both spatial and nonspatial skills as well as object cognition, event perception, and relational learning. As such, documenting the processes infants recruit to form abstract representations of spatial relations and investigating the role of multiple exemplars in shaping these processes may yield insights into a broad array of cognitive skills.

Do Infants Require Multiple Exemplars for Forming Spatial Categories?

Infants' daily experiences provide them with ample exposure to varied examples of spatial relations. However, we now know from the empirical work that has been done that the degree to which infants benefit from these experiences depends on an array of factors, including the spatial relation in question, the characteristics and familiarity of the objects depicting the spatial relation, the structure of the categorization task, and, of course, the perceptual and cognitive abilities of the infant. Over their first year, infants become more adept at forming spatial categories across increasingly diverse examples of the specific spatial relation. Although a 1-month-old infant may not appreciate that a cup in the sink depicts the same containment relation as carrot sticks in a bowl, infants of 6 months of age can do so, generalizing a containment relation across distinct pairs of objects (e.g., Casasola et al., 2003). However, the 1-month-old can form a spatial category of left-right spatial relations if it is depicted by a specific pair of perceptually simple objects.

Understanding the progression by which infants learn to form spatial categories sets the stage for appreciating how experience with multiple exemplars may shape this development. First, when it comes to forming spatial categories, not all spatial relations are equivalent. Infants form categories of particular spatial relations earlier in development than others. For example, as we have mentioned, neonates form a perceptual category of left-right relations (Antell & Caron, 1985; Gava et al., 2009).

There is also evidence that young infants can mentally subdivide small-scale spaces even in the absence of a midline, discriminating between objects situated to the left as distinct from those situated on the right. This result suggests a primacy to how we organize space along a vertical axis (Quinn, 2012). By 4 months, infants provide evidence for the formation of categories of other spatial relations, such as above and below (Quinn, 1994), and by 6 months, this skill has expanded to include the spatial relations of containment, between, and those that depict a tight-fit relation between two objects (Casasola et al., 2003; Hespos & Spelke, 2004; Quinn, Norris, Pasko, Schmader, & Mash, 1999). Finally, by 8 months, infants provide the first evidence that they can form a spatial category of support relations (Park et al., 2012; Park & Casasola, 2013). Thus, infants form categories of some types of spatial relations earlier in development than others. Which spatial relations infants first categorize has been attributed to core concepts of events (e.g., Hespos & Spelke, 2004), biases in perceptual and cognitive systems (Quinn, 2012), their own experience in manipulating objects (Casasola et al., 2017), or those that are linguistically encoded by the infants' language (Choi & Bowerman, 1991).

An important caveat, however, is that infants' ability to form a spatial category in one study should not be taken as evidence that they will successfully do so in all other contexts. Consider, for example, results reported by Casasola et al. (2003) with infants of 6 months and those reported by Rigney and Wang (2015) with infants of 8 months. Whereas Casasola et al. (2003) found that infants of 6 months formed a category of containment events as distinct from support, Rigney and Wang (2015) found that infants failed to do so when the test events contrasted containment with occlusion (but did form the spatial category when the novel relation was support, replicating Casasola et al.). Similarly, Quinn et al. (1999) found that infants of 6–7 months could form a category of the spatial relation of between when familiarized with exemplars of this spatial relation in which the two lines were always in the same orientation (either always vertical or always horizontal). However, when infants were familiarized to two lines in a distinct orientation (e.g., vertical), they no longer had formed the spatial category of between when tested with the referent frame in a novel orientation (e.g., horizontal) (Quinn, Doran, & Papafragou, 2011). At slightly older ages, 11 months for Rigney and Wang and 9–10 months for Quinn, Doran, and Papafragou, infants were able to form each spatial category regardless of which spatial relation was depicted as the novel spatial relation and whether the orientation of the lines varied from familiarization to test. In sum, at younger ages, infant success in one spatial categorization task may not extend to a task that depicted more challenging or even simply distinct conditions; at older ages, they become more resilient in their spatial categorization across a wider array of conditions.

Developmental changes in infant spatial categorization are evident not only in which spatial relations they can organize into a category but also in the breadth of their generalizations when tested on their categorization of a spatial relation. As we have noted, infants progress from forming spatial categories across exemplars that share a high level of similarity to forming ones in which there is minimal similarity across the exemplars. That is, they advance from forming spatial categories that are more perceptually based to ones that are increasingly abstract. To illustrate this point, compare the scope of a spatial category formed by young versus older infants. In the case of very young infants, the only difference across exemplars in several studies was in the specific location of one or both of the objects (Antell et al., 1985; Gava et al., 2009; Quinn, 1994). Because the objects remained constant across all phases of the experiment, we could predict that any generalization would be narrow in scope and that the category formed would be a perceptual one (e.g., Gava et al., 2009). At the other end of the continuum are studies in which the spatial events include more perceptually complex objects, such as those that infants may encounter in their everyday interactions, objects such as a cup, bowl, or toy rather than the monochromatic geometric shapes or symbols presented to younger infants (see Fig. 3.1). In these studies, the objects also varied across trials, creating a more difficult categorization task (Casasola & Ahn, 2017; Choi, 2006; McDonough, Choi, & Mandler, 2003). To form the spatial category with these more variable exemplars, infants had to generalize the spatial relation across changes in the objects, a skill that they do not demonstrate until about 5–6 months of age and even so, only with particular spatial relations, such as containment, above versus below, and, in some cases, tight-fit (Casasola et al., 2003; Hespos & Spelke, 2004; Quinn et al., 1996). When infants recognize a spatial relation independent of specific objects, they are argued to have formed the abstract representation of that spatial relation.

In the next section, we review the methods for testing infants' spatial categorization, with particular attention to the use of multiple exemplars in these tasks to test the spatial categories that infants form. In particular, we note how the stimuli and structure of a particular experiment may shape whether infants form a spatial category.

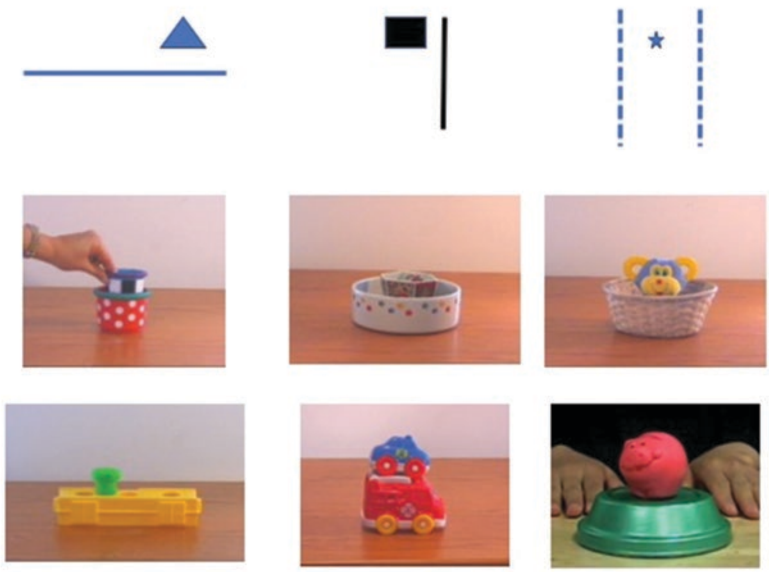


Fig. 3.1 Adaptation of stimuli used in distinct studies of infant spatial categorization

Procedures for Testing Infants' Categorization of Spatial Relations

As demonstrated by the contrasting results reported by Rigney and Wang (2015) and Casasola et al. (2003), the stimuli and structure of the spatial categorization task can have a significant impact on whether infants provide evidence of forming a particular spatial category. Most studies probing whether infants can organize the spatial configuration between objects into a category have used infant looking time as the dependent variable. Studies also typically include several examples of the target relation as part of the spatial categorization task. Once infants have met a criterion of decreased looking during a habituation phase or viewed a fixed number of familiarization trials, they are presented with test events that depict the familiarized spatial relation as well as a novel spatial relation, one not seen during familiarization. Typically, infants tested in a habituation task are expected to demonstrate longer looking to the novel than familiarized spatial relation, but in studies that use a familiarization procedure, they can show a familiarity preference for the familiarized spatial relation. Infants' preference for one type of event over the other during the test phase is taken as evidence that they have formed a categorical representation of the spatial relation. This approach has been used successfully with a wide range of infant ages, from neonates to young children of 36 months, and with a diverse array of spatial categories, including above, below, left, right, between, containment, support, and tight-fit (Antell et al., 1985; Behl-Chadha & Eimas, 1995; Choi, 2006; Gava et al., 2009; Hespos & Spelke, 2004; Quinn, 1994; Quinn et al., 1999).

Despite sharing these commonalities in general structure, studies of infant spatial categorization have differed, often drastically, in the type of stimuli, including the amount of variability in the stimuli, and structure of the categorization task, differences which can shape whether infants form the spatial category in the task. One notable difference across studies is in the representational depiction of the spatial relation. In studies with young infants, those in their first 6 months, the spatial relation is often depicted as a static image (e.g., Antell & Caron, 1985; Gava et al., 2009; Quinn et al., 2011). In studies with somewhat older infants (such as those of about 6 months and older), the spatial relation is instead usually depicted in a dynamic event in which a figure object, such as a cup, is seen being placed into the target relation to a second object (e.g., Casasola & Cohen, 2002; Hespos & Spelke, 2004; McDonough et al., 2003). Regardless of whether infants have viewed the static or dynamic versions of a spatial relation, if they look longer at the novel instance at test, they have provided evidence of forming a spatial category, suggesting that perhaps the level of representation of the spatial relation may not significantly impact their performance on the categorization task. However, to date, very young infants have not been tested with the dynamic versions of spatial relations, nor have any studies directly compared infants on their spatial categorization across static versus dynamic instances of a spatial relation. Similarly, most studies present spatial relations as two-dimensional representations of the images or event, but in a few cases, infants have been tested with live presentations of a dynamic spatial events (Hespos & Piccin, 2009; Hespos & Spelke, 2004). It is not clear if one type of event may be easier for infants to encode and to categorize or if instead infants are attuned to the spatial relations in their environment whether they are in motion or static and whether they are viewed on a screen or in person.

One distinct advantage of the studies that present static images of the spatial relation is that these studies depict the familiarized and novel spatial relations side by side at test, possibly facilitating infants' ability to compare the two images and attend to the novel relation. In studies that use dynamic events of spatial relations, the test events are instead presented sequentially (but see Choi, 2006, and McDonough et al., 2003, for exceptions). This difference in the presentation of the test events has been shown to matter for infants' categorization of objects. When given the opportunity to view exemplars side by side during familiarization, even if the overlap in items is brief, infants are more likely to form the objects' category than when familiarized to objects sequentially (Kovack-Lesh & Oakes, 2007). This difference in results has been attributed to the greater memory demands created when test items are presented sequentially. It is not surprising then that infants can provide evidence of category formation with paired presentations than sequential ones. Of course, this direct comparison has not been conducted with infants' categorization of spatial relations, but it stands to reason that the same pattern of results may emerge with infant spatial categorization, with the inclusion of paired test events possibly facilitating young infants' spatial categorization.

Studies of infant spatial categorization have also differed in the type of objects used to depict the spatial relation of interest. There seems to be a divide in the perceptual complexity and amount of variability in the stimuli presented to younger versus older infants. Studies with the youngest infants have used monochromatic symbols or geometric forms, in relation to a fixed referent object, such as a bar, to depict the spatial relation, such as those depicted in the first row of Fig. 3.1 (e.g., Gava et al., 2009; Quinn, 1994; Quinn et al., 1996). For example, Quinn (1994) selected a dot and line to test infants on their categorization of above versus below, Gava et al. (2009) used a flashing gray square against a dark background in relation to a vertical bar, and Antell and Caron (1985) depicted a plus sign above a square when testing infants in one of the first studies to explore infants' ability to encode the spatial arrangement between two objects. In contrast, in studies with older infants, the spatial relations have been depicted with realistic objects in dynamic events, such as tubes, cups, and toys, such as those depicted in the second and third rows of Fig. 3.1 (e.g., Casasola & Cohen, 2002; Hespos & Spelke, 2004; McDonough et al., 2003). There has not been much discussion about the choice of objects used to depict the spatial relation in studies of infant spatial categorization. Presumably, the stimuli are created or chosen because they are considered most appropriate for the goals of a study and for the age of infant tested. For young infants, particularly those with limited visual acuity, perceptually simple, monochromatic stimuli with high contrast seem ideal for ensuring that infants can parse the objects from their background. Nonetheless, it is not clear whether very young infants would demonstrate the same ability to form a spatial category if provided with more perceptually rich objects.

The stimuli in spatial categorization task can differ in the amount of variability they introduce throughout the experiment, with implications for the scope of the spatial category that infants form. In some cases, there is minimal variability across the exemplars with infants viewing familiarization and test objects that are identical to each other. For example, Gava et al. (2009) presented neonates with a square that flashed in one of three locations on one side of a vertical bar. During the test phase, infants viewed pairs of stimuli in which the blinking square appeared in a novel location, not seen during the familiarization phase. In one test stimulus, infants viewed the square on the same side of the vertical line as during familiarization, whereas in the other test stimulus, it appeared on the alternate side of the bar. Infants demonstrated a significant preference for the test stimulus that depicted a change in the left-right placement of the square relative to the vertical line. That is, very young infants demonstrate the ability to generalize across changes in absolute location between a single, simple shape (e.g., a square) and a single referent (e.g., a bar). However, because the only variation between the familiarized and test events was in the specific location of the object in relation to the referent line, this type of spatial category has been described as a perceptual one because the scope of generalization is a narrow one.

The spatial categorization tasks with older infants have presented greater amount of variability across the exemplars than the tasks presented to younger infants. For example, in the study by Rigney and Wang (2015), infants viewed three distinct pairs of objects in a containment relation and, during the test phase, viewed novel objects in the familiarized relation as well as novel objects in an unfamiliar relation. McDonough et al. (2003) presented infants of 9, 11, and 14 months with three pairs of dynamic events to depict a spatial relation. Each pair depicted distinct objects such that infants were given exposure to six exemplars of a spatial relation and each exemplar depicted very different objects. A similar design has been used to explore when infants can form a more abstract representation of above versus below (Quinn et al., 1996), between (Quinn et al., 2003), support (Park et al., 2012; Park & Casasola, 2015), and tight-fit (Casasola & Ahn, 2017). Because infants must generalize the spatial relation from one set of objects seen during familiarization to new objects seen during the test phase, they are argued to have formed an abstract categorical representation of a spatial relation (Casasola, 2008; Quinn et al., 1996) and to have formed a spatial category that is broader in the scope of categorization relative to the task with neonates that present the same objects throughout the task.

Finally, studies also have differed in the number of distinct exemplars presented during the familiarization or habituation phase. Many studies familiarize infants with four exemplars of a spatial relation (e.g., Casasola & Cohen, 2002; Park et al., 2012; Park & Casasola, 2013; Quinn, 1994; Quinn et al., 1996, 2003), but several use just two or three exemplars (e.g., Casasola, 2005; Park & Casasola, 2015; Rigney & Wang, 2015), and a few use as many as six exemplars (Casasola, 2005; Park & Casasola, 2015; Choi, 2006; McDonough et al., 2003). Presenting infants with multiple exemplars of a spatial relation in the spatial categorization task is a useful tool for probing the nature of infants' spatial categories and, in particular, their ability to generalize a spatial relation from the array of distinct, familiarized examples of the relation to novel ones. In particular, manipulating the similarity among the exemplars makes it possible to test the breadth of events that infants can organize into a spatial category.

Interestingly, not all studies of infant spatial categorization have relied on presenting infants with a set of distinct exemplars during the familiarization phase to examine their ability to generalize a spatial relation. In some studies, infants are familiarized to a single event. For example, in their first study by Gava et al. (2009), neonates viewed a single exemplar of the spatial relation, a flashing square in a single location on one side of the bar. Even after viewing just this single exemplar, neonates generalized the left-right arrangement to a novel location on the same side of the vertical bar as presented during familiarization, looking longer at the display with the square in the new location on the alternate side of the vertical bar. Thus, even within days after birth, infants demonstrate the ability to generalize across changes in absolute locations that maintain the left-right spatial arrangement between two objects.

In another study, Hespos and Spelke (2004) presented infants with a single event and then tested infants with changes in the type of spatial relation. The goal of this study was to document infants' ability to discriminate changes in the degree of fit between two objects, testing infants' categorical perception of tight-fit, a relation that is the basis of a semantic category in Korean. Given this goal, the inclusion of a single familiarization event was well suited to exploring young infants' sensitivity to when there is an exact, interlocking fit between two objects. Similarly, Hespos and Piccin (2009) demonstrated that infants of 5 months could generalize the type of fit from one type of event to a distinct type of spatial event, even when familiarized to a single event. In this study, infants were familiarized to a single exemplar of an event, a covering event for some infants and an occlusion event for other infants. Infants then viewed test events that depicted a containment relation. Critically, the containers were much wider than the object, depicting a loose-fit, or were only slightly wider than the lowered object, depicting a tight-fit once inserted. Even with habituation to a single event, infants of 5 months generalized the type of fit when habituated to the covering events. They looked significantly longer at the test events that depicted the change in fit. Thus, although presenting multiple exemplars is a common practice when testing infant spatial categorization, infants can sometimes generalize their spatial knowledge without it. More specifically, infants can still generalize a spatial relation following experience with just a single exemplar, pointing to the robustness of this skill across variation in the structure of the spatial categorization task.

In sum, the procedures for studying infant spatial categorization all share the same dependent variable, infant looking to a familiarized versus novel spatial relation. Yet these paradigms also can differ in how they represent the spatial relation (e.g., static vs. dynamic, two-dimensional vs. three-dimensional), the type of objects presented (symbols, simple tubes or container, or more perceptually detailed household items or toys), the variability in the objects across the phases of the categorization task (no variability to low variability to relatively high variability), and the number of exemplars (from a single exemplar to as many as six exemplars), variations which can shape infants' ability to form the spatial category. Impressively, infants often have provided converging evidence of forming a spatial category across versions of the task which can be quite distinct from each other, including in the number and similarity of the exemplars. In cases where results have not converged, the discrepancy in results offers valuable insights into the processes that contribute to the development of this spatial skill. Indeed, one can consider each study of infant spatial categorization as its own exemplar of this ability, and by comparing across exemplars, it becomes possible to identify which aspects of infant spatial categorization are robust and which aspects may be emerging and, thus, sensitive to variation in how this skill is tested. In addition, discrepancies across results can signal which aspects of infant spatial categorization are currently emerging as well as point to the underlying processes guiding the development of this skill.

Does Infant Spatial Categorization Benefit from Multiple Examples?

Thus far in our discussion, the impact of exemplars on infant spatial categorization task has been inferred by comparing differences in the learning phase that impact results across individual studies. In these studies, infants form the spatial category regardless of whether they view a single versus multiple exemplars of the relation during familiarization, suggesting that the number of exemplars may not have a significant impact on infants' ability to form a spatial category. However, there is a sizable literature reporting a significant effect of varying the number of exemplars during the learning phase, suggesting that, although exemplar number may not seem to have a strong effect across individual studies of infant spatial categorization, it may nonetheless promote infants' generalization when manipulated within a study.

To explore this possibility, we conducted several studies to test how manipulating the exemplars provided to infants during habituation shapes their categorization of spatial relations. In one set of studies, we manipulated the number of exemplars provided during habituation and tested infants on their categorization of containment and support relations (Casasola, 2005; Park & Casasola, 2015). In previous studies, we had found that infants of 10 and 18 months demonstrated greater ease in forming a spatial category of containment than support (Casasola & Cohen, 2002). We reasoned that if manipulating the number of exemplars has a significant effect on how infants form spatial categories, then infants might benefit from this manipulation and form a spatial category of support. However, if manipulating the number of exemplars creates a more difficult spatial categorization task, then we might instead see infants struggle to form a spatial category that they have formed in previous studies, that is, containment. Thus, the inclusion of two spatial categories made it possible to explore the directionality of the effect of manipulating exemplar number during the habituation phase. We also included two age groups of infants, those of 10 and 14 months, to document whether the effect of this manipulation might vary with infants' cognitive abilities.

Infants of each age were randomly assigned to one of two exemplar conditions, viewing either two or six exemplars of a specific spatial relation, and to one of the two spatial relations, either containment or support. During habituation, infants in the two-exemplar condition viewed two pairs of objects in a spatial relation (containment for one group of infants, support for another group of infants). In comparison, infants in the six-exemplar condition viewed six pairs of objects in the spatial relation. Each pair of objects depicted a unique figure, a clay figurine generally resembling an animal or person, placed in a containment or support relation to a referent object, such as a box, bowl, or cup, also unique within each pair. Although the pairs were distinct from each other, the figures did share a similar shape, vaguely resembling the overall structure of a snowman. If greater familiarity with a limited number of exemplars facilitates infants' spatial categorization, then infants provided with only two exemplars of the relation should be at an advantage to those provided with six exemplars; these infants would have more opportunities to view the two object pairs in their spatial relation throughout habituation. In contrast, if infants benefit from experience with more examples of a spatial relation, even though they would have less exposure to a specific exemplar, then infants in the six-exemplar condition should demonstrate more robust generalization of the spatial relation. Of particular interest was how infants in each condition would generalize their habituation of the spatial relation to a novel pair of objects in that relation (Fig. 3.2).

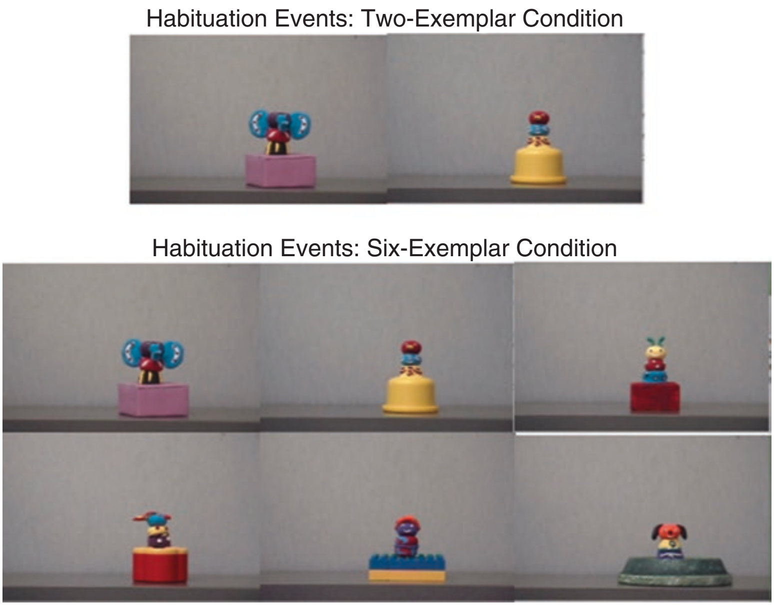


Fig. 3.2 Examples of the habituation support events used across the exemplar conditions of Park & Casasola (2015)

Following familiarization to either two or six exemplars of a containment or support relation, infants viewed four test trials, in a randomized order. One test event was an event seen during habituation. This familiarized event served as a baseline to infant looking during the test phase because both the objects and the spatial relation had been presented during habituation and were expected to be familiar to infants by the test phase. The remaining three test events presented a change in either the objects, the spatial relation, or both the objects and spatial relation. More specifically, across the four test events, two of the test events depicted objects seen during habituation, with one pair depicting the familiarized relation and another pair depicted in a novel relation. For these test events with familiarized objects, longer looking to the novel than familiarized relation is taken as evidence that infants discriminate the spatial relation. Two other test events presented novel objects, seen for the first time during the test phase. One of these novel object pairs depicted the familiarized spatial relation and the other a novel spatial relation. For these two test events with novel objects, a significant increase in looking time to the novel than familiarized spatial relation is taken as evidence that infants have generalized their learning of the familiarized spatial relation to a novel exemplar. That is, infants would provide evidence of having formed an abstract representation of the spatial relation (Fig. 3.3).

For infants of 10 months, there was a significant effect of exemplar number on their spatial categorization. Whether infants were habituated to a containment or support relation, they formed the spatial category when habituated to six exemplars of the relation. Specifically, these infants demonstrated significantly longer looking times to spatial events with an unfamiliar than familiarized spatial relation and did so whether the objects in the test events were familiar or novel. In contrast,

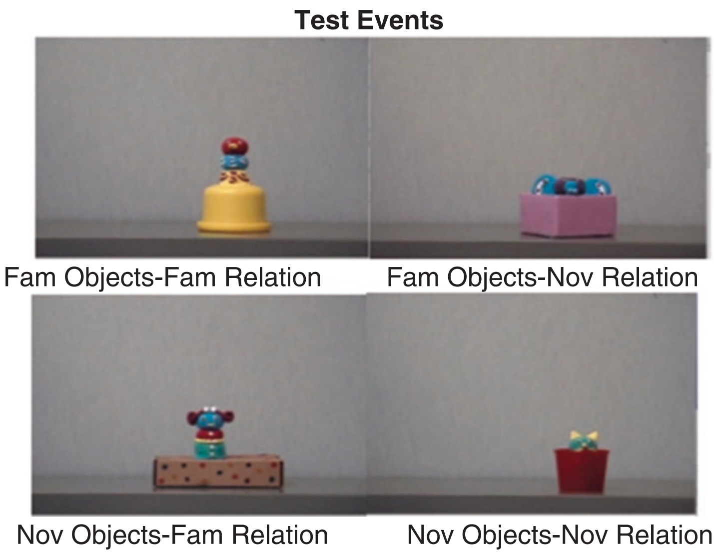


Fig. 3.3 Sample test events used in Casasola and Park (2013)

when familiarized with only two exemplars of the spatial relation, infants of 10 months failed to provide any evidence that they had formed either spatial category. They did not show a significant increase in looking time to the novel relative to the familiarized spatial relation, whether the objects were familiar or novel. That is, limiting infants' experience to two examples made it difficult for them to form a spatial category.

The findings with infants of 10 months are consistent with results across many other domains. For example, Needham, Dueker, and Lockhead (2005) found that if provided with only a single exemplar of a cylinder next to a box, infants did not segregate the display into two objects. However, if given three distinct exemplars of a cylinder and box, infants were able to parse the display into two objects. Quinn and Bhatt (2005) noted that 3- and 4-month-old infants organized elements when provided with three exemplars but failed to do so when given only a single example of the elements. Similarly, Bomba and Siqueland (1983) noted that infants required experience with 12 exemplars of irregular geometric forms to form the prototype of the form, and Gómez (2002) also found that significantly increasing the number of exemplars made it possible for both adults and infants to attend to nonadjacent dependencies in an artificial language. Infants could learn to attend to a previously ignored object feature if given experience with multiple exemplars in a study of infant physical reasoning (Wang & Baillargeon, 2008). In the domain of verb learning, Childers (2011) found that toddlers of 2.5 years were more effective in generalizing a novel verb if viewing multiple unique exemplars than when given multiple exposure to a single event. Finally, Vukatana, Graham, Curtin, and Zepeda (2015) noted that infants of 11 months could generalize a sound pairing if familiarized to three exemplars, but not when familiarized with a single exemplar. As these examples illustrate, increasing variability by providing a greater number of unique exemplars can aid learning, facilitating generalization. As the number of unique exemplars increase, infants may be better able to attend to the relational commonality of the spatial relation and form the spatial category. That is, increasing the number of exemplars may guide infants' attention to what is relevant for the category (the spatial relation) and to disregard what is not (the objects in the spatial relation).

Despite the many findings outlining the benefits of multiple exemplars on learning, there are instances in which providing additional exemplars has not promoted infant spatial categorization. In contrast to the infants of 10 months, the 14-month-old infants tested by Casasola and Park (2013) formed the spatial categories of containment and support regardless of whether they viewed two or six exemplars of the relation. That is, the number of exemplars did not have a significant effect on their spatial categorization. One interpretation of this result is that by 14 months, infants have developed the skills to form a spatial category under more diverse conditions than infants of 10 months. However, in a previous study with infants of 14 months, infants best formed a spatial category of support when habituated to two exemplars and failed to do so when habituated to six exemplars of the relation (Casasola, 2005). This pattern of results not only contrasts with what was found for infants of 10 months but also is inconsistent with Casasola and Park (2013), who found no effect of the number of habituation exemplars on 14-month-old infants' spatial categorization.

A closer inspection of the results for infants of 14 months in each study does show some degree of consistency in how infants of this age form spatial categories when provided with two versus six exemplars of a spatial relation. In the Casasola (2005) study, infants of 14 months only formed the spatial category of support when provided with two exemplars (and failed to do so when provided with six exemplars of support). This pattern of results suggested a benefit of few exemplars of a support relation for infants of 14 months. Similarly, Casasola and Park (2013) found that more infants of 14 months formed the spatial category of containment or support when habituated to two exemplars than when habituated to six exemplars. That is, there was some indication in each study that fewer exemplars may be more beneficial in forming a spatial category than more exemplars for infants of 14 months.

Why does the effect of exemplar number differ across the two age groups? One possibility may have to do with the degree to which infants encode the objects in the spatial events, which in turn shapes their ability to look beyond the objects to form the spatial category. Infants may benefit from additional exemplars of a spatial relation when the increased variability in the irrelevant features can easily shift attention to the relevant feature, in this case from the objects to the spatial relation. For example, Casasola and Park (2013) reported a significant effect of age on infants' discrimination of the objects in the spatial events. Although infants at each age discriminated changes in the objects during the test phase, infants of 14 months demonstrated a significantly greater increase in looking to the novel than familiarized objects than did infants of 10 months. Perhaps because infants of 10 months were less attentive to the change in objects than the infants of 14 months, viewing more exemplars of the spatial relation facilitated their spatial categorization. In contrast, because infants of 14 months were more attentive to objects, increasing the number of objects (as in the six exemplars) may have created a more challenging categorization task for them. Maguire, Hirsh-Pasek, Golinkoff, and Brandone (2008) outlined a similar point when examining young children's ability to generalize a novel label for an action. The children of 2.5 and 3 years in their study demonstrated more robust generalization of the novel verb when familiarized to a single actor depicting the target action than when four actors depicted the action. These authors suggest that the actors were more salient than the action, and for this reason, fewer exemplars provided the conceptual scaffolding to focus on the action rather than on the agent. When additional exemplars were included during the learning phase, the increase in the number of actors may have overwhelmed the young children with the irrelevant information. Childers, Paik, Flores, Lai, and Dolan (2016) outlined a similar pattern of results in which the young children from three distinct language groups displayed greater difficulty learning a verb when the complex actions were depicted by multiple agents.

The results from our studies of exemplar number document a clear benefit of increasing the number of exemplars for infants of 10 months. For infants of 14 months, however, infants either did not show any effect of exemplar number (Casasola & Park, 2013) or, instead, demonstrated more robust generalization when habituated to two rather than six exemplars (Casasola, 2005). Additional research is needed to better understand when multiple exemplars can promote generalization and when, instead, it may create a more challenging spatial categorization task. As these results suggest, the effect of multiple exemplars varies with the cognitive skills of infants. Ironically, advances in one domain, such as greater attention to objects, may shift how they are forming categories of spatial relations and may complicate whether infants do indeed benefit from an increase in the number of exemplars. We suspect that infants may most benefit from additional exemplars when they can easily attend to the relational commonality as the number of object pairs in the target relation increases, although this assertion remains to be tested. In the next section, we outline some of the mechanisms that guide infant spatial categorization and provide further discussion as to why variation in the features and number of exemplars has an impact on how infants form spatial categories.

What Mechanisms Are Central to Infant Spatial Categorization?

How do infants generalize a spatial relation from one or multiple instances to another? How do infants form an abstract categorical representation of a spatial relation in which objects' specific features are ignored? Like many other researchers, we postulate that forming an abstract categorical representation of a spatial relation involves abstraction, the process of extracting the essential commonalities among category members. We assume that it is the human mind's ability to align two instances physically or mentally side by side (i.e., comparison) that highlights the essential commonalities. If comparison causes abstraction, exactly what is going on during comparison? In comparing spatial events, does the infant directly go to comparing the relational structures of the instances or shift from a component-level comparison to a structure-level comparison? We consider the process of structural alignment as a likely central mechanism underlying infants' comparing, abstracting, and categorizing spatial relations.

Structural alignment has been originally proposed as the underlying mechanism of analogies (Gentner, 1983), which enables the learner to note the deeper relational structure shared between instances in different domains (Falkenhainer, Forbus, & Gentner, 1989; Gentner & Markman, 1997). Two critical pieces of structural alignment are to establish object correspondences between two instances and to compare the correspondences across common relational structure. Through this process, relations between objects are mapped from a base domain to a target domain. Furthermore, the structural alignment theory posits that the learner pursues the maximum number of commonalities. That is, the learner seeks for the largest and deepest commonalities between two instances; consequently, the particular relations mapped are determined by systematicity (i.e., the attention to higher-order relations).

Although the theory as initially proposed was more focused on analogies between two instances from different domains, it also presented the possibility that through structural alignment a person can form a *general rule* when the learner goes beyond creating a temporary correspondence and creates a new relational structure whose objects are lacking in specific attributes so that it can be applied across widely different domains (Forbus & Gentner, 1983; Gentner, 1983; Gick & Holyoak, 1980, 1983).

Structural alignment includes the process of progressive alignment (Kotovsky & Gentner, 1996). Progressive alignment refers to the idea that comparisons made between highly similar elements bolster learners to make subsequent alignments between instances having low surface-level similarity. That is, according to Kotovsky and Gentner (1996), progressive alignment acts as a mechanism of representational change, by allowing children to make similarity comparisons over concrete, perceptual similarities (e.g., monotonic increase in size across differently shaped stimuli). Then, these similarity comparisons facilitate children's ability to notice higher-order relational commonalities across stimuli possessing fewer surface-level features in common (e.g., increase in size as compared to saturation of color across differently shaped stimuli). Thus, alignment allows children to recognize "richer and deeper" abstract relational similarities across mental representations that may not have been immediately apparent before similarity comparisons were made (Kotovsky & Gentner, 1996), and progressive alignment helps children learn how to make more difficult alignments (across more varied stimuli).

According to structural alignment theory, the characteristics of components in a relation (i.e., objects) influence how easily the alignment process occurs. That is, when the objects have a high degree of literal similarity across instances, that readily invites the alignment process, a process that highlights the shared relational structure between two instances (Gentner, 1983; Gentner & Markman, 1997; Markman & Gentner, 1993). In particular, young children who have rich knowledge of objects and relatively sparse knowledge of relations (Gentner, 2005) may rely heavily on the presence of highly similar concrete object matches to carry out structural alignment and reason about relational structure. Further, the object correspondences made by the learner generate a candidate set of inferences about the relational structure in the target domain that can be extended from the base domain (Gentner, 1983).

We suggest that infants undergo the process of structural alignment when they compare exemplars of a spatial relation during familiarization and, if able to encode the spatial relation, can then compare it from a familiar instance to a novel one. That is, infants presented with two or more exemplars of a spatial relation make correspondences between the objects and the relations across the examples. For example, in order to generalize a support relation from "a car on a truck" to "a pig on a stand" as shown in Fig. 3.1, infants should align (or match) the car to the pig and the truck to the stand. Moreover, the supporting role of the truck should be compared to the supporting role of the stand. Also, infants' ability to notice the relation between the objects in one instance can affect whether they can generate inferences about the relation between the objects in the second instance. As shown by previous findings, infants can go beyond creating a temporary match between objects across two instances and create a common spatial relational structure whose objects are lacking in specific attributes, which is an abstract categorical representation of a spatial relation.

However, infants do not always succeed in the categorization or generalization of spatial relations. The theory of structural alignment provides two possibilities about where difficulties may arise. Infants may have a difficulty in establishing the object correspondences across the two examples (possibly due to a lack of surface-level similarities), or infants may be misled by object matches across the two instances that are in different relational structures. Later in this section, we revisit these possibilities to explain infants' success or failure in spatial categorization.

If the process of structural alignment underlies infants' categorization of spatial relations, then the objects used to depict a spatial relational structure should play a central role in infants' ability to align and compare scenes, a process that highlights the shared relational structure between two instances (Gentner, 1983; Gentner & Markman, 1997; Markman & Gentner, 1993). For instance, the more similar the corresponding objects are, the more likely one is to align the instances and note the common relational structure.

Many results from other labs support this view, and we have results from our lab from both infants and young children. From infants, we find that the objects do shape infant spatial categorization. First, comparison of studies that employed objects having different levels of similarity supports the view that similarity between elements influences infant spatial categorization. In some studies (e.g., Casasola & Cohen, 2002), infants were habituated to exemplars of support relations in which corresponding objects shared a low degree of similarity (e.g., car, cup, roundman, and turtle). Infants at 18 months failed to form the support category in this study. In other studies (e.g., Casasola, 2005; Casasola & Park, 2013), infants were habituated to exemplars where corresponding objects were snowman-shaped animals or human figures of similar size (see Fig. 3.2). In this case, infants of 14 months successfully formed the support category. In still other studies (e.g., Park et al., 2012; Park & Casasola, 2015), corresponding objects were highly similar in that the referent objects (the supportive objects) were larger tissue boxes and the figure objects (the supported objects) were smaller square blocks. In these studies, not only 14-month-old but also 8-month-old infants successfully formed the support category. Together, the findings support the view that high similarity between elements results in high alignability (easier establishment of object correspondences) and that objects have an important role in recognizing a common relational structure.

Similar patterns of results have been found with preschool-aged children: Objects depicting a spatial relation impact preschoolers' transfer of spatial relations. Park and Casasola (2017) examined 4- and 5-year-old children's spatial reasoning in a match-to-sample task, manipulating the objects in the task (abstract geometric shapes, line drawings of realistic objects, or both). Children generalized the target spatial configuration (i.e., on, in, above) more easily when the sample used geometric shapes and the response options used realistic objects than the reverse (i.e., realistic-object sample to geometric-shape choices). With within-type stimuli (i.e., sample and choices were both geometric shapes or both realistic objects), 5-year-old, but not 4-year-old, children generalized the spatial relations more easily with geometric shapes than realistic objects. This advantage of perceptually sparse, abstract objects in relational transfer was consistent with findings from other relational transfer tasks (Kaminski & Sloutsky, 2010; Kaminski, Sloutsky, & Heckler, 2008; Uttal et al., 2013; Uttal, Scudder, & DeLoache, 1997). We posit that the symbolic simplicity of abstract objects allows these object to be flexibly represented and makes it easier for them to be easily conceived as referring to something else (e.g., a circle can refer to the sun, a clock, and so on). Consequently, when an abstract object is used as the sample in a task, object correspondences between the compared scenes may be more readily formed, enabling the transfer of the common relation. In contrast, realistic objects may be less flexible in their representations and, consequently, are less likely to be treated as referring to other objects and serving as a symbol (e.g., Uttal et al., 1997), creating a more difficult task than the comparable task with abstract objects. Therefore, even if the learner has recognized the intended relation in a realistic instance, that relational knowledge may not translate as readily to another instance (Kaminski, Sloutsky, & Heckler, 2006).

In addition, Park and Casasola (unpublished dissertation) administered the same spatial analogy task to 3-year-old children. Interestingly, 3-year-old children did not show the advantage of abstract geometric shapes over realistic objects and in fact failed to transfer the spatial configurations, *on* and *in*, when they were depicted with abstract geometric shapes. They could transfer the relations of *in*, when they were depicted with realistic objects. Anecdotal records of children's linguistic responses show that 3-year-old children may not have perceived the scene (e.g., a circle in a larger circle, a circle on a larger circle) as being composed of two objects (the figure and referent). Rather, they perceived them as one object such as a donut and a snowman. These findings suggest that younger preschool children's failure in noticing the spatial relation between two objects may cause difficulties in forming correspondences across examples, which in turn causes their failure to transfer spatial relations to new examples.

Similarity between the corresponding elements (i.e., objects) also strongly influences preschoolers' appreciation of relational commonalities, including commonalities in spatial relations. The more similar the corresponding objects are, the more likely young children are to note the common relational structure between the elements (e.g., DeLoache, Kolstad, & Anderson, 1991; Gentner & Toupin, 1986). When the object surface similarity conflicts with similarity of relational roles, preschool children are drawn to the object matches rather than relational matches, with mapping of relational roles hindered (Loewenstein & Gentner, 2005). Also, another example of object similarity facilitating mapping comes from DeLoache et al. (1991), who showed that preschoolers performed better when the component objects in two spaces were similar to each other than when they were dissimilar. Thus, the appearance, features, and similarity of objects can either support or hinder young children's recognition of spatial relational commonalities.

Furthermore, preschool children's far mapping of spatial relations is facilitated by their close mapping, a finding that supports progressive alignment in the domain of spatial relations (Loewenstein & Gentner, 2001). To our knowledge, a direct test of progressive alignment in infants' spatial categorization has not been conducted yet, outlining a task for future research. For example, a study can be designed to test whether infants of 18 months can generalize the support relations to the widely different objects and to the various subtypes of support relations if the stimuli present the support relations from the highly similar ones (in terms of both object similarity and relational similarity) gradually to the highly dissimilar ones.

Thus far, we have reviewed the structural alignment theory and emphasized the importance of objects in noticing relational structures of instances. However, does the theory mean that learners always move from a component-level comparison to a structure-level comparison when seeing two instances? Or can the learner start by comparing the structures of the instances? Perhaps both are possible: In some cases, the comparison may occur in a top-down manner, from a higher-order commonality to lower-level commonalities; in other cases, comparison may occur in a bottom-up manner. It might depend on the relative salience of the common structure and the learner's prior relational knowledge (allowing easy access to the relational structure information from a single exemplar). Thus, one possibility is that the younger or more inexperienced we are (novice cognizer), the more surface-level commonalities (objects' perceptual commonalities such as shape and color) we attend to, rather than larger structural commonalities. Therefore, when comparing two instances, younger minds rely more on starting with building correspondences between the elements of the relational structure.

How Could the Other Theories in This Area Impact or Contribute to Your Findings?

Although we posited structural alignment as the central mechanism underlying infants' spatial categorization and young children's transfer of spatial relations, several other mechanisms might have worked together to explain some of these findings of infant spatial categorization. First, it is possible that infants' formation of an abstract categorical representation of a spatial relation benefits from their statistical learning (Marcus, Vijayan, Bandi Rao, & Vishton, 1999; Saffran, Aslin, & Newport, 1996), the ability to extract a rule or a set of regularities from input, using statistical properties of input. Research has shown infants' ability to track sequential statistical information with visual shapes (Fiser & Aslin, 2002; Kirkham, Slemmer, & Johnson, 2002). Thus, our infant participants might have similarly learned a set of sequential rules during the habituation phase, the rules such as the presence of two objects (a smaller one and a larger one, placed side by side), a hand coming in, and then the smaller object being picked up and placed on top of the larger object. Following the habituation phase, when the infant witnessed the rule violated, they responded to it with increased looking time. Studies to date have not tested between these two possibilities, and it remains to be seen if structural alignment or statistical learning best captures the pattern of results reported in infants' spatial categorization.

Another possible contributing factor to our findings is infants' forgetting. The spacing effect refers to findings showing that abstraction and generalization can be facilitated by interleaving the exemplar learnings and allowing learners time for forgetting (Vlach, 2014). To our knowledge, this effect has not been tested yet in the studies of infant spatial categorization. However, the theory raises an interesting possibility that forgetting (due to the wide variety of factors rather than a temporal gap between learning sessions) also contributes to our finding. More specifically, note that increasing the number of exemplars promotes 10-month-old infants' generalization of spatial relations to a novel instance. Possibly, with two exemplars repeatedly shown during the habituation phase, 10-month-old infants may be able to remember all the details of the objects and store object-specific representations in their memory. In contrast, with six exemplars, although they too are repeatedly shown, it may be hard for them to retain all the object details, namely, forgetting the continuously variant (or less frequently appearing) elements of the scenes. This forgetting may have led the infants to consider the test stimulus presenting a novel (but somewhat similar) pair of objects in a familiar spatial relation as familiar. The same explanation can also be applied to the finding about the impact of object perceptual features on infants' spatial categorization of support relations (Park & Casasola, 2015). Namely, infants formed the abstract categorical representation of support relations with perceptually complex objects, but not with simple objects.

Some of our findings regarding infant spatial categorization may be better explained by considering not only structural alignment theory but also other theories alongside. For example, our finding that infants, particularly those of 8 months, formed an abstract categorical representation of support relations in the complex condition but not in the simple condition may be explained by the goldilocks effect (Kidd, Piantadosi, & Aslink, 2012). That is, a moderate level of visual complexity may be ideal for the abstraction of a spatial relation. Indeed, in a preliminary study where we added more perceptual features including googly eyes to the complex objects, infants of 10 months failed to generalize a support relation to a novel pair of objects. Moreover, the object categorization literature suggests that nearly identical exemplars yield a highly inclusive category and inhibit abstraction (Namy, Gentner, & Clepper, 2007), showing that too close exemplars do not promote abstraction and generalization. Given that the stimuli in the simple condition varied only in color, infants' abstraction of the spatial relation in this condition may have been inhibited.

Can Infant Spatial Categorization Inform Other Types of Spatial Learning?

In conclusion, we consider how the theories and findings with infant spatial categorization may be extended to other skills. Although infants' attention to and categorization of spatial relations can be argued to be part of spatial cognition, the results from studies of infant spatial categorization are more often discussed with respect to non-spatial than spatial skills. For example, studies examining infants' and children's analogical reasoning have noted similarities in how generalization is shaped by the type of objects in the events (Ferry et al., 2015; Park & Casasola, 2017). Studies examining infant categorization of manner and path have noted similarities in the developmental changes in this skill and infant spatial categorization, including noting infants' greater ease in attending to some features of these motion events over others and in the facilitative effect of including labels during the familiarization phase (e.g., Pruden et al., 2012, 2013; Pulverman et al., 2008). How infants encode and form concepts of spatial relations also has been discussed in relation to advances in young children's skill in using spatial knowledge with real objects (e.g., accurately inserting shapes through openings) and discussed extensively with respect to the early acquisition of spatial language (e.g., Choi & Bowerman, 1991; Örnkloo & von Hofsten, 2007), extending infant spatial categorization to aspects of young children's motor skills and manipulation of objects and to their emerging language skills.

Interestingly, there has been less discussion of how infant spatial categorization relates to the development of other types of spatial skills. Can results from how infants form abstract representations of spatial relations be extended to consider developmental changes in the spatial domain more broadly? We argue that it can. Infant spatial categorization and, in particular, how multiple exemplars can shape infant spatial categorization can lend insight into children's performance on other types of spatial skills. One example, mentioned previously, is young children's spatial reasoning. Park and Casasola (2017) documented an effect of the objects in young children's ability to generalize a target image with a particular spatial relation to a novel instance, similar to way that specific object characteristics can shape infant spatial categorization. As another example, consider mental rotation, the ability to imagine the appearance of an object when in a different orientation. An interesting question is whether there may be some degree of overlap between infant spatial categorization and their later mental rotation. For both skills, the objects that depict a spatial relation and those to be mentally rotated can shape performance on each task (e.g., Dalecki, Hoffmann, & Bock, 2012; Thomas, Dalecki, & Abeln, 2013; Park & Casasola, 2015), pointing to a degree of intersection between the two skills.

Extrapolating from studies of infant spatial categorization to the study of other spatial skills allows us to outline possibilities for how to promote these other types of spatial skills, creating a road map of how we might manipulate experience with particular exemplars to promote spatial reasoning and mental rotation. Although manipulating these exemplars may not ultimately promote other types of spatial skills to the same degree as infant spatial categorization, they do offer an exciting approach to begin to bridge work in different areas of spatial cognition.

References

Antell, S. E., Caron, A. J., & Myers, R. S. (1985). Perception of relational invariants by newborns. *Developmental Psychology, 21*, 942–948. https://doi.org/10.1037/0012-1649.21.6.942

Antell, S. G., & Caron, A. J. (1985). Neonatal perception of spatial relationships. *Infant Behavior and Development, 8*, 15–23.

Baillargeon, R., & Wang, S. (2002). Event categorization in infancy. *Trends in Cognitive Sciences, 6*, 85–93. https://doi.org/10.1016/S1364-6613(00)01836-2

Baillargeon, R. (2004). Infants' physical world. *Current Directions in Psychological Science, 13*, 89–94. https://doi.org/10.1111/j.0963-7214.2004.00281.x

Behl-Chadha, G., & Eimas, P. D. (1995). Infant categorization of left-right spatial relations. *British Journal of Developmental Psychology, 13*, 69–79.

Bomba, P. C., & Siqueland, E. R. (1983). The nature and structure of infant form categories. *Journal of Experimental Child Psychology, 35*, 294–328. https://doi.org/10.1016/0022-0965(83)90085-1

Casasola, M. (2005). When less is more: How infants learn to form an abstract categorical representation of support. *Child Development, 76*, 279–290. https://doi.org/10.1111/j.1467-8624.2005.00844.x

Casasola, M. (2008). The development of infants' spatial categories. *Current Directions in Psychological Science, 17*, 21–25. https://doi.org/10.1111/j.1467-8721.2008.00541.x

Casasola, M. (2017). Above and beyond objects: The development of infants' spatial concepts. *Advances in Child Development and Behavior, 54*, 87–121. https://doi.org/10.1016/bs.acdb.2017.10.007

Casasola, M., & Ahn, Y. A. (2017). What develops in infants' spatial categorization? Korean infants' categorization of containment and tight-fit relations. *Child Development, 89*(4), e382–e396. https://doi.org/10.1111/cdev.12903

Casasola, M., & Cohen, L. B. (2002). Infant categorization of containment, support and tight-fit spatial relationships. *Developmental Science, 5*, 247–264. https://doi.org/10.1111/1467-7687.00226

Casasola, M., Cohen, L. B., & Chiarello, E. (2003). Six-month-old infants' categorization of containment spatial relations. *Child Development, 74*, 679–693. https://doi.org/10.1111/1467-8624.00562

Casasola, M., & Park, Y. (2013). Developmental changes in infant spatial categorization: When more is best and when less is enough. *Child Development, 84*, 1004–1019. https://doi.org/10.1111/cdev.12010

Casasola, M., Bhagwat, J., Doan, S. N., & Love, H. (2017). Getting some space: Infants' and caregivers' containment and support spatial constructions during play. Journal of Experimental Child Psychology, 159, 110–128. https://doi.org/10.1016/j.jecp.2017.01.012

Childers, J. B. (2011). Attention to multiple events helps 2 1/2-year-olds extend new verbs. *First Language, 31*, 3. https://doi.org/10.1177/0142723710361825

Childers, J. B., Paik, J. H., Flores, M., Lai, G., & Dolan, M. (2016). Does variability across events affect verb learning in English, Mandarin, and Korean? *Cognitive Science, 41*, 1–23. https://doi.org/10.1111/cogs.12398

Choi, S. (2006). Influence of language-specific input on spatial cognition: Categories of containment. *First Language, 26*, 207–232.

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.

Dalecki, M., Hoffmann, U., & Bock, O. (2012). Mental rotation of letters, body parts and complex scenes: Separate or common mechanisms? *Human Movement Science, 31*, 1151–1160. https://doi.org/10.1016/j.humov.2011.12.001

DeLoache, J. S., Kolstad, V., & Anderson, K. N. (1991). Physical similarity and young children's understanding of scale models. *Child Development, 62*, 111–126. https://doi.org/10.2307/1130708

Falkenhainer, B., Forbus, K. D., & Gentner, D. (1989). The structure-mapping engine: Algorithm and examples. *Artificial Intelligence, 41*, 1–63. https://doi.org/10.1016/0004-3702(89)90077-5

Ferry, A. L., Hespos, S. J., & Gentner, D. (2015). Prelinguistic relational concepts: Investigating analogical processing in infants. *Child Development, 86*, 1386–1405. https://doi.org/10.1111/cdev.12381

Fiser, J., & Aslin, R. N. (2002). Statistical learning of higher-order temporal structure from visual shape sequences. *Journal of Experimental Psychology. Learning, Memory & Cognition, 28*, 458–467. https://doi.org/10.1037/0278-7393.28.3.458

Forbus, K. D., & Gentner, D. (1983). Learning physical domains: Towards a theoretical framework. *Proceedings of the 1983 International Machine Learning Workshop*, Monticello, IL, June 1983.

Gava, L., Valenza, E., & Turati, C. (2009). Newborns' perception of left – right spatial relations. *Child Development, 80*, 1797–1810.

Gentner, D. (1983). Structure-mapping: A theoretical framework for analogy. *Cognitive Science, 7*, 155–170.
Gentner, D. (2005). The development of relational category knowledge. In L. Gershkoff-Stowe & D. H. Rakison (Eds.), *Building object categories in developmental time. Carnegie Mellon symposia on cognition* (pp. 245–275). Mahwah, NJ: Lawrence Erlbaum Associates Publishers.
Gentner, D., Loewenstein, J., & Thompson, L. (2003). Learning and transfer: A general role for analogical encoding. *Journal of Educational Psychology, 95*, 393–408. https://doi.org/10.1037/0022-0663.95.2.393
Gentner, D., & Markman, A. B. (1997). Structure mapping in analogy and similarity. *American Psychologist, 52*, 42–56. https://doi.org/10.1037/0003-066X.52.1.45
Gentner, D., & Toupin, C. (1986). Systematicity and surface similarity in the development of analogy. *Cognitive Science, 10*, 277–300. https://doi.org/10.1207/s15516709cog1003_2
Gick, M. L., & Holyoak, K. J. (1980). Analogical problem solving. *Cognitive Psychology, 12*, 306–355. https://doi.org/10.1016/0010-0285(80)90013-4
Gick, M. L., & Holyoak, K. J. (1983). Schema induction and analogical transfer. *Cognitive Psychology, 15*, 1–38. https://doi.org/10.1016/0010-0285(83)90002-6
Göksun, T., Hirsh-Pasek, K., & Golinkoff, R. M. (2010). Trading spaces: Carving up events for learning language. *Perspectives in Psychological Science, 13*, 431–436. https://doi.org/10.1177/1745691609356783
Gómez, R. L. (2002). Variability and detection of invariant structure. *Psychological Science, 13*, 431–436. https://doi.org/10.1111/1467-9280.00476
Hespos, S. J., & Baillargeon, R. (2001). Infants' knowledge about occlusion and containment events: A surprising discrepancy. *Psychological Science, 12*, 141–147.
Hespos, S. J., & Piccin, T. B. (2009). To generalize or not to generalize: Spatial categories are influenced by physical attributes and language. *Developmental Science, 12*, 88–95. https://doi.org/10.1111/j.1467-7687.2008.00749.x
Hespos, S. J., & Spelke, E. S. (2004). Conceptual precursors to language. *Nature, 430*, 453–456. https://doi.org/10.1038/nature02634
Kaminski, J. A., & Sloutsky, V. M. (2010). Concreteness and relational matching in preschoolers. In S. Ohlsson & R. Catrambone (Eds.), *Proceedings of the XXXII annual conference of the cognitive science society* (pp. 335–340). Mahwah, NJ: Erlbaum.
Kaminski, J. A., Sloutsky, V. M., & Heckler, A. F. (2006). Do children need concrete instantiations to learn an abstract concept? In R. Sun & N. Miyake (Eds.), *Proceedings of the XXVIII annual conference of the cognitive science society* (pp. 411–416). Mahwah, NJ: Erlbaum.
Kaminski, J. A., Sloutsky, V. M., & Heckler, A. F. (2008). The advantage of abstract examples in learning math. *Science, 320*, 454–455. https://doi.org/10.1126/science.1154659.
Kidd, C., Piantadosi, S. T., & Aslink, R. N. (2012). The goldilocks effect: Human infants allocate attention to visual sequences that are neither too simple nor too complex. *PLoS One, 7*, e36399.
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. https://doi.org/10.1016/S0010-0277(02)00004-5
Konishi, H., Pruden, S. M., Golinkoff, R. M., & Hirsh-Pasek, K. (2016). Categorization of dynamic realistic motion events: Infants form categories of path before manner. *Journal of Experimental Child Psychology, 152*, 54–70. https://doi.org/10.1016/j.jecp.2016.07.002
Kotovsky, L., & Gentner, D. (1996). Comparison and categorization in the development of relational similarity. *Child Development, 67*, 2797–2822. https://doi.org/10.1111/1467-8624.ep9706244834
Kovack-Lesh, K. A., & Oakes, L. M. (2007). Hold your horses: How exposure to different items influences infant categorization. *Journal of Experimental Child Psychology, 98*, 69–93. https://doi.org/10.1016/j.jecp.2007.05.001
Loewenstein, J., & Gentner, D. (2001). Spatial mapping in preschoolers. Close comparisons facilitate far mappings. *Journal of Cognition & Development, 2*, 189–219.
Loewenstein, J., & Gentner, D. (2005). Relational language and the development of relational mapping. *Cognitive Psychology, 50*, 315–353. https://doi.org/10.1016/j.cogpsych.2004.09.004

Maguire, M. J., Hirsh-Pasek, K., Golinkoff, R. M., & Brandone, A. C. (2008). Focusing on the relation: Fewer exemplars facilitate children's initial verb learning and extension. *Developmental Science, 11*, 628–634. https://doi.org/10.1111/j.1467-7687.2008.00707.x

Marcinowski, E. C., & Campbell, J. M. (2017). Building on what you have learned. *International Journal of Behavioral Development, 41*, 341–349. https://doi.org/10.1177/0165025416635283

Marcus, G. F., Vijayan, S., Bandi Rao, S., & Vishton, P. M. (1999). Rule learning by seven-month-old infants. *Science, 283*, 77–80. https://doi.org/10.1126/science.283.5398.77

Markman, A. B., & Gentner, D. (1993). Structural alignment during similarity comparisons. *Cognitive Psychology, 25*, 431–467. https://doi.org/10.1006/cogp.1993.1011

McDonough, L., Choi, S., & Mandler, J. M. (2003). Understanding spatial relations: Flexible infants, lexical adults. *Cognitive Psychology, 46*, 229–259. https://doi.org/10.1016/S0010-0285(02)00514-5

Namy, L. L., Gentner, D., & Clepper, L. E. (2007). How close is too close? Alignment and perceptual similarity in children's categorization. *Cognition, Brain, Behavior, 11*, 647–659.

Needham, A. W., Dueker, G., & Lockhead, G. (2005). Infants' formation and use of categories to segregate objects. *Cognition, 94*, 215–240. https://doi.org/10.1016/j.cognition.2004.02.002

Newcombe, N. S., & Huttenlocher, J. (2000). *Making space: The development of spatial representation and reasoning*. Cambridge, MA: The MIT Press.

Örnkloo, H., & von Hofsten, C. (2007). Fitting objects into holes: On the development of spatial cognition skills. *Developmental Psychology, 43*, 404–416. https://doi.org/10.1037/0012-1649.43.2.404

Oudgenoeg-Paz, O., Boom, J., Volman, M. C. J. M., & Leseman, P. P. M. (2016). Development of exploration of spatial-relational object properties in the second and third years of life. *Journal of Experimental Child Psychology, 146*, 137–155. https://doi.org/10.1016/j.jecp.2016.02.005

Oudgenoeg-Paz, O., Leseman, P. P. M., & Volman, M. C. J. M. (2015). Exploration as a mediator of the relation between the attainment of motor milestones and the development of spatial cognition and spatial language. *Developmental Psychology, 51*, 1241–1253. https://doi.org/10.1037/a0039572

Park, Y., & Casasola, M. (2017). The impact of object type on the spatial analogies in Korean preschoolers. *Cognitive Psychology, 94*, 53–66. https://doi.org/10.1016/j.cogpsych.2017.02.001

Park, Y., & Casasola, M. (2015). Plain or decorated? Object visual features matter in infant spatial categorization. *Journal of Experimental Child Psychology, 140*, 105–119. https://doi.org/10.1016/j.jecp.2015.07.002

Park, Y., & Casasola, M. (2013). The simplest objects may not be the best for infant spatial learning. Oral presentation given at the biennial meeting of the Society for Research in Child Development, Seattle, WA.

Park, Y., Casasola, M., & Kim, J. W. (2012). Do simple objects facilitate infants' formation of a spatial category? *Child Studies in Diverse Contexts, 2*, 77–90. https://doi.org/10.5723/csdc.2012.2.2.077

Pruden, S. M., Roseberry, S., Göksun, T., Hirsh-Pasek, K., & Golinkoff, R. M. (2013). Infant categorization of path relations during dynamic events. *Child Development, 84*, 331–345. https://doi.org/10.1111/j.1467-8624.2012.01843.x

Pruden, S. M., Goksun, T., Roseberry, S., Hirsh-Pasek, K., & Golinkoff, R. M. (2012). Find your manners: How do infants detect the invariant manner of motion in dynamic events? Child Development, 83, 977–991. https://doi.org/10.1111/j.1467-8624.2012.01737.x

Pulverman, R., Song, L., Hirsh-Pasek, K., Pruden, S. M., & Golinkoff, R. M. (2013). Preverbal infants' attention to manner and path: Foundations for learning relational terms. *Child Development, 84*, 241–252. https://doi.org/10.1111/cdev.12030

Pulverman, R., Golinkoff, R. M., Hirsh-Pasek, K., & Buresh, J. S. (2008). Infants discriminate manners and paths in non-linguistic dynamic events. Cognition, 108, 825–830. https://doi.org/10.1016/j.cognition.2008.04.009

Quinn, P. C. (1994). The categorization of above and below spatial relations by young infants. *Child Development, 65*, 58–69.

Quinn, P. C. (2003). Concepts are not just for objects: Categorization of spatial relation information by infants. In D. H. Rakison & L. M. Oakes (Eds.), *Early category and concept development: Making sense of the blooming, buzzing confusion* (pp. 50–76). New York: Oxford University Press.

Quinn, P. C. (2005). Developmental constraints on the representation of spatial relation information: Evidence from preverbal infants. In L. Carlson & E. van der Zee (Eds.), *Functional features in language and space: Insights from perception, categorization, and development* (pp. 293–309). New York: Oxford University Press.

Quinn, P. C. (2007). On the infant's prelinguistic conception of spatial relations: Three developmental trends and their implications for spatial language learning. In J. M. Plumert & J. P. Spencer (Eds.), *The emerging spatial mind* (pp. 117–141). New York: Oxford University Press. https://doi.org/http://dx.doi.org.proxy.library.cornell.edu/10.1093/acprof:oso/9780195189223.003.0006.

Quinn, P. C. (2012). Evidence for mental subdivision of space by infants: 3- to 4-month-olds spontaneously bisect a small-scale area into left and right categories. *Psychonomic Bulletin & Review, 19*, 449–455. https://doi.org/10.3758/s13423-012-0243-9

Quinn, P. C., Adams, A., Kennedy, E., Shettler, L., & Wasnik, A. (2003). Development of an abstract category representation for the spatial relation between in 6- to 10-month-old infants. *Developmental Psychology, 39*, 151–163. https://doi.org/10.1037/0012-1649.39.1.151

Quinn, P. C., & Bhatt, R. S. (2005). Learning perceptual organization in infancy. *Psychological Science, 16*, 511–515. https://doi.org/10.1111/j.0956-7976.2005.01567.x

Quinn, P. C., Cummins, M., Kase, J., Martin, E., & Weissman, S. (1996). Development of categorical representations for above and below spatial relations in 3- to 7-month-old infants. *Developmental Psychology, 32*, 942–950. https://doi.org/10.1037/0012-1649.32.5.942

Quinn, P. C., Doran, M. M., & Papafragou, A. (2011). Does changing the reference frame affect infant categorization of the spatial relation BETWEEN? *Journal of Experimental Child Psychology, 109*, 109–122. https://doi.org/10.1016/j.jecp.2010.10.001

Quinn, P. C., Norris, C. M., Pasko, R. N., Schmader, T. M., & Mash, C. (1999). Formation of a categorical representation for the spatial relation between by 6- to 7-month-old Infants. *Visual Cognition, 6*, 569–585. https://doi.org/10.1080/135062899394948

Quinn, P. C., Polly, J. L., Furer, M. J., & Dobson, V. (2002). Young infants' performance in the object-variation version of the above-below categorization task: A result of perceptual distraction or conceptual limitation? *Infancy, 3*, 323–347. https://doi.org/10.1207/S15327078IN0303_3

Rakison, D. H., & Oakes, L. M. (2003). *Early category and concept development: Making sense of the blooming, buzzing confusion*. New York: Oxford University Press.

Rigney, J., & Wang, S. H. (2015). Delineating the boundaries of infants' spatial categories: The case of containment. *Journal of Cognition and Development, 16*, 420–441. https://doi.org/10.1080/15248372.2013.848868

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science, 274*, 1926–1928. https://doi.org/10.1126/science.274.5294.1926

Song, L., Pruden, S. M., Golinkoff, R. M., & Hirsh-Pasek, K. (2016). Prelinguistic foundations of verb learning: Infants discriminate and categorize dynamic human actions. *Journal of Experimental Child Psychology, 151*, 77–95. https://doi.org/10.1016/j.jecp.2016.01.004

Shutts, K., € Ornkloo, H., Von Hofsten, C., Keen, R., & Spelke, E. S. (2009). Young children's representations of spatial and functional relations between objects. Child Development, 80, 1612–1627. https://doi.org/10.1111/j.1467-8624.2009.01357.x

Thomas, M., Dalecki, M., & Abeln, V. (2013). EEG coherence during mental rotation of letters, hands and scenes. *International Journal of Psychophysiology, 89*, 128–135. https://doi.org/10.1016/j.ijpsycho.2013.06.014

Uttal, D. H., Amaya, M., Maita, M. R., Hand, L. L., Cohen, C. A., O'Doherty, K., et al. (2013). It works both ways: Transfer difficulties between manipulatives and written subtraction solutions. *Child Development Research, 2013*, 1–13. https://doi.org/10.1155/2013/216367

Uttal, D. H., Scudder, K. V., & DeLoache, J. (1997). Manipulatives as symbols: A new perspective on the use of concrete objects to teach mathematics. *Journal of Applied Developmental Psychology, 18*, 37–54.

Vlach, H. (2014). The spacing effect on children's generalization of knowledge: Allowing children time to forget promotes their ability to learn. *Child Development Perspectives, 8*, 163–168. https://doi.org/10.1111/cdep.12079

Vukatana, E., Graham, S. A., Curtin, S., & Zepeda, M. S. (2015). One is not enough: Multiple exemplars facilitate infants' generalizations of novel properties. *Infancy, 20*, 548–575. https://doi.org/10.1111/infa.12092

Wang, S., & Baillargeon, R. (2008). Can infants be "taught" to attend to a new physical variable in an event category? The case of height in covering events. *Cognitive Psychology, 56*, 284–326. https://doi.org/10.1016/j.cogpsych.2007.06.003

Chapter 4
How the Demands of a Variable Environment Give Rise to Statistical Learning

Erik D. Thiessen

Abstract Language inherently requires learners to process variability in the input, as no two utterances, sentences, or speakers sound identical. Statistical learning, the ability to identify structure in the input by detecting regular patterns, is a potential mechanism that may help infants and adults cope with, and benefit from, the variability in linguistic input. In this chapter, I provide an overview of statistical learning phenomena, including identifying units (such as words) from the co-occurrence of sounds and discovering category membership from the frequency and variability of exemplars in the input. While there are many different statistical learning tasks, I propose that they share many commonalities that can be explained by viewing statistical learning as an emergent property of the way that information is stored, accessed, and integrated in memory. This perspective makes novel predictions about the process of language development and how it is related to more domain-general cognitive processes.

Language acquisition is a domain in which learning necessarily depends on multiple exemplars. In part, this is due to pragmatic constraints. Children are typically exposed to – and learning from – multiple caregivers and even from multiple languages. Moreover, even a single source of linguistic input (e.g., a single parent) will produce a significant amount of variability, due to factors like speech register (e.g., Biber, 1999), coarticulatory environment (Iskarous & Kavitskaya, 2010; Perkell & Matthies, 1992), and changes in the child's own maturational and linguistic development resulting in different language use contexts (e.g., Englund & Behne, 2006; Stern, Spieker, Barnett, & MacKain, 1983). Perhaps even more importantly, the nature of language itself requires multiple exemplars to learn. This is due to the fact that linguistic knowledge is generative (Chomsky, 1956). That is, a fluent language speaker uses their knowledge of language to produce novel, grammatically acceptable utterances. As such, language consists of a potentially infinite set of utterances and cannot be learned by rote rehearsal of a fixed set of exemplars.

E. D. Thiessen (✉)
Carnegie Mellon University, Pittsburgh, PA, USA
e-mail: thiessen@andrew.cmu.edu

J. B. Childers (ed.), *Language and Concept Acquisition from Infancy Through Childhood*, https://doi.org/10.1007/978-3-030-35594-4_4

Learning from multiple exemplars requires learners to deal with variability. The same underlying construct (e.g., a word) will rarely be expressed in an identical manner across different exemplars of that construct. The variability in language has been conceptualized as both an obstacle and a benefit to learning. Consider, as an example, the influence of coarticulation on phoneme production and perception. Coarticulation refers to the fact that production of a speech sound, in fluent speech, is influenced by the identity of preceding and following sounds (Hardcastle & Hewlett, 2006). This means that there is no single invariant form of a phoneme. Instead, every time a phoneme is produced, it will differ as a function of the sounds surrounding it (e.g., the phoneme /t/ is produced with more pronounced aspiration in syllable initial position, as in "top," than inside a consonant cluster, as in "stop"). This variability in the way a phoneme is articulated leads to a complex relationship between acoustic cues and phonemic identity. Cues vary greatly as a function of context, and there is often no single invariant cue that can serve to infallibly identify a phoneme. Indeed, the variability in acoustic realization of a phoneme is so great that some theories hold it to be intractable, arguing that identification of phonemes is instead done by reference to the (relatively less variable) neuromotor commands that give rise to articulation (Liberman, Cooper, Shankweiler, & Studdert-Kennedy, 1967). While a comprehensive exploration of the motor theory of speech perception is beyond the scope of the current discussion (for a review, see Galantucci, Fowler, & Turvey, 2006), it serves to illustrate the challenges that variability poses to speech perception.

Despite the difficulty posed by coarticulation-induced variability in the production of phonemes, coarticulation also plays an important facilitative role in speech perception. The effects of coarticulation provide listeners with cues about the identity of the preceding and following speech segments (for review, see Diehl, Lotto, & Holt, 2004). For example, when presented with a sound ambiguous between a/d/ and a/g/, listeners are able to use the characteristics of the preceding sound to help them identify the identity of the ambiguous phoneme (Lotto & Kluender, 1998). Similarly, coarticulation plays a role in word segmentation, as sounds within a word show more profound coarticulatory effects than sounds across word or (especially) phrase boundaries (e.g., Johnson & Jusczyk, 2001). Indeed, the effects of coarticulation are so informative that when listeners are presented with words removed from their coarticulatory context, word recognition rates drop by approximately half, even for familiar high-frequency words (Pollack & Pickett, 1964).

As the example of coarticulation indicates, the effect of variability across exemplars has a double-edged impact on learning. On one hand, variability presents a challenge to learners. It increases the complexity of the input and requires learners to identify those aspects of variation that are informative and those that are irrelevant (note that there is not a "one size fits all" solution to this problem, as in many cases variability that is irrelevant to one task is critical to another). On the other hand, variability facilitates some aspects of learning, such as generalization to novel exemplars (e.g., Lively, Logan, & Pisoni, 1993; Perry, Samuelson, Malloy, & Schiffer, 2010). Nor are the effects of variability across exemplars limited to speech perception. Indeed, it appears to be the case that this characterization of the role of variability – as both a boon to learning and an obstacle to be overcome – applies to many levels of linguistic organization. Consider, for example, the variability in the surface realizations of a unit such as "noun phrase." Just as in speech perception, learning of phrase structure appears to be affected by the variability of the input (e.g., Gómez & Gerken, 2000; Mintz, 2003).

A further challenge presented by variability in linguistic input is that variability has different implications as a function of context. This is clearly the case across different languages (and thus for bilingual learners); for example, the same variance in frequency that is informative about the difference between/r/ and /l/ in English is uninformative in Japanese (e.g., Best & Strange, 1992). But it is also the case within even a single language. Consider variation in voice onset time. In English (as in many other languages), the difference between voiced and voiceless consonants is primarily signaled by differences in voice onset time. Voiced consonants typically have a voice onset time around 0 msec; voiceless consonants typically have a voice onset time around 40 msec (e.g., Allen & Miller, 1999). When making categorical distinctions between phoneme identity, differences "within" a category (such as the difference between 5 and 10 msec) are typically ignored (e.g., Gerrits & Schouten, 2004). However, while variation within a category is uninformative for identifying phonemes, it is quite useful for identifying individual speakers, as different speakers have different idiosyncratic "preferred" or "modal" voice onset times (Allen, Miller, & DeSteno, 2003). That is, for some tasks, within-category differences in voice onset time are uninformative, while in other tasks (such as identifying a familiar speaker), within-category differences in voice onset time present useful information. This means that language users must learn when to rely on, and when to ignore, variation in the input.

In our work, we have proposed that language learners – especially, but not only, infant language learners – benefit from the distribution of information across multiple exemplars to discover which aspects of variation in the input are informative (Thiessen, Kronstein, & Hufnagle, 2013). The ability to identify informative structure across surface variability arises in part due to an ability often referred to as "statistical learning" (Saffran, Aslin, & Newport, 1996). In the remainder of this chapter, we will explore how statistical learning helps infants to discover regularities presented across multiple exemplars and the possibility that statistical learning results from general principles organizing human memory. We will begin with an overview of conditional statistical learning (e.g., discovering co-occurrence relations) and turn to an overview of distributional statistical learning (e.g., discovering category boundaries), before discussing how these different forms of statistical learning can be unified in a memory-based framework.

Conditional Statistical Learning and Language Acquisition

The term "statistical learning" is often taken to mean sensitivity to probabilistic relations among sequential elements in a continuous stream of input (e.g., Johnson & Seidl, 2008; Saffran et al., 1996). Indeed, there is now a great deal of evidence to suggest that infants and adults are sensitive to sequential co-occurrence probability in linguistic (and nonlinguistic) stimuli. This sensitivity can be used to identify words in a stream of fluent speech, insofar as sounds within a word are more likely to co-occur than sounds across word boundaries (e.g., Aslin, Saffran, & Newport, 1998; Hayes & Clark, 1970). Similarly, the co-occurrence of words can be a useful cue to phrase boundaries, to the extent that words within a phrase are more likely to co-occur than words across phrase boundaries (e.g., Thompson & Newport, 2007). Linguistic structure is characterized by a variety of sequential regularities at different levels of analysis, including phonemic, syllabic, and lexical (e.g., Swingley, 1999; Taylor & Black, 1998; Vitevitch & Luce, 2004).

As this discussion indicates, sensitivity to sequential probabilities can be useful for acquiring a variety of linguistic regularities. Even so, the role of sensitivity to conditional relations in language acquisition has been most thoroughly explored with respect to word segmentation, and a discussion of this role can help to illuminate the process of learning from probabilistic input. In a typical word segmentation study, participants are presented with a stream of syllables. These syllables are often computer synthesized, to ensure that there are no acoustic cues to the structure of the input. Unbeknownst to the participants, these syllables are organized into a set of words, such that when the first syllable of a word is heard, the other syllables invariably follow; at the end of a word, any of the other words in the language may follow. This means that syllables within a word are very likely to co-occur, while syllables across word boundaries are less likely to co-occur.

This likelihood of co-occurrence is often described in terms of *transitional probability*. Transitional probability is calculated as the frequency with which a pair (XY) occurs, in comparison with the rate at which an element within that pair (e.g., X) occurs; as such, transitional probabilities can range between 0 (XY never occurs) and 100% (XY occurs every time X occurs). Consider, for example, an artificial language made up of four words: *golabu*, *padoti*, *daropi*, and *bidaku*. Whenever the syllable "go" occurs, the syllables "la" and "bu" will always occur next. As such, the transitional probabilities within the word are 100%. By contrast, after the syllable "bu," any of the other three words in the language (*padoti*, *daropi*, or *bidaku*) can occur next; as such, the transitional probability at word boundaries is 33%. When presented with an artificial language like this, infants and adults learn to differentiate between sequences with high transitional probabilities (i.e., words) and items with lower transitional probabilities (e.g., a grouping of syllables that occurs across word boundaries, such as *tidaro)*, even with relatively brief exposure times (e.g., Aslin et al., 1998; Saffran et al., 1996; Thiessen & Saffran, 2004). Notice that this type of statistical structure requires that learners have multiple exposures to the words in different contexts. If the words are only heard once (e.g., *padotidaropibidakugolabu)*, transitional probabilities within words are identical (100%) to transitional probabilities across word boundaries. This suggests that hearing multiple exemplars of words, in different contexts, can be facilitative for word learning (c.f. Rost & McMurray, 2009).

Of course, the results of these kinds of experiments are only informative about natural language learning to the extent that statistical learning is involved in the acquisition of natural language. There are now a variety of converging lines of evidence to suggest that this is the case. One such line of evidence indicates that individual differences in statistical learning ability are correlated with individual differences in language ability (Misyak & Christiansen, 2012; Misyak, Christiansen, & Tomblin, 2010). This correlation suggests that, at minimum, some of the factors that influence variance in statistical learning also influence variance in language ability, indicating some shared set of processes between statistical learning and language. Alternatively, the correlation could indicate a relatively more direct link; if it is the case that statistical learning is a critical process in language acquisition (c.f. Romberg & Saffran, 2010), then individuals who are better at statistical learning may have an advantage in learning to use a language that persists into adulthood.

Another line of evidence that statistical learning is involved in the acquisition of natural language is that the representations that emerge from statistical segmentation tasks behave like lexical representations. That is, participants respond to the "words" that they have acquired from exposure to an artificial segmentation language in a similar manner as they respond to words from their native language. For example, infants are unsurprised when words from an artificial grammar appear surrounded by familiar English words (Saffran, 2001). Similarly, infants and adults learn labels for novel objects more easily when they have segmented those labels from an artificial language (Estes, Evans, Alibali, & Saffran, 2007; Mirman, Magnuson, Estes, & Dixon, 2008). Finally, one of the key features of lexical representations is that they are treated holistically (this is why, for example, participants can often identify a familiar word faster than the letters contained within the word; for discussion, see Paap, Newsome, McDonald, & Schvaneveldt, 1982). The "words" that participants learn from statistical segmentation experiments appear to be associated with representations of a similarly holistic nature. After learning a word like *golabu* from an artificial grammar, participants do not recognize components contained within the word such as *gola* or *labu* (Giroux & Rey, 2009; Orbán, Fiser, Aslin, & Lengyel, 2008).

Taken together, these disparate lines of evidence suggest that learning from sequential probabilistic relations is not just something that participants can do in a laboratory task. Rather, these results indicate that statistical learning is involved in at least some aspects of language acquisition. At current, the evidence for this claim is stronger for lexical and phonemic learning (e.g., Giroux & Rey, 2009; Saffran, 2001; Thiessen & Saffran, 2007) than it is for syntactic learning. That is, there are relatively fewer studies investigating how statistical learning is involved in the learning of syntactic regularities (though see Mintz, 2003; Onnis & Thiessen, 2013; Thompson & Newport, 2007; van den Bos, Christiansen, & Misyak, 2012). In part, this is due to the complexity of syntax, making artificial analogs of syntactic systems difficult to instantiate in a laboratory setting.

Another set of unsettled questions about the link between statistical learning and language acquisition relate to the nature of the processes or mechanisms underlying statistical learning. Sensitivity to conditional statistical relations has been studied in a wide variety of tasks. These tasks differ in their stimulus modality (e.g., audio vs. visual stimuli; Conway & Christiansen, 2006), the response expected of participants (e.g., passively listening vs. regularly responding to some aspect of the input; Hunt & Aslin, 2001), and the kind of statistical structure they present to be learned (e.g., words vs. categories; Finn, Lee, & Hudson Kam, 2014). The diversity of these tasks, and the differences between them, raise the question of whether the "statistical learning" studied in these different tasks is accomplished by the same underlying mechanism; an alternative possibility is that different tasks assess different (potentially completely independent) aspects of statistical learning.

To illustrate this point, consider the difference between tasks that assess conditional statistical relations (such as transitional probability) in sequential or simultaneous input. Most experiments assessing statistical learning have focused on sequential conditional relations – that is, whether an element of the input predicts a *subsequent* element of the input (Saffran et al., 1996). But learners are also sensitive to conditional relations among simultaneously presented elements – that is, whether two items are likely to co-occur together, as in a visual scene (e.g., Fiser & Aslin, 2005). Simultaneous conditional relations might be especially useful for discovering aspects of linguistic structure that are expected to co-occur, such as the relation between a referent and its label (e.g., Smith & Yu, 2008). But the extent to which simultaneous and sequential statistical structure are learned by the same (or different) underlying mechanism is an open question. Similarly, conditional statistical structure is present not only in linguistic or auditory input. Learners are sensitive to conditional statistical relations in many other kinds of stimuli, including musical tones (e.g., Saffran, Johnson, Aslin, & Newport, 1999), visual images (e.g., Kirkham, Slemmer, & Johnson, 2002; see also Johnson, this volume), and action sequences (e.g., Baldwin, Andersson, Saffran, & Meyer, 2008). Across these different kinds of input, however, there are important differences in the time course and output of learning (e.g., Conway & Christiansen, 2005; Slone & Johnson, 2015). It is as yet unclear whether these differences are due to differences in the underlying process of statistical learning or differences in the way that these stimuli are processed or represented independently of statistical learning (for discussion, see Frost, Armstrong, Siegelman, & Christiansen, 2015).

Distributional Statistics

The question of whether statistical learning can be thought of as a single, unified process is magnified when we consider that conditional relations (such as those described by transitional probabilities) are not the only kind of statistical structure to which learners are sensitive. In addition to conditional structure, humans are also sensitive to a set of statistical properties that have been termed *distributional statistics* (e.g., Thiessen et al., 2013) or *cross-situational statistics* (e.g., Smith & Yu, 2008) or *summary statistics* (e.g., Zhao, Ngo, McKendrick, & Turk-Browne, 2011). Broadly speaking, these terms refer to the ability to respond to the central tendency and variability of a set of exemplars (for a more extensive discussion, see Thiessen & Pavlik, 2013). Maye, Werker, and Gerken's (2002) research on phonological category learning provides a paradigmatic example of learning from distributional statistics. In their experiments, participants listened to a set of phonemes ranging in voice onset time from /d/ to (unaspirated) /t/. When exposed to a bimodal distribution, such that a prototypical /d/ and a prototypical /t/ occurred most frequently, infants were more likely to discriminate between /d/ and /t/. When exposed to a monomodal distribution, such that a sound intermediate between /d/ and /t/ occurred most frequently, infants were less likely to discriminate between /d/ and /t/. This work suggests that categorical boundaries can be shifted as a function of the distribution of exemplars in the input (e.g., Maye, Weiss, & Aslin, 2008; Imai & Childers, this volume; Oakes & Spalding, 1997).

Sensitivity to the frequency with which phonemic exemplars occur may play a role in infants' adaptation to the phonemic structure of the native language in the first year of life (e.g., Werker & Tees, 1984). Sounds near the prototypical center of a phonemic category that is productive within a language occur more frequently than sounds further from the center of the category or sounds that are ambiguous between two categories (Werker et al., 2007). The role of distributional statistics in language acquisition is not, however, limited to perceptual learning of sound categories. The frequency of exemplars in the input has been suggested to play a role in word learning (e.g., Thiessen & Yee, 2010; Tomasello, 2000) and in discovering syntactic regularities (e.g., Reber & Lewis, 1977). For each of these aspects of language, identifying categories in the input is key, and sensitivity to the frequency of exemplars in the input provides a useful cue to category membership.

In addition to frequency of exemplars, learners are also sensitive to the variability of exemplars in the input. When exposed to distributions with high variability, learners accept a wider range of exemplars as members of a category and are correspondingly less certain about category membership when asked to make judgments about stimuli near a category boundary for many kinds of stimuli, including both audio and visual stimuli (e.g., Clayards, Tanenhaus, Aslin, & Jacobs, 2008; Quinn, Eimas, & Rosenkrantz, 1993). When there is very low variability in the input, category boundaries are relatively sharper. Note that this discussion of sensitivity to variation across exemplars suggests that learners are able to detect within-category variation. Within-category differences were thought to be ignored according to classical theories of categorical perception (e.g., Liberman, Harris, Hoffman, & Griffith, 1957). Subsequent research, however, has demonstrated that learners are sensitive to within-category variation even with stimuli over which categorical perception can easily be obtained (e.g., McMurray, Tanenhaus, & Aslin, 2002; Miller & Volaitis, 1989; Pisoni & Tash, 1974). This suggests that learners encode both the category identity of exemplars and idiosyncratic features that are not necessary for category identification. We will return to this notion of encoding specificity in the next section.

A final distributional feature of the input to which learners are sensitive is the context in which exemplars occur; these contextual details can also serve as cues to category membership. For example, Thiessen (2007) found that infants were better able to use the categorical distinction between phonemes (such as /d/ vs. /t/) in word-learning contexts if they had previously experienced the phonemes in distinct lexical contexts (e.g., /d/ in *doggy* and *diaper*; /t/ in *tiger* and *toothbrush*). This is an example of a phenomenon known as *acquired distinctiveness*: when two similar stimuli occur in distinct contexts, learners distinguish between them more easily (Honey & Hall, 1989; James, 1890). This, in turn, makes learners more likely to treat the stimuli as members of different categories.

As with conditional statistical learning, there are a variety of unsettled questions about the nature of the mechanisms or processes underlying distributional statistical learning. Distributional learning, like conditional learning, has been measured with a variety of different tasks, and it is not clear that each of these tasks is tapping into the same set of processes (for discussion, see Perruchet & Pacton, 2006; Thiessen & Pavlik, 2016). Also, as with conditional statistical learning, it is clear that distributional statistical learning is available for many kinds of stimuli beyond linguistic stimuli, including nonlinguistic audio and visual stimuli (e.g., Dougherty & Haith, 2002; Lotto, Kluender, & Holt, 1997). It is less clear that the same set of processes are at work in each of these modalities, as different types of stimuli can yield different learning outcomes (e.g., Aslin & Newport, 2014; Johnson & Tyler, 2010; Saffran, Pollak, Seibel, & Shkolnik, 2007).

As this discussion indicates, many outstanding questions – for both conditional and distributional statistical learning – relate to the nature of the mechanisms or processes underlying learning (e.g., Frank, Goodman, & Tenenbaum, 2007; Thiessen et al., 2013). In the remainder of this chapter, we will argue that human memory presents a unifying framework for thinking about these different forms of statistical learning. Our claim is that statistical learning arises from processes that are necessary for memory, especially similarity-based activation, interference, and decay.

Representing Variability Across Exemplars

In an environment with multiple, variable exemplars, learners face a fundamental challenge: the need to recognize familiar exemplars while also generalizing prior experience to novel exemplars. This tension is well illustrated by the comparison of two fundamental theoretical approaches to human memory, prototype models and exemplar models. In a prototype model, memory consists of a set of prototypes, which are the “most average” or “most representative” members of a category (e.g., Posner & Keele, 1968; Rosch, 1975). For example, the representation of the category “dog” would consist of something like the “weighted average” of all the dogs a learner had ever seen in their life. This kind of representational system is excellent for generalization, because the representations that are stored are maximally similar to the distribution of category members in the environment. However, it is less useful for recognizing individual exemplars that one has seen in the past, because the representations formed of a category consist solely of the average of the category, rather than of individual exemplars of that category. Exemplar memory models show a complementary pattern of strengths and weaknesses. In an exemplar memory model, each individual exemplar a learner has experienced is stored in memory (e.g., Hintzman, 1984; Medin & Schaffer, 1978). This kind of memory system is excellent at recognizing familiar tokens, as each individual exemplar is stored in memory. However, it is less well suited to generalization. Because individual exemplars are stored in memory, rather than the average of a set of exemplars, novel exemplars may not be recognized as category members.

Exemplar memory models and prototype memory models were initially conceived as in opposition to each other (e.g., Smith & Minda, 1998). However, decades of research assessing these competing theoretical accounts failed to yield conclusive evidence in favor of one approach over the other (for discussion, see Vanpaemel, 2016). What this extensive body of research has done, though, is provide compelling evidence supporting the claims of *both* exemplar and prototype theory. For example, exemplar theory is consistent with speaker specificity effects; that is, utterances are better recognized when they are spoken by the same speaker as produced them initially, rather than by a different speaker, indicating that our representations for speech incorporate speaker-specific idiosyncrasies (e.g., Goldinger, 1998; Houston & Jusczyk, 2003). Conversely, there is ample evidence that humans form prototypical representations (e.g., Kruschke, 2005; Langlois & Roggman, 1990; Principe & Langlois, 2012) and indications that for some experimental tasks, prototype models provide a better fit to human learning than exemplar models (e.g., Minda & Smith, 2002; though see Zaki, Nosofsky, Stanton, & Cohen, 2003).

The fact that there is evidence in favor of both exemplar theory and prototype theory makes a degree of intuitive sense. Both recognizing familiar items (a strength of exemplar models) and generalizing to related but novel items (a strength of prototype models) are problems that organisms regularly face in the environment. As such, it is no great surprise that memory has developed in ways that facilitate solving both of these problems. Indeed, many current models of learning and memory involve representation of both exemplars and prototypes (e.g., Abbot-Smith & Tomasello, 2006; Smith, 2009; Thiessen & Pavlik, 2013). Some models accomplish this by using a single representational system, in which the representations can be recombined into prototypical, abstracted composites or maintained as separate exemplars, depending on the nature of the input and the task (e.g., McClelland & Rumelhart, 1985; Thiessen & Pavlik, 2016). Other models accomplish this via dual representational systems: one system for encoding exemplars and the other for encoding prototypical representations (e.g., McClelland, McNaughton, & O'Reilly, 1995).

As this discussion indicates, the idea of representing both exemplars and prototypes is not a novel one. We propose that the mechanisms and constraints provided by a system that encodes both exemplars and prototypes give rise to sensitivity to statistical structure in the input. That is to say, statistical learning is an emergent property of the more general characteristics of memory. If this is the case, the function and development of the neurological systems supporting memory should have clear implications for our understanding of statistical learning. This is consistent with an emerging body of literature suggesting that the activity of the human memory system is related to performance in a wide array of statistical learning paradigms composed of diverse stimulus characteristics (Batterink, 2017; Bornstein & Daw, 2012; Schapiro, Gregory, Landau, McCloskey, & Turk-Browne, 2014; Schapiro, Kustner, & Turk-Browne, 2012; Schapiro, Rogers, Cordova, Turk-Browne, & Botvinick, 2013; Schapiro, Turk-Browne, Norman, & Botvinick, 2016; Turk-Browne, Scholl, Chun, & Johnson, 2009; Turk-Browne, Scholl, Johnson, & Chun, 2010). In the next section, we will provide a broad overview of the neuroscience of memory as well as explore the consequences of the organization of this system for our understanding of statistical learning.

Encoding and Generalization in Memory

To account for sensitivity to prototypicality, the memory system must have some way of combining information across multiple representations. To illustrate how these processes work together, we will focus on a classic theory that attempts to explain how neural function gives rise to processing that is needed in statistical learning, the complementary learning system (CLS) hypothesis (McClelland et al., 1995). CLS theory posits that memory is divided into two distinct systems with complementary properties. The first system creates sparsely coded representations (i.e., all stimuli represented by a unique subset of neurons) that are robust to interference from future learning. This system is crucial for maintaining separate representations of events that share many similarities – for instance, remembering one's parking spot in a frequently visited parking lot – but it is not well suited for learning about the regularities that persist across these events because highly similar events are represented as distinct entities. The medial temporal lobe (MTL), especially the hippocampal formation (HF), is generally agreed to be the locus of pattern-separated representation (Bakker, Kirwan, Miller, & Stark, 2008; Clelland et al., 2009; Gilbert, Kesner, & Lee, 2001; Leutgeb, Leutgeb, Moser, & Moser, 2007; McHugh et al., 2007). A separate system with highly overlapping distributed representations, the neocortex, is ideal for uncovering statistical regularities because memories that share common features are represented similarly (Hinton, McClelland, & Rumelhart, 1986). As such, the CLS framework posits that the MTL and neocortex work together to support a highly flexible and adaptive memory system.

The CLS framework argues that statistical learning is the outcome of information consolidation in the neocortex (O'Reilly, Bhattacharyya, Howard, & Ketz, 2014; Winocur, Moscovitch, & Bontempi, 2010). The MTL rapidly encodes detailed, context-specific representations of the episodic content using a sparse coding system. The DG and CA3 subregions can employ a sparse coding system in which similar stimuli are represented by a unique subset of neurons because of the incredibly large number of neurons present in these subregions. When the neocortex receives projections from the MTL, the information is slowly transformed to accommodate a distributed coding system. The neocortex employs a distributed coding system because of the relatively small number of neurons available in this region compared to the DG and CA3. A distributed coding system, in which stimuli are represented over a larger number of neurons and the same set of neurons can participate in the representation of multiple pieces of information (i.e., population coding), is ideal for brain regions with fewer neurons. In this case, the most common features of the input are encoded across many neurons, and idiosyncratic details that are rarely encountered are discarded. This coding scheme naturally highlights similarity structures, enabling the neocortex to uncover the underlying statistical regularities that characterize the input. For instance, consider the parking lot example. The MTL encodes a particular day's parking location, whereas the neocortex stores all memories of previous parking locations within this parking lot. The neocortex serves as a general map of the most frequent parking locations. The CLS framework takes advantage of the unique connectivity between MTL and neocortex to address two issues commonly found with distributed coding schemes: to integrate new memories without overwriting old memories, there must be a system in place to mitigate interference from competing memories (e.g., AB-AC learning paradigm) and prevent catastrophic forgetting of old memories (McClelland et al., 1995; Rogers & McClelland, 2004). However, the CLS framework has been criticized for its claim that the discovery of regularities (via overlapping representations) relies on the cortex, which is thought to be a slower learning system – requiring hours or days to consolidate information (for discussion, see Norman & O'Reilly, 2003). This is somewhat inconsistent with evidence that learners can identify regularities from overlapping representations over the course of only a few minutes (e.g., Schapiro et al., 2016).

Memory and Statistical Learning

While recent research on the neuroscience of memory (e.g., Schapiro et al., 2016) suggests that CLS theory (McClelland et al., 1995) provides an overly simplified view of the memory system, the theory illustrates the important tension and interaction between representing specific instances (sparse coding) and generalizing from those instances (distributed coding). In the final section of this chapter, we will provide a brief mechanistic outline of how statistical learning emerges from this dynamic interaction between representation and generalization and highlight a set of questions that this perspective raises. To ground this discussion, we will first attempt to explain conditional statistical learning (as in a word segmentation task), before turning our attention to distributional statistical learning.

When presented with a stream of continuous input, as in the sequence of syllables that comprise an artificial language in a segmentation task (e.g., *golabubidakutupiro…*), learners extract a series of discrete representations (e.g., *go*, *labu*, *bidaku*, *tupi*, and *ro*). Initially, these representations are extracted randomly, as learners have no knowledge of the statistical structure of the input (e.g., Perruchet & Vinter, 1998). As the learner proceeds through the input, these representations are strengthened if the learner experiences them again and decay if the learner does not. Syllable groupings that are less likely to occur (i.e., syllable pairs that only occur across word boundaries) are less likely to be strengthened than are syllable groupings that occur frequently (i.e., the syllable groupings contained within a word). Over time, then, the representations that are most active in memory are those representations that correspond to frequent groupings of syllables seen in the input.

However, humans are sensitive not only to frequency; they are also able to distinguish high-probability sequences from low-probability sequences, even when those sequences occur equally often (e.g., Aslin et al., 1998). To capture this ability, we must incorporate interference, a well-known property of memory (e.g. Keppel & Underwood, 1962; Melton & Von Lackum, 1941). If two representations overlap with each other (e.g., they contain a syllable in common), then there may be interference between them, such that the strengthening of one representation results in the weakening of another. Because the foil items used in statistical learning tasks invariably involve a syllable drawn from multiple words, these items face interference from each of the words from which their syllables were drawn. Conversely, because words in these statistical learning tasks typically do not contain syllables that overlap with other words, these items do not face interference from other strong representations. This results in a set of representations where words are more strongly active than foil items (for discussion, see Perruchet & Vinter, 1998; Thiessen et al., 2013). Thus, memory processes can explain conditional probability learning (or learning the co-occurrence of items).

Similarly, memory processes are capable of explaining *distributional* statistical learning (or learning about the central tendency and variation of a set of items). To see how, we need to invoke one more critical component of memory: similarity-based activation. That is, when a cue is presented, the items in memory that are activated are those most similar to the cue (for discussion, see Hintzman, 1984). When multiple memory traces are activated, we propose (Thiessen & Pavlik, 2013) that information across the traces is integrated together – consistent with the principles of a distributed coding system – such that features which are consistent across the traces are enhanced while features that are inconsistent tend to cancel each other out; the resulting integrated information captures the central tendency of the integrated traces.

To see how this can explain distributional statistical learning, consider acquired distinctiveness, in which experience with two similar stimuli in distinct contexts makes them more discriminable. For example, children of around 18 months struggle to use phonemic contrasts (such as voicing in /d/ vs. /t/) in a label-object pairing task (Stager & Werker, 1997). But exposure to the contrast in distinct lexical contexts (e.g., /d/ and /t/ in *dawbow* and *tawgoo)* improves their performance, while exposure to the contrast in identical lexical contrasts (e.g., *dawgoo* and *tawgoo*) does not (Thiessen, 2007). We propose (Thiessen & Pavlik, 2013) that this is due to the effect of integrating information across multiple memory traces. When infants are exposed to /d/ and /t/ in identical lexical contexts, the resulting memory traces that are associated with both phonemes are quite similar. By contrast, when infants are exposed to /d/ and /t/ in distinct lexical contrasts, testing with /d/ mostly activates memories of one lexical context, while testing with /t/ mostly activates memories of a different lexical context. As such, the representations that arise from exposure to /d/ and /t/ are "pulled apart," becoming more distinct. This increased distinctiveness makes it easier for infants to take advantage of the contrast.

A memory-based approach to statistical learning also provides a unified framework for thinking about the relation between conditional and distributional statistical learning. These two processes are clearly related, as conditional statistical learning is influenced by the regularities across a set that learners have extracted (via distributional statistical learning) from the input. One piece of evidence for this is that learners from different language backgrounds extract different statistical clusters from the same input (e.g., Onnis & Thiessen, 2013). These effects can even be observed in a brief laboratory intervention. For example, after exposure to a set of bisyllabic words, learners are better at extracting bisyllabic words from statistical relations than at extracting trisyllabic words (Lew-Williams, Pelucchi, & Saffran, 2011). Similarly, after learning that words are likely to follow a specific lexical stress pattern (e.g., trochaic stress), learners are better able to identify words in fluent input that follow that stress pattern than words that violate it (Thiessen & Saffran, 2007).

Viewing both conditional and distributional learning within a memory-based framework provides a way to explain how these processes interact with each other. Once learners have extracted a set of words from the input, they can identify the regularities that characterize these words via distributional learning – that is, because these words are represented by overlapping neural populations, the features that are consistent across these words become strengthened. Once learners have identified a regularity that characterizes the words that they know, subsequently learning further words that conform to this regularity is facilitated. We propose that this is due to the effect of attention, which can influence the process of segmenting words from fluent input (e.g., Toro et al., 2005). Only elements of the input that are simultaneously held together in working memory can be bound together and extracted as a representation (e.g., Baker et al., 2004). Once a learner has discovered a phonological regularity that characterizes several familiar words, they begin to preferentially attend to syllable groupings that obey this regularity (e.g., Jusczyk, Cutler, & Redanz, 1993).

Moreover, viewing statistical learning as an emergent property of memory processes brings several questions and avenues of exploration into sharp relief. As we have discussed before, one of the primary questions related to statistical learning is about the nature of the underlying mechanisms. Our proposal is that these mechanisms are processes endemic to memory, such as activation, decay, and interference. If this proposal is correct, the characteristics of the memory system should be informative about the process and outcome of statistical learning. For example, differences in statistical learning tasks as a function of modality (e.g., Conway & Christiansen, 2005) may be explicable in terms of differences in the way that input is encoded into memory as a function of modality (e.g., Jensen, 1971; Penney, Gibbon, & Meck, 2000). Similarly, the developmental time course of statistical learning should be shaped by the maturation of the neurological systems underlying memory.

References

Abbot-Smith, K., & Tomasello, M. (2006). Exemplar-learning and schematization in a usage-based account of syntactic acquisition. *The Linguistic Review, 23*(3), 275–290.

Allen, J. S., & Miller, J. L. (1999). Effects of syllable-initial voicing and speaking rate on the temporal characteristics of monosyllabic words. *The Journal of the Acoustical Society of America, 106*(4), 2031–2039.

Allen, J. S., Miller, J. L., & DeSteno, D. (2003). Individual talker differences in voice-onset-time. *The Journal of the Acoustical Society of America, 113*(1), 544–552.

Aslin, R. N., & Newport, E. L. (2014). Distributional language learning: Mechanisms and models of category formation. *Language Learning, 64*(s2), 86–105.

Aslin, R. N., Saffran, J. R., & Newport, E. L. (1998). Computation of conditional probability statistics by 8-month-old infants. *Psychological Science, 9*(4), 321–324.

Baker, C., Olson, C. R., & Berhmann, M. (2004). Role of attention and perceptual grouping in visual statistical learning. *Psychological Science, 15*, 460–466.

Bakker, A., Kirwan, C. B., Miller, M., & Stark, C. E. (2008). Pattern separation in the human hippocampal CA3 and dentate gyrus. *Science, 319*(5870), 1640–1642.

Baldwin, D., Andersson, A., Saffran, J., & Meyer, M. (2008). Segmenting dynamic human action via statistical structure. *Cognition, 106*(3), 1382–1407.

Batterink, L. J. (2017). Rapid statistical learning supporting word extraction from continuous speech. *Psychological Science, 28*(7), 921–928.

Best, C. T., & Strange, W. (1992). Effects of phonological and phonetic factors on cross-language perception of approximants. *Journal of Phonetics, 20*(3), 305–330.

Biber, D. (1999). A register perspective on grammar and discourse: Variability in the form and use of English complement clauses. *Discourse Studies, 1*(2), 131–150.

Bornstein, A. M., & Daw, N. D. (2012). Dissociating hippocampal and striatal contributions to sequential prediction learning. *European Journal of Neuroscience, 35*(7), 1011–1023.

Chomsky, N. (1956). Three models for the description of language. *IRE Transactions on Information Theory, 2*(3), 113–124.

Clayards, M., Tanenhaus, M. K., Aslin, R. N., & Jacobs, R. A. (2008). Perception of speech reflects optimal use of probabilistic speech cues. *Cognition, 108*(3), 804–809.

Clelland, C. D., Choi, M., Romberg, C., Clemenson, G. D., Fragniere, A., Tyers, P., et al. (2009). A functional role for adult hippocampal neurogenesis in spatial pattern separation. *Science, 325*(5937), 210–213.

Conway, C. M., & Christiansen, M. H. (2005). Modality-constrained statistical learning of tactile, visual, and auditory sequences. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31*(1), 24–39.

Conway, C. M., & Christiansen, M. H. (2006). Statistical learning within and between modalities: Pitting abstract against stimulus-specific representations. *Psychological Science, 17*, 905–912.

Diehl, R. L., Lotto, A. J., & Holt, L. L. (2004). Speech perception. *Annual Review of Psychology, 55*, 149–179.

Dougherty, T. M., & Haith, M. M. (2002). Infants' use of constraints to speed information processing and to anticipate events. *Infancy, 3*(4), 457–473.

Englund, K., & Behne, D. (2006). Changes in infant directed speech in the first six months. *Infant and Child Development, 15*(2), 139–160.

Estes, K. G., Evans, J. L., Alibali, M. W., & Saffran, J. R. (2007). Can infants map meaning to newly segmented words? Statistical segmentation and word learning. *Psychological Science, 18*(3), 254–260.

Finn, A. S., Lee, T., & Hudson Kam, C. L. (2014). When it hurts (and helps) to try: The role of effort in language learning. *PLOS ONE, 9*(7), e101806.

Fiser, J., & Aslin, R. N. (2005). Encoding multielement scenes: Statistical learning of visual feature hierarchies. *Journal of Experimental Psychology: General, 134*(4), 521–537.

Frank, M. C., Goodman, N. D., & Tenenbaum, J. B. (2007). A Bayesian framework for cross-situational word-learning. In J. C. Platt, D. Koller, Y. Singer, & S. Roweis (Eds.), *Advances in neural information processing systems, volume 20* (pp. 1212–1222). Cambridge, MA: MIT Press.

Frost, R., Armstrong, B. C., Siegelman, N., & Christiansen, M. H. (2015). Domain generality versus modality specificity: The paradox of statistical learning. *Trends in Cognitive Sciences, 19*(3), 117–125.

Galantucci, B., Fowler, C. A., & Turvey, M. T. (2006). The motor theory of speech perception reviewed. *Psychonomic Bulletin & Review, 13*(3), 361–377.

Gerrits, E., & Schouten, M. E. H. (2004). Categorical perception depends on the discrimination task. *Perception & Psychophysics, 66*(3), 363–376.

Gilbert, P. E., Kesner, R. P., & Lee, I. (2001). Dissociating hippocampal subregions: A double dissociation between dentate gyrus and CA1. *Hippocampus, 11*(6), 626–636.

Giroux, I., & Rey, A. (2009). Lexical and sublexical units in speech perception. *Cognitive Science, 33*(2), 260–272.

Goldinger, S. D. (1998). Echoes of echoes? An episodic theory of lexical access. *Psychological Review, 105*(2), 251–279.

Gómez, R. L., & Gerken, L. (2000). Infant artificial language learning and language acquisition. *Trends in Cognitive Sciences, 4*(5), 178–186.

Hardcastle, W. J., & Hewlett, N. (Eds.). (2006). *Coarticulation: Theory, data and techniques*. Cambridge, UK: Cambridge University Press.

Hayes, J. R., & Clark, H. H. (1970). Experiments in the segmentation of an artificial speech analog. In J. R. Hayes (Ed.), *Cognition and the development of language* (pp. 221–234). New York: Wiley.

Hinton, G. E., McClelland, J. L., & Rumelhart, D. E. (1986). Distributed representations. *Parallel Distributed Processing: Explorations in the Microstructure of Cognition, 1*(3), 77–109.

Hintzman, D. L. (1984). MINERVA 2: A simulation model of human memory. *Behavior Research Methods, Instruments, & Computers, 16*(2), 96–101.

Honey, R. C., & Hall, G. (1989). Acquired equivalence and distinctiveness of cues. *Journal of Experimental Psychology: Animal Behavior Processes, 15*(4), 338–346.

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 and Performance, 29*(6), 1143–1154.

Hunt, R. H., & Aslin, R. N. (2001). Statistical learning in a serial reaction time task: Access to separable statistical cues by individual learners. *Journal of Experimental Psychology: General, 130*, 658–680.

Iskarous, K., & Kavitskaya, D. (2010). The interaction between contrast, prosody, and coarticulation in structuring phonetic variability. *Journal of Phonetics, 38*(4), 625–639.

James, W. (1890). *The principles of psychology*. New York: H. Holt and Company.

Jensen, A. R., (1971). Individual differences in visual and auditory memory. *Journal of Educational Psychology, 62*, 123–131.

Johnson, E. K., & Jusczyk, P. W. (2001). Word segmentation by 8-month-olds: When speech cues count more than statistics. *Journal of Memory and Language, 44*(4), 548–567.

Johnson, E. K., & Seidl, A. (2008). Clause segmentation by 6-month-old infants: A crosslinguistic perspective. *Infancy, 13*(5), 440–455.

Johnson, E. K., & Tyler, M. D. (2010). Testing the limits of statistical learning for word segmentation. *Developmental Science, 13*(2), 339–345.

Jusczyk, P. W., Cutler, A., & Redanz, N. J. (1993). Infants' preference for the predominant stress patterns of English words. *Child Development, 64*, 675–687.

Keppel, G., & Underwood, B. J. (1962). Proactive inhibition in short-term retention of single items. *Journal of Verbal Learning and Verbal Behavior, 1*(3), 153–161.

Kirkham, N. Z., Slemmer, J. A., & Johnson, S. P. (2002). Visual statistical learning in infancy: Evidence for a domain general learning mechanism. *Cognition, 83*(2), B35–B42.

Kruschke, J. K. (2005). Category learning. In K. Lamberts & R. L. Goldstone (Eds.), *The handbook of cognition* (pp. 183–201). London, UK: Sage.
Langlois, J. H., & Roggman, L. A. (1990). Attractive faces are only average. *Psychological Science, 1*, 115–121.
Leutgeb, J. K., Leutgeb, S., Moser, M. B., & Moser, E. I. (2007). Pattern separation in the dentate gyrus and CA3 of the hippocampus. *Science, 315*(5814), 961–966.
Lew-Williams, C., Pelucchi, B., & Saffran, J. R. (2011). Isolated words enhance statistical language learning in infancy. *Developmental Science, 14*, 1323–1329.
Liberman, A. M., Cooper, F. S., Shankweiler, D. P., & Studdert-Kennedy, M. (1967). Perception of the speech code. *Psychological Review, 74*(6), 431–461.
Liberman, A. M., Harris, K. S., Hoffman, H. S., & Griffith, B. C. (1957). The discrimination of speech sounds within and across phoneme boundaries. *Journal of Experimental Psychology, 54*(5), 358–368.
Lively, S. E., Logan, J. S., & Pisoni, D. B. (1993). Training Japanese listeners to identify English/r/ and/l/. II: The role of phonetic environment and talker variability in learning new perceptual categories. *The Journal of the Acoustical Society of America, 94*(3), 1242–1255.
Lotto, A. J., & Kluender, K. R. (1998). General contrast effects in speech perception: Effect of preceding liquid on stop consonant identification. *Perception & Psychophysics, 60*(4), 602–619.
Lotto, A. J., Kluender, K. R., & Holt, L. L. (1997). Perceptual compensation for coarticulation by Japanese quail (Coturnix coturnix japonica). *The Journal of the Acoustical Society of America, 102*(2), 1134–1140.
Maye, J., Weiss, D. J., & Aslin, R. N. (2008). Statistical phonetic learning in infants: Facilitation and feature generalization. *Developmental Science, 11*(1), 122–134.
Maye, J., Werker, J. F., & Gerken, L. (2002). Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition, 82*(3), B101–B111.
McClelland, J. L., McNaughton, B. L., & O'Reilly, R. C. (1995). Why there are complementary learning systems in the hippocampus and neocortex: Insights from the successes and failures of connectionist models of learning and memory. *Psychological Review, 102*(3), 419–457.
McClelland, J. L., & Rumelhart, D. E. (1985). Distributed memory and the representation of general and specific information. *Journal of Experimental Psychology: General, 114*(2), 159–188.
McHugh, T. J., Jones, M. W., Quinn, J. J., Balthasar, N., Coppari, R., Elmquist, J. K., et al. (2007). Dentate gyrus NMDA receptors mediate rapid pattern separation in the hippocampal network. *Science, 317*(5834), 94–99.
McMurray, B., Tanenhaus, M. K., & Aslin, R. N. (2002). Gradient effects of within-category phonetic variation on lexical access. *Cognition, 86*(2), B33–B42.
Medin, D. L., & Schaffer, M. M. (1978). Context theory of classification learning. *Psychological Review, 85*(3), 207–238.
Melton, A. W., & Von Lackum, W. J. (1941). Retroactive and proactive inhibition in retention: Evidence for a two-factor theory of retroactive inhibition. *The American Journal of Psychology, 54*(2), 157–173.
Miller, J. L., & Volaitis, L. E. (1989). Effect of speaking rate on the perceptual structure of a phonetic category. *Perception & Psychophysics, 46*(6), 505–512.
Minda, J. P., & Smith, J. D. (2002). Comparing prototype-based and exemplar-based accounts of category learning and attentional allocation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28*(2), 275–292.
Mintz, T. H. (2003). Frequent frames as a cue for grammatical categories in child directed speech. *Cognition, 90*(1), 91–117.
Mirman, D., Magnuson, J. S., Estes, K. G., & Dixon, J. A. (2008). The link between statistical segmentation and word learning in adults. *Cognition, 108*(1), 271–280.
Misyak, J. B., & Christiansen, M. H. (2012). Statistical learning and language: An individual differences study. *Language Learning, 62*(1), 302–331.
Misyak, J. B., Christiansen, M. H., & Tomblin, J. B. (2010). On-line individual differences in statistical learning predict language processing. *Frontiers in Psychology, 1*, 31.

Norman, K. A., & O'Reilly, R. C. (2003). Modeling hippocampal and neocortical contributions to recognition memory: a complementary-learning-systems approach. *Psychological Review, 110*, 611–646.
O'Reilly, R. C., Bhattacharyya, R., Howard, M. D., & Ketz, N. (2014). Complementary learning systems. *Cognitive Science, 38*(6), 1229–1248.
Oakes, L. M., & Spalding, T. L. (1997). The role of exemplar distribution in infants' differentiation of categories. *Infant Behavior and Development, 20*(4), 457–475.
Onnis, L., & Thiessen, E. (2013). Language experience changes subsequent learning. *Cognition, 126*(2), 268–284.
Orbán, G., Fiser, J., Aslin, R. N., & Lengyel, M. (2008). Bayesian learning of visual chunks by human observers. *Proceedings of the National Academy of Sciences, 105*(7), 2745–2750.
Paap, K. R., Newsome, S. L., McDonald, J. E., & Schvaneveldt, R. W. (1982). An activation–verification model for letter and word recognition: The word-superiority effect. *Psychological Review, 89*(5), 573–594.
Penney, T., Gibbon, J., & Meck, W. (2000). Differential effects of auditory and visual signals on clock speed and temporal memory. *Journal of Experimental Psychology: Human Perception and Performance, 26*, 1770–1787.
Perkell, J. S., & Matthies, M. L. (1992). Temporal measures of anticipatory labial coarticulation for the vowel/u: Within-and cross-subject variability. *The Journal of the Acoustical Society of America, 91*(5), 2911–2925.
Perruchet, P., & Pacton, S. (2006). Implicit learning and statistical learning: One phenomenon, two approaches. *Trends in Cognitive Sciences, 10*(5), 233–238.
Perruchet, P., & Vinter, A. (1998). PARSER: A model for word segmentation. *Journal of Memory and Language, 39*(2), 246–263.
Perry, L. K., Samuelson, L. K., Malloy, L. M., & Schiffer, R. N. (2010). Learn locally, think globally: Exemplar variability supports higher-order generalization and word learning. *Psychological Science, 21*(12), 1894–1902.
Pisoni, D. B., & Tash, J. (1974). Reaction times to comparisons within and across phonetic categories. *Perception & Psychophysics, 15*(2), 285–290.
Pollack, I., & Pickett, J. M. (1964). Intelligibility of excerpts from fluent speech: Auditory vs. structural context. *Journal of Verbal Learning and Verbal Behavior, 3*(1), 79–84.
Posner, M. I., & Keele, S. W. (1968). On the genesis of abstract ideas. *Journal of Experimental Psychology, 77*(3p1), 353–363.
Principe, C. P., & Langlois, J. H. (2012). Shifting the prototype: Experience with facesinfluences affective and attractiveness preferences. *Social Cognition, 30*, 109–120.
Quinn, P. C., Eimas, P. D., & Rosenkrantz, S. L. (1993). Evidence for representations of perceptually similar natural categories by 3-month-old and 4-month-old infants. *Perception, 22*, 463–475.
Reber, A. S., & Lewis, S. (1977). Implicit learning: An analysis of the form and structure of a body of tacit knowledge. *Cognition, 5*(4), 333–361.
Rogers, T. T., & McClelland, J. L. (2004). *Semantic cognition: A parallel distributed processing approach*. Cambridge, MA: MIT Press.
Romberg, A. R., & Saffran, J. R. (2010). Statistical learning and language acquisition. *Cognitive Science, 1*(6), 906–914.
Rosch, E. (1975). Cognitive representations of semantic categories. *Journal of Experimental Psychology: General, 104*(3), 192–233.
Rost, G. C., & McMurray, B. (2009). Speaker variability augments phonological processing in early word learning. *Developmental Science, 12*(2), 339–349.
Saffran, J. R. (2001). Words in a sea of sounds: The output of infant statistical learning. *Cognition, 81*(2), 149–169.
Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science, 274*(5294), 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*(1), 27–52.

Saffran, J. R., Pollak, S. D., Seibel, R. L., & Shkolnik, A. (2007). Dog is a dog is a dog: Infant rule learning is not specific to language. *Cognition, 105*(3), 669–680.
Schapiro, A. C., Gregory, E., Landau, B., McCloskey, M., & Turk-Browne, N. B. (2014). The necessity of the medial temporal lobe for statistical learning. *Journal of Cognitive Neuroscience, 26*(8), 1736–1747.
Schapiro, A. C., Kustner, L. V., & Turk-Browne, N. B. (2012). Shaping of object representations in the human medial temporal lobe based on temporal regularities. *Current Biology, 22*(17), 1622–1627.
Schapiro, A. C., Rogers, T. T., Cordova, N. I., Turk-Browne, N. B., & Botvinick, M. M. (2013). Neural representations of events arise from temporal community structure. *Nature Neuroscience, 16*(4), 486–492.
Schapiro, A. C., Turk-Browne, N. B., Norman, K. A., & Botvinick, M. M. (2016). Statistical learning of temporal community structure in the hippocampus. *Hippocampus, 26*(1), 3–8.
Slone, L. K., & Johnson, S. P. (2015). Infants' statistical learning: 2-and 5-month-olds' segmentation of continuous visual sequences. *Journal of Experimental Child Psychology, 133*, 47–56.
Smith, E. R. (2009). Distributed connectionist models in social psychology. *Social and Personality Psychology Compass, 3*(1), 64–76.
Smith, J. D., & Minda, J. P. (1998). Prototypes in the mist: The early epochs of category learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24*(6), 1411–1436.
Smith, L., & Yu, C. (2008). Infants rapidly learn word-referent mappings via cross-situational statistics. *Cognition, 106*(3), 1558–1568.
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.
Stern, D. N., Spieker, S., Barnett, R. K., & MacKain, K. (1983). The prosody of maternal speech: Infant age and context related changes. *Journal of Child Language, 10*(1), 1–15.
Swingley, D. (1999). Conditional probability and word discovery: A corpus analysis of speech to infants. In M. Hahn & S. C. Stoness (Eds.), *Proceedings of the 21st annual conference of the cognitive science society* (pp. 724–729). Mahwah, NJ: LEA.
Taylor, P., & Black, A. W. (1998). Assigning phrase breaks from part-of-speech sequences. *Computer Speech & Language, 12*(2), 99–117.
Thiessen, E. D. (2007). The effect of distributional information on children's use of phonemic contrasts. *Journal of Memory and Language, 56*(1), 16–34.
Thiessen, E. D., Kronstein, A. T., & Hufnagle, D. G. (2013). The extraction and integration framework: A two-process account of statistical learning. *Psychological Bulletin, 139*(4), 792–814.
Thiessen, E. D., & Pavlik, P. I. (2013). iMinerva: A mathematical model of distributional statistical learning. *Cognitive Science, 37*(2), 310–343.
Thiessen, E. D., & Pavlik, P. I. (2016). Modeling the role of distributional information in children's use of phonemic contrasts. *Journal of Memory and Language, 88*, 117–132.
Thiessen, E. D., & Saffran, J. R. (2004). Spectral tilt as a cue to word segmentation in infancy and adulthood. *Perception & Psychophysics, 66*(5), 779–791.
Thiessen, E. D., & Saffran, J. R. (2007). Learning to learn: Infants' acquisition of stress-based strategies for word segmentation. *Language Learning and Development, 3*(1), 73–100.
Thiessen, E. D., & Yee, M. N. (2010). Dogs, bogs, labs, and lads: What phonemic generalizations indicate about the nature of Children's early word-form representations. *Child Development, 81*(4), 1287–1303.
Thompson, S. P., & Newport, E. L. (2007). Statistical learning of syntax: The role of transitional probability. *Language Learning and Development, 3*(1), 1–42.
Tomasello, M. (2000). Do young children have adult syntactic competence? *Cognition, 74*, 209–253.
Toro, J. M., Sinnett, S., & Soto-Faraco, S. (2005). Speech segmentation by statistical learning depends on attention. *Cognition, 97*, B25–B34.
Turk-Browne, N. B., Scholl, B. J., Chun, M. M., & Johnson, M. K. (2009). Neural evidence of statistical learning: Efficient detection of visual regularities without awareness. *Journal of Cognitive Neuroscience, 21*(10), 1934–1945.

Turk-Browne, N. B., Scholl, B. J., Johnson, M. K., & Chun, M. M. (2010). Implicit perceptual anticipation triggered by statistical learning. *Journal of Neuroscience, 30*(33), 11177–11187.

van den Bos, E., Christiansen, M. H., & Misyak, J. B. (2012). Statistical learning of probabilistic nonadjacent dependencies by multiple-cue integration. *Journal of Memory and Language, 67*(4), 507–520.

Vanpaemel, W. (2016). Prototypes, exemplars and the response scaling parameter: A Bayes factor perspective. *Journal of Mathematical Psychology, 72*, 183–190.

Vitevitch, M. S., & Luce, P. A. (2004). A web-based interface to calculate phonotactic probability for words and nonwords in English. *Behavior Research Methods, Instruments, & Computers, 36*(3), 481–487.

Werker, J. F., Pons, F., Dietrich, C., Kajikawa, S., Fais, L., & Amano, S. (2007). Infant-directed speech supports phonetic category learning in English and Japanese. *Cognition, 103*(1), 147–162.

Werker, J. F., & Tees, R. C. (1984). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development, 7*(1), 49–63.

Winocur, G., Moscovitch, M., & Bontempi, B. (2010). Memory formation and long-term retention in humans and animals: Convergence towards a transformation account of hippocampal–neocortical interactions. *Neuropsychologia, 48*(8), 2339–2356.

Zaki, S. R., Nosofsky, R. M., Stanton, R. D., & Cohen, A. L. (2003). Prototype and exemplar accounts of category learning and attentional allocation: A reassessment. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29*(6), 1160–1173.

Zhao, J., Ngo, N., McKendrick, R., & Turk-Browne, N. B. (2011). Mutual interference between statistical summary perception and statistical learning. *Psychological Science, 22*(9), 1212–1219.

Chapter 5
Structure-Mapping Processes Enable Infants' Learning Across Domains Including Language

Susan J. Hespos, Erin Anderson, and Dedre Gentner

Abstract Humans have an astounding ability to acquire new information. Like many other animals, we can learn by association and by perceptual generalization. However, unlike most other species, we also acquire new information by means of relational generalization and transfer. In this chapter, we explore the origins of a uniquely developed human capacity—our ability to learn relational abstractions through analogical comparison. We focus on whether and how infants can use analogical comparison to derive relational abstractions from examples. We frame our work in terms of structure-mapping theory, which has been fruitfully applied to analogical processing in children and adults. We find that young infants show two key signatures of structure mapping: first, relational abstraction is fostered by comparing alignable examples, and second, relational abstraction is hampered by the presence of highly salient objects. The studies we review make it clear that structure-mapping processes are evident in the first months of life, prior to much influence of language and culture. This finding suggests that infants are born with analogical processing mechanisms that allow them to learn relations through comparing examples.

Turning to very early learning, we augmented our account by considering the nature of young infants' encoding processes, leading to two counterintuitive predictions. First, we predicted that young infants (2–3 months old) would be better able to form a relational abstraction when given two alternating exemplars than when given six different exemplars (Anderson et al. Cognition 176:74–86, 2018). This is based on the assumption that young infants may initially focus on the individual objects and shift to noticing the relation between them after repetition of the

This research was supported by National Science Foundation grants BCS-1423917 and BCS-1729720 awarded to Susan Hespos and Dedre Gentner, NSF SLC grant SBE-1041707 awarded to the Spatial Intelligence and Learning Center (SILC), ONR grant N00014-92-J-1098 awarded to Dedre Gentner, and US Department of Education's Institute of Education Sciences training grant (Multidisciplinary Program in Education Sciences) no. R305B140042 awarded to Erin Anderson.

S. J. Hespos (✉) · E. Anderson · D. Gentner
Northwestern University, Evanston, IL, USA
e-mail: hespos@northwestern.edu

J. B. Childers (ed.), *Language and Concept Acquisition from Infancy Through Childhood*, https://doi.org/10.1007/978-3-030-35594-4_5

exemplar (Casasola. Child Development 76(1):279–290, 2005a; Casasola. Developmental Psychology 41:183–192, 2005b). Second, we predicted that younger, but not older, infants would be able to form a relational abstraction from one repeated exemplar; this follows from the assumption that young infants have unstable encoding processes, so identical exemplars may be variably encoded (Anderson et al. 2019).

Next, we revisited Premack's insight from 1983 that the tasks used to measure analogical abilities (RMTS, MTS, and *same/different* discrimination) are vastly different from each other. The takeaway from this section is that while many species can learn through association and perceptual generalizations, there are relatively few species that can succeed in the *same/different* discrimination task. Of these species that can succeed in the *same/different* task, humans are unique in that they need fewer than 10 trials to learn such relations. In the final sections, we reviewed how structure mapping extends to language acquisition, artificial grammar learning, and physical reasoning. The value of investigating the origins of our analogical abilities is that we will be in a better position to understand how language and culture capitalize on cognitive abilities. More broadly, we can address whether essential differences between humans and other species are evident from the earliest points in development.

Humans have an astounding ability to acquire new information. Like many other animals, we learn by association and by perceptual generalization. However, unlike most other species, we also acquire new information by means of relational generalization and transfer. In this chapter, we will explore the origins of a uniquely developed human capacity—our ability to learn relational abstractions through analogical comparison. We focus on whether and how infants can use analogical comparison processes to derive relational abstractions from examples.

By analogical comparison, we mean a comparison process in which the relational structure of the two items is aligned, as described in Gentner's structure-mapping theory (Gentner, 1983, 1989, 2010; Markman & Gentner, 1997). At first glance, the idea that infants might use analogical processes may seem absurd. After all, analogy is considered a sophisticated process even in adults. Further, there is a methodological challenge in studying whether prelinguistic infants can make analogical comparisons. Fortunately, decades of research have revealed general signatures of relational alignment and learning; thus, we can compare the performance of infants with established signatures of analogical processing.

The value of this pursuit is in allowing us to discover the roots of relational cognition. Adults' ability to use abstract categories and rules is supported by a vast store of conceptual knowledge, influenced by the culture that surrounds us and the languages we speak, as well as by real-world experience. To gain an understanding of the nature and origin of our extraordinary relational ability, we must investigate infants who have had less exposure to culture and language. If we can specify how infants learn relations from multiple examples, then we will be in a better position to understand how language and culture capitalize on these existing cognitive abilities and how our cognitive processes compare to those of other species.

Our field is only beginning to study the origins of human analogical ability. However, there has been considerable research on the development of analogical ability from preschool to adulthood. We briefly review this research to set the stage for examining what characteristics might be evident in infants. In general, children's comparison processing shows a relational shift whereby children focus on object matches early in learning and focus increasingly on relational commonalities as they gain in domain knowledge (Gentner, 1988; Gentner, Anggoro, & Klibanoff, 2011; Gentner & Rattermann, 1991; Gentner & Toupin, 1986; Paik & Mix, 2006, 2008; Richland, Morrison, & Holyoak, 2006). For example, Gentner and Rattermann (1991, Experiment 1) asked 3- to 5-year-old children to find a hidden sticker. The experimenter had three pots of increasing sizes in a row in front of them. The child had a similar series of three pots in front of them. On each trial, the experimenter secretly hid an object under one of the child's pots. Then, while the child watched, the experimenter placed a sticker under one of her pots. The stickers were always placed in the same relative position—left (smallest), middle (medium), or right (largest)—and the child was told that by watching where the experimenter put a sticker, they could find their sticker. Three-year-olds succeeded when identical objects occupied the same relational roles. The interesting manipulation was when the sizes of the pots were shifted, such that the experimenter had a small, medium, and large pot and the child had a medium, large, and extra-large pot. This arrangement sets up a cross-mapping—a case in which there is an object match that competes with the desired relational match (Gentner & Toupin, 1986; Ross, 1989). Cross-mapped analogies provide a stringent test of children's understanding of the relational match. In Gentner and Rattermann's study, younger (3-year-olds) children performed at chance, repeatedly choosing the object match, while older children (5-year-olds) chose the relational match.

Gentner and colleagues have argued that the relational shift is not age-linked, but rather results primarily from increases in relational knowledge (see also Gentner, 1989, 2003, 2010; Gentner & Medina, 1998; Gentner & Rattermann, 1991). As evidence for this claim, Gentner and Rattermann (1991, Table 7.5, p. 250) offered examples of the relational shift taking place between 4 and 6 months in an occlusion event (Baillargeon, 1994), between 3 and 4 years old (the Gentner and Rattermann task described above), and between 6 and 9 years old in a story-enacting task involving social causation (Gentner & Toupin, 1986). As further evidence that the shift is largely driven by knowledge, rather than maturational processes, Gentner and Rattermann showed that 3-year-olds could succeed on this task when provided with relational labels for the object sets (e.g., daddy, mommy, baby). This suggests that the 3-year-olds in the initial study were limited not by age-related processing constraints, but by the lack of a relational knowledge schema in this task (see also Loewenstein & Gentner, 2005). Other researchers have linked the relational shift to maturational increases in processing capacity (Halford, 1992) and to increases in executive ability, including inhibitory control (Doumas, Hummel, & Sandhofer, 2008; Richland et al., 2006; Thibaut, French, & Vezneva, 2010), and it is possible that all three factors play a role.

This work has also revealed characteristic patterns of analogical learning, including factors that facilitate and hinder the relational learning process. One signature component of relational learning is that the ability to perceive abstract relational matches can be enhanced by comparing instances of a relation. For example, Gick and Holyoak (1983) found that comparing two stories that had the same abstract causal structure enabled people to generalize that structure and to transfer it to a further situation. Similar effects of comparison have been found for preschool children in relational tasks (e.g., Christie & Gentner, 2010; Gentner et al., 2011; Gentner & Namy, 1999; Namy & Clepper, 2010). These findings are consistent with the structure-mapping theory (Gentner, 1983; Gentner & Forbus, 2011; Gentner & Markman, 1997) account that the act of comparison entails a structural alignment process. In structural alignment, the two analogs are aligned in such a way that the common relations are placed into correspondence. Once a structural alignment is achieved, the relational commonalities between the items are highlighted (Markman & Gentner, 1993; see also Gentner, 2010). In addition, further inferences may be projected, and certain differences may be highlighted; however, in this chapter, we focus on the role of structural alignment in revealing commonalities. The influence of structural alignment is a defining characteristic of analogical reasoning in adults (Doumas & Hummel, 2013; Forbus, Ferguson, Lovett, & Gentner, 2017; Gentner, Holyoak, & Kokinov, 2001) and the evidence of its influence in children as young as 3 years of age suggests a possible continuity in relational processing through human development. In this chapter, we explore whether this continuity extends to infants.

The Gentner and Rattermann study also exemplifies a second signature of analogical processing: namely, that attention to individual objects can interfere with relational processing. The 3-year-olds in these studies were able to carry out the mapping quite well when the objects matched but failed when there were competing object matches (unless given support from relational language). There are many studies showing that preschool children perform far worse on relational matching tasks when competing object matches are present (Gentner & Toupin, 1986; Richland et al., 2006), especially if the objects involved are rich and distinctive (DeLoache, 1995; Gentner & Rattermann, 1991; Paik & Mix, 2006). For example, children can pass the relational match-to-sample (RMTS) task (exemplified on the left in Fig. 5.1) at 4.5 (Christie & Gentner, 2014) or 5 years of age (Hochmann, Mody, & Carey, 2016). However, when Christie and Gentner (2007) gave children and adults a version of the RMTS task in which there was a competing object match (see the right side of Fig. 5.1), the results showed a steep gradient across age: 4.5-year-olds chose the relational match only 17% of the time, 8.5-year-olds performed at chance (50%), and adults chose the relational match 90% of the time.

The finding that attention to objects can overshadow attention to relations has also been found in word-learning tasks (Casasola, 2005a; Maguire, Hirsh-Pasek, Golinkoff, & Brandone, 2008). In the work we will describe below, we focus on infant relational learning and ask whether it is similarly facilitated by comparison and hindered by object focus. Finding substantive evidence for the signatures of analogical reasoning in infants would suggest that the relational process is continuous through development.

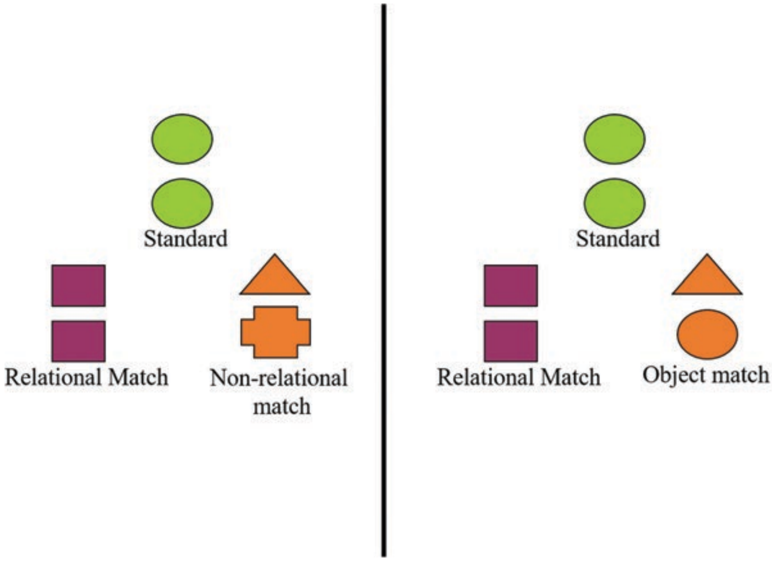


Fig. 5.1 On the left is a sample triad from Christie and Gentner (2014). They found 4.5-year-olds chose the relational match significantly more often than chance. On the right is a sample trial that contains a competing object match from Christie and Gentner (2007), and the pattern of results was very different. In trials with a competing object match (the green and orange circles), 4.5-year-olds preferred the object match, adults chose the relational match, and 8.5-year-olds were in the middle; their performance was not different from chance

The existing literature on the development of analogical abilities highlights the role of linguistic symbols in facilitating relational learning (Gentner, 2003, 2010; Gentner & Rattermann, 1991). There is evidence that children's relational insight is improved by having symbolic labels for relations and relational systems (Carey, 2010; Christie & Gentner, 2014; Gentner, 2005; Hermer & Spelke, 1994; Loewenstein & Gentner, 2005; Pyers, Shusterman, Senghas, Spelke, & Emmorey, 2010; Son, Doumas, & Goldstone, 2010). There is also considerable evidence that common labels can prompt children (and adults; see Lupyan, 2012) to compare referents and abstract the commonalities they share, for concrete nouns (Ferry, Hespos, & Waxman, 2010; Gentner & Namy, 1999; Liu, Golinkoff, & Sak, 2001; Namy & Gentner, 2002), relational nouns (Gentner, 2005; Gentner et al., 2011), adjectives (Waxman & Klibanoff, 2000), and verbs (Haryu, Imai, & Okada, 2011; Waxman et al., 2013; see Gentner & Namy, 2006, for a review). More specifically, there is evidence that relational language, such as verbs, prepositions, and comparative adjectives, can foster retaining and transferring relational patterns (Casasola, 2005b; Childers, 2011; Christie & Gentner, 2014; Gentner et al., 2011; Hermer & Spelke, 1994; Jamrozik & Gentner, under review; Loewenstein & Gentner, 2005; Pyers & Senghas, 2007; Son et al., 2010). While it is clear that language plays a critical role in relational learning and reasoning in children and adults, the focus of this chapter will be on infants' prelinguistic abilities, prior to much influence from language and culture. If we see evidence of relational learning in early infancy, then we can infer that these processes exist prior to the acquisition of language. Moreover, such findings would put us in a better position to understand how language learning may capitalize on this preexisting relational ability.

The central question in this chapter is how human analogical ability arises. More specifically, when and how does our ability to derive relational abstractions from examples arise? One possibility is that we are born with a core set of abstract relations, which we can perceive in specific examples. Such a set would almost certainly include the relations *same* or *different* (Christie & Gentner, 2014; Hochmann et al., 2017; Wasserman, Castro, & Fagot, 2017). A second possibility is that infants are born with an analogical processing mechanism that allows them to learn relations through comparing examples. A third possibility is that analogical ability develops by combining other abilities through cultural and linguistic experience. To decide among these proposals, we focus on the *same-different* relation. The relations of *same* and *different* are among the simplest and most basic relations in the human repertoire and are therefore a logical starting point.

Possibility (1)—that we are born with a core set of abstract relations—has been widely assumed, based on a highly cited study by Tyrrell, Stauffer, and Snowman (1991). Tyrrell et al. (1991), using a preferential looking paradigm, reported that 7-month-old infants encode abstract *same* and *different* relations without training, simply from exposure to a single exemplar. However, examination of the reported results revealed ambiguity as to whether infants genuinely abstracted the relation. We therefore replicated Tyrrell et al.'s methods with the same age group (Ferry, Hespos, & Gentner, 2015). The results showed no evidence for relational abstraction. Infants showed a novelty response when comparing the identical pairs they had seen (e.g., AA) with a new pair (BC), but when the familiarized relation and the competing relation were tested with new objects (e.g., XX vs. YZ), the infants showed no preference. Thus, there is no evidence that these infants formed a relational abstraction from one exposure.

Next, we tested the second proposal: whether infants are capable of learning an abstract relation by structural alignment across exemplars. We showed infants a sequence of four exemplars of *same* or *different* toys. Half the infants saw *same* pairs (e.g., AA, BB, CC, DD), and half saw *different* pairs (AB, CD, BC, DA), repeated until infant looking declined sufficiently to demonstrate habituation (about 6–9 pairs). We then showed infants a sequence of six test trials. On alternating trials, infants saw pairs of objects that were either the *same* or *different*, and the dependent measure was the duration of infants' looking times. The key question was whether infants would look longer at the novel relation (AA vs. AB), even when instantiated with new objects (XX vs. YZ). Indeed, that is what happened, both for infants habituated to *same* and for those habituated to *different*—evidence that they had abstracted the common relation across the habituation pairs (see Fig. 5.2).

This ability to learn an abstract relation from a series of examples is one signature of analogical learning in older children and adults. We also tested the second signature of relational learning—whether object salience would interfere with structural alignment. Prior to the experiment, we gave infants a brief exposure to a

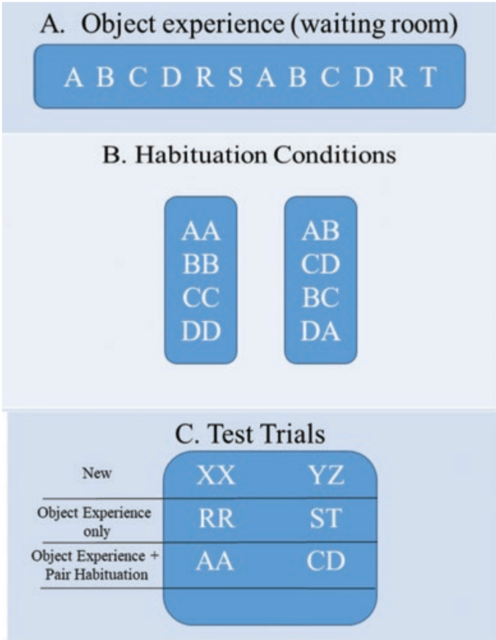


Fig. 5.2 Schematic of events in Ferry et al. (2015). (**a**) In the waiting room, infants saw a subset of the individual toys before the experiment. (**b**) Infants were habituated to four pairs of objects, either *same* or *different*. (**c**) In six sequential test trials, looking time was recorded to the novel and familiar relational pairs in three different types of test trials

subset of the objects used in test trials, thus increasing the salience of these individual objects. We found that infants failed to discriminate between the *same* and *different* relations when the test pairs contained objects that had been rendered individually salient prior to habituation—consistent with the findings among older children, for whom object salience interferes with analogical comparison (Gentner & Toupin, 1986; Paik & Mix, 2006). These findings suggest that by 7 months, infants show the basic characteristics of analogical learning—their learning was facilitated by comparison across examples and hindered by object focus. We interpret these findings as showing that the analogical processing ability is present in the first year of life and may be continuous through development.

Given our non-replication of the Tyrrell et al. study, we cannot assume that infants have a preexisting relational vocabulary that they can apply to examples in the world. Rather, our studies provide evidence that infants have a relational processing mechanism that can compare across examples to form abstract relations. These findings also argue against the third possibility that analogical ability arises through combining other capacities and experiences. Although language and conceptual learning refine and extend our analogical abilities, these abilities are present before extensive cultural and linguistic experience.

Our next study tested for relational abstraction at the earliest age possible to serve as a base for capturing developmental changes and variability in the learning process across age groups. Anderson, Chang, Hespos, and Gentner (2018) tested 3-month-old infants—the earliest age at which infants have the neck control to participate in a looking-time paradigm. As in the prior study, the key dependent measure is whether infants are able to differentiate the familiar relation (e.g., *same*, if habituated to *same*) from the unfamiliar one (e.g., *different*) when they see test pairs composed of new objects. The specific predictions were that if infants are learning by comparison, then (1) relational learning should benefit from comparing a series of analogous exemplars and (2) performance on test pairs should be hampered for pairs that contain objects that were rendered individually salient through object experience in the waiting room prior to the experiment.

Learning theories broadly agree that increasing the variability in a set of exemplars should lead to a greater range of transfer (Markman & Wisniewski, 1997; Rogers & McClelland, 2005; Wasserman, Young, & Fagot, 2001; Xu & Tenenbaum, 2007). Following this logic, young infants may require a larger training set than the four exemplars given to older infants in Ferry et al.'s (2015) study. Therefore, in one study, we increased the number of exemplars seen in habituation to six.

But there is an alternate possibility. Because alignment of relational structure is the sine qua non for discovering new relational commonalities, the ability to successfully compare and align is a prerequisite for relational learning. As discussed below, some studies have found that increasing the number and variability of examples can be detrimental to young children's relational learning (Casasola, 2005a; Maguire et al., 2008). To allow for this possibility, in our second experiment, we gave infants two exemplars that alternated across habituation (see Fig. 5.3).

The results revealed no evidence of learning the relation when 3-month-old infants were presented with six exemplars. However, the infants did learn the relation when they were presented with two alternating exemplars during habituation trials. In the two-exemplar condition, the 3-month-olds showed the key signature of analogical abstraction: they looked significantly longer at the novel relation during test even when that relation was instantiated with new objects, thus suggesting that they were able to transfer the relation to objects that they had not seen previously. In addition, there was evidence that object focus hindered learning. As in our prior studies, there was no difference in looking time between the novel and familiar relations when instantiated by objects that had been made individually salient through pre-exposure. Further, there was a significant difference in performance across test trial types that contrasted pairs seen in the waiting room before the experiment and new objects. These findings show that the signatures of analogical learning are present not only at 7 months (Ferry et al., 2015) but also by 3 months of age (Anderson et al., 2018). Clearly, language is not a necessary prerequisite for relational processing—the ability to carry out structural alignment and abstraction is in place prior to and independent of language. In contrast to the possibility that relational knowledge depends on language, we speculate that language may capitalize on the relational processes and may be used in learning grammatical structures (Gentner, 2010; Gentner & Namy, 2006).

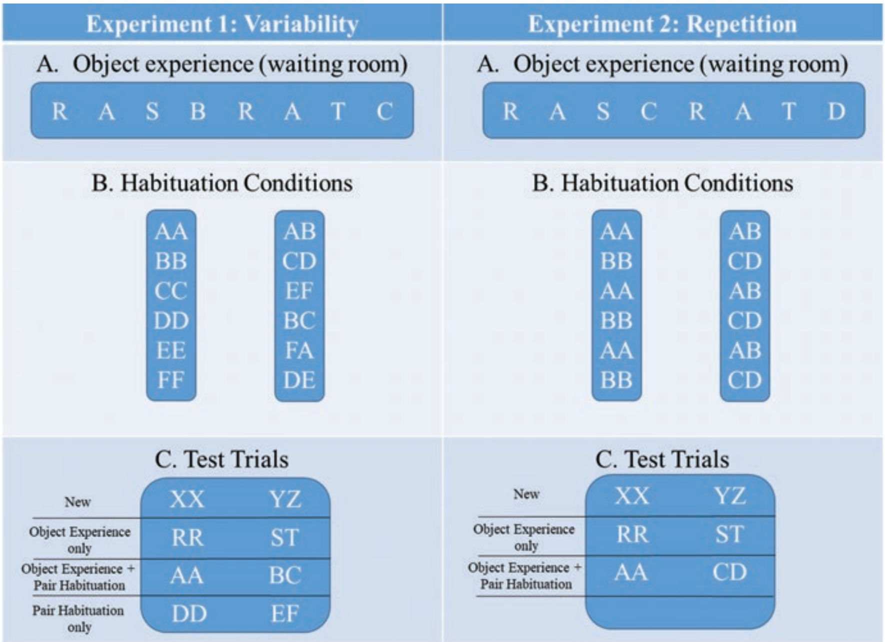


Fig. 5.3 Schematic of events in Anderson et al. (2018). In Experiment 1 on the left, infants saw six exemplars during habituation trials. In Experiment 2 on the right, infants saw an alternation between two exemplars. (**a**) In the waiting room, infants saw a subset of the individual toys before the experiment. (**b**) Infants were habituated to pairs of objects, either same or different. (**c**) In sequential test trials, looking time was recorded to the novel and familiar relational pairs across different types of test trials

When Is High Variability Helpful and When Not?

Across these studies, we have found evidence that infants can abstract a common relation from a sequence of examples. At 7–9 months, infants formed a relational abstraction from four exemplars. At 3 months, infants formed a relational abstraction with two alternating exemplars but not with six exemplars. This second finding—that 3-month-olds were better at forming an abstraction with two exemplars than with six—seems at odds with the many findings in both animal and human learning that have found that increasing the number and variability of exemplars promotes generalization (Cooper, Heron, & Heward, 2007; Thompson, Oden, & Boysen, 1997; Wasserman & Young, 2010).

The existing developmental literature reveals many studies that have found better learning with more exemplars (Bomba & Siqueland, 1983; Casasola & Park, 2013; Castro, Kennedy, & Wasserman, 2010; Gerken, 2006; Gerken & Bollt, 2008; Gomez, 2002; Needham, Dueker, & Lockhead, 2005; Quinn & Bhatt, 2005). Yet, there are a few studies that align with the "less is more" pattern (Bulf, Johnson,

& Valenza, 2011; Casasola, 2005a; Gerken & Quam, 2017; Maguire et al., 2008). These findings suggest a divide between studies in which the desired generalization depends on common object properties and those in which the desired generalization depends on relational commonalities. In the former case, more variability generally helps to broaden the generalization. But in order to form a relational abstraction, the learner must be able to carry out structural alignment over the exemplars. If the exemplars look very different from one another, the learner may fail to align them. For example, in our studies with 3-month-olds, infants could form a relational abstraction when given two alternating exemplars, but not when given six examples. We suggest that repeated exposure to two exemplars allowed the infants to go beyond noticing only the individual objects to also encode the relations, which could then be aligned across exemplars (see Casasola, 2005a, for a similar account).

The standard learning principle—"breadth of training predicts breadth of transfer"—is a useful rule, widely applicable for relatively concrete categories. But because alignment of relational structure is essential for discovering new relational commonalities, the ability to successfully compare and align is a prerequisite for relational learning (Anderson et al., 2018; Gentner & Hoyos, 2017). Thus, as Gentner and Hoyos (2017) noted, the standard principle must be amended for relational learning to be "breadth of *alignable* training predicts breadth of transfer."

Promoting Relational Learning

As noted above, structural alignment is essential to relational abstraction. But it remains true that breadth of training (in this case, alignable training) will increase generalization. Is there a way to have it both ways? Can we ensure alignment while increasing the number and variability of exemplars in infant relational learning tasks? Research on older children suggests that progressive alignment (Kotovsky & Gentner, 1996) provides a way to do this. In progressive alignment, relational learning is facilitated by initially giving children highly similar (and readily alignable) exemplars of a relation before presenting them with more surface-dissimilar pairs (Childers, Parrish, Olson, Fung, & McIntyre, 2016; Gentner et al., 2011; Gentner, Loewenstein, & Hung, 2007; Haryu et al., 2011; Hoyos, Horton, & Gentner, under review; Kotovsky & Gentner, 1996; Loewenstein & Gentner, 2001). These initial pairs with their highly similar corresponding objects are likely to be spontaneously aligned, and this alignment boosts the salience of the common relation (Gentner & Namy, 1999; Namy & Gentner, 2002). Note that progressive alignment operates quite differently from the alternation technique used by Anderson et al. (2018), in which repetition reduced the salience of the objects. In progressive alignment, close surface matches are used to seed comparison and promote initial alignment and thereby increase relational focus. Thus, in the progressive alignment condition, the infants would be presented with a series of six pairs in which the first pairs are highly similar to each other; then the variability will increase (a schematic depiction for habituation to *same* would be OO, QQ, CC, SS, WW, FF).

A second prediction is based on the idea that if comparison is critical for relational learning, then infants would never be able to learn a relation from a single example. This is consistent with our non-replication of Tyrrell et al. (1991). However, it is possible that the higher-order process of analogical comparison could interact with low-level encoding processes. In a recent set of experiments, we made the following counterintuitive prediction: for very early learners, even one example might be perceived as many due to immature/inconsistent encoding (Anderson, Hespos, & Gentner, 2019). This prediction is based on the assumption that infants' early encoding processes are unstable, resulting in variable encodings of the same external situation. This means that for the young infant, multiple exposures of a single example could be perceived as a series of highly similar pairs that share an alignable relational structure. In contrast, older infants who have a more stable ability to encode would recognize the repeated single example and would fail to learn the relation. We found that 3-month-old infants were indeed able to generalize a same or different relation from a single pair that was repeated over the course of habituation. In contrast, 7- and 9-month-olds did not generalize, though they did successfully distinguish the habituation pair from a novel pair. These findings are consistent with the idea that comparison is important to relational abstraction but highlight that comparison processes operate over representations that vary according to the learner's level.

What Paradigms Are Usually Used to Test Our Theory?

Relational learning paradigms have a diverse and extensive history stretching far back into the comparative literature. As Premack (1983) pointed out, three tasks that might seem to recruit similar processes are in fact vastly different in the ease with which animals can master them. The easiest is the object match-to-sample (MTS) task (given A, choose A over B), which can be passed by many species, including pigeons, macaques, and honeybees, as well as by 14-month-old human infants (Fagot & Thompson, 2011; Flemming, Beran, & Washburn, 2007; Giurfa, Zhang, Jenett, Menzel, & Srinivasan, 2001; Hochmann et al., 2016; Thompson et al., 1997; Wasserman & Young, 2010). In contrast, the relational match-to-sample (RMTS) task (given AA, choose XX over YZ; given BC, choose YZ over XX)[1] is far more challenging. The set of species that succeeds in the RMTS task is far smaller than the set that succeeds in object matching. So far, this set includes humans above the age of about four (without special training), chimpanzees with symbol training (Premack, 1983; Thompson et al., 1997), and hooded crows, also with considerable training (Smirnova, Zorina, Obozova, & Wasserman, 2015). The fact that success with MTS is evident across many species and success with RMTS is sparse calls for

[1] We follow Premack (1983) in restricting the term "relational match-to-sample (RMTS)" to the two-item version and refer to matches of four or more identical (or nonidentical) items (e.g., Wasserman et al., 2001) as "array match-to-sample."

an analysis of what each task requires. The MTS task requires recognizing an object match. In contrast, the RMTS task requires encoding the relation between each pair of objects and choosing the alternative that shares a relation with the standard. The third similarity task Premack discussed is the *same-different* task. Although making a same-different judgment might seem rather like making a match-to-sample, far fewer species are able to master the same-different task than can master the match-to-sample task (Premack, 1983).

Our chief reason for focusing on the same-different relation is the centrality of sameness and difference in conceptual thought. Wasserman and Young (2010) quote William James as follows: "the recognition and integration of the 'sense of sameness is the very keel and backbone of our thinking' (p. 459) as well as 'the most important of all the features of our mental structure' (p. 460)." A second reason for choosing the *same-different* task is that it has been used extensively with nonhuman primates, offering the possibility of cross-species comparison. A third, more pragmatic reason is that it can be tested without language and is therefore feasible for use with infant populations. Of course, many researchers in the comparative arena have found this option attractive for the same reason; the *same-different* task has been used with a wide variety of species (Fagot & Thompson, 2011; Fagot, Wasserman, & Young, 2001; Flemming et al., 2007; Shields, Smith, & Washburn, 1997; Thompson et al., 1997; Thompson & Oden, 2000; Truppa, Mortari, Garofoli, Privitera, & Visalberghi, 2011; Wright & Katz, 2006; Young & Wasserman, 1997, 2002; Zentall, Singer, & Miller, 2008). There appears to be broad cross-species continuity in the ability to carry out *same-different* judgments on arrays of multiple objects (Zentall et al., 2008). For example, pigeons can be trained to successfully differentiate between an array of 16 all-identical objects and an array of 16 all-different objects (Young & Wasserman, 2002). However, other research by this group indicates that pigeons could be responding to differences in degree of entropy. Studies by Young and Wasserman (1997) varied the degree of sameness within arrays of 16 objects and showed that pigeons are highly sensitive to the degree of entropy within an array (where entropy is high if all the objects are different and low if all are identical). Therefore, if we define relational ability as requiring the ability to distinguish *same pairs* (AA, BB, etc.) from *different pairs* (AB, CD), then this ability is extremely rare in nonhuman species. Nevertheless, human infants can succeed in the *same-different* task.

If we focus on the rare nonhuman species capable of making the *same-different* distinction for pairs, we find that extensive training is required for successful performance. For example, Wright and Katz (2006) were able to train rhesus monkeys, capuchin monkeys, and pigeons to distinguish *same* pairs from *different* pairs; however, to show full transfer to novel pairs, the two species of monkey required over 4700 training trials, and the pigeons required nearly 14,000 training trials. Flemming et al. (2007) showed that rhesus monkeys could learn the *same-different* task with larger arrays and that they could subsequently succeed on the *same-different* task with pairs. In general, apes—notably chimpanzees—have shown greater success in learning abstract *same-different* relations than have monkeys. The RMTS task has proved highly challenging for monkeys (but see Fagot et al., 2001) and is difficult even for young children (Christie & Gentner, 2014; Hochmann et al., 2017). However, adult humans readily pass the RMTS task.

Researchers have differed in how to interpret this difference across species. Gentner (2003, 2010) and colleagues have proposed that there is a continuum of relational ability between humans and primates. They cite work showing that chimpanzees who have learned symbols (either distinctive tokens or some other differential response) for *same* and *different* can pass the RMTS task—generally considered strong test of relational ability (Premack, 1983; Thompson et al., 1997). In contrast, Penn, Holyoak, and Povinelli (2008) propose that humans are the only species that possesses any relational ability. They discount evidence that chimpanzees can pass the RMTS task, arguing that the task can be passed via entropy detection and therefore does not indicate the ability to carry out relational matching. In making this argument, they are extrapolating from Young and Wasserman's (1997) demonstration that pigeons are responding to entropy when matching large arrays of same vs. different. However, this argument appears to be incorrect—recent research demonstrates that while the multi-item array match-to-sample can be passed via entropy detection, the classic two-item RMTS task cannot (Hochmann et al., 2017). More direct evidence comes from other recent studies that have found that chimpanzees (and bonobos) can pass relational tasks (Christie, Gentner, Call, & Haun, 2016; Haun & Call, 2009).

A more general point is that tasks that aim to measure sameness—such as MTS, same-different discrimination, and RMTS—may call on very different processes and knowledge. This is important for understanding what we can infer from these tasks. For example, passing the object MTS task does not require forming the relation of *same*. We know this because many animals can pass the MTS task but will fail to learn a same-different discrimination. All we can infer when an animal (or infant) passes the MTS task is that seeing two identical objects feels different from seeing two distinct objects[2]. Likewise, being able to pass the RMTS task does not require forming a higher-order relation of sameness between the two SAME relations. To spell out this analogy:

> *Matching X with X instead of Y does not imply that the animal has formed a relation of SAME (X,X).*
> *Likewise, Matching (X,X) with (A,A) instead of (B,C) does not imply that the animal has formed a higher-order relation of SAME {SAME (X,X), SAME (A,A)}.*

In any case, it is clear that humans excel in relational ability, even compared to our nearest cousins among the great apes. This examination of the comparative literature reveals two important points for understanding infant relational ability. First, focusing on the same-different task, human infants readily learn *same-different* discrimination. This contrasts with the difficulty many other species experience with

[2] Further, Hochmann et al. (2016) have found evidence suggesting that 14-month-olds in a non-match to same task pass the "different" task by first finding the match and then choosing the other one.

these relations. Second, infants learn the relation in very few trials (six to nine habituation trials), whereas nonhuman species often require extensive training.

How Could Structure-Mapping Theory Extend Beyond Contexts?

The work that we described in this chapter differs from most work on infant cognition in that it focuses on the nature of the learning process not the nature of the representation. Research on infants' expectations about how objects behave and interact has made enormous progress in the last 30 years and has revealed impressive early capacities in several different arenas, including spatial relations (Casasola, 2005b; Casasola & Cohen, 2002; Hespos, Grossman, & Saylor, 2010; Hespos & Piccin, 2009; Hespos, Saylor, & Grossman, 2009; Hespos & Spelke, 2004; Kibbe & Feigenson, 2015; McDonough, Choi, & Mandler, 2003; Moher, Tuerk, & Feigenson, 2012; Quinn, Cummins, Kase, Martin, & Weissman, 1996) and physical reasoning (Baillargeon, 1994; Hespos & Baillargeon, 2001, 2006, 2008; Needham & Baillargeon, 1993; Wang & Baillargeon, 2008 for reviews, see Baillargeon, Li, Gertner, & Wu, 2011; Baillargeon, Li, Ng, & Yuan, 2009; Spelke, Breinlinger, Macomber, & Jacobson, 1992). This work has focused on tracing the early development of understanding of spatial and physical events. Thus, the focus of this prior research is on revealing the knowledge infants have acquired in the world and how that knowledge supports infants' expectations. In contrast, the focus for this chapter is on the *learning processes during the experiment itself*. We suggest that the structure-mapping approach to learning has implications for many other arenas of human learning. Here we discuss two such areas: language learning and learning about the physical world.

Structure-mapping theory leads to a set of predictions concerning how comparison can benefit language learning (Gentner, 2010; Gentner et al., 2007; Gentner & Christie, 2010):

- Comparing two things engages a structural alignment process that renders their commonalties more salient—and this effect is greatest for common relational structure (Gentner & Namy, 1999).
- Structural alignment also renders *alignable differences*—differences that play the same role in the common relational structure—more salient (Gentner & Markman, 2006; Markman & Gentner, 1993; Sagi, Gentner, & Lovett, 2012).
- Progressive alignment is beneficial in early learning. Early in learning, when domain knowledge is weak, alignment purely on the basis of relations is often impossible. In progressive alignment, learners are first given a close overall similarity match that instantiates the desired relational structure, as exemplified below.

As Gentner and Namy (2006) reviewed, there is considerable evidence that language learning benefits from these processes. Studies of word learning have demonstrated the power of comparison to reveal common relational structure. For example, Gentner and Namy (1999) taught 4-year-olds a new noun (e.g., "blicket" for a bicycle) and asked them to choose another blicket. Children mostly chose a perceptually similar alternative (eyeglasses) instead of a perceptually dissimilar object from the same conceptual category (a skateboard). The same result occurred for children who were told that "blicket" was a name for a tricycle. But when a third group of 4-year-olds was shown both the bicycle and the tricycle, told that they were both blickets, and asked "can you see why these are both blickets?," the results were strikingly different. Despite the fact that they had twice as much evidence for the matching perceptual features, they chose the conceptual match (the skateboard). Gentner and Namy (1999) (see also Namy & Gentner, 2002) concluded that structurally aligning the two standards had highlighted their common causal and functional relations. Gentner et al. (2011) found that comparison aided children aged 3–6 years in learning the meanings of relational nouns—nouns such as *container*, whose meanings are determined not by common features but by common relations.

As Childers (2011) and colleagues have noted, this feature of structural alignment—that it preferentially highlights common relational structure—suggests that it would be particularly applicable to verb learning (see Imai & Childers, this volume). Learning verb meanings is challenging to young children. Not only are verbs slower to enter the vocabulary than nouns ((Bates et al., 1994; Bornstein, 2004; Gentner, 1982; Gentner & Boroditsky, 2001; Gleitman, Cassidy, Nappa, Papafragou, & Trueswell, 2005; Imai, Haryu, & Okada, 2005; MacNamara, 1972), but also even when children do learn a new verb, they often initially use it in a highly restrictive way (Forbes & Poulin-Dubois, 1997; Huttenlocher, Smiley, & Charney, 1983; Tomasello, 1992, 2000). Thus, an important question is how—by what processes—children acquire and extend new verbs. There is a growing body of research and theory that supports the idea that structure-mapping processes are integral to this learning (Childers, 2011; Childers, Hirshkowitz, & Benavides, 2014; Childers & Paik, 2009; Haryu et al., 2011; Tomasello, 2000). For example, Childers and colleagues have shown that children benefit from seeing multiple enactments of a given verb, rather than repeated enactments with the same objects (Childers & Paik, 2009). In another study, Childers, Heard, Ring, Pai, and Sallquist (2012) found that 2.5-year-olds taught a new verb performed as well after seeing a set of comparable enactments as they did after receiving direct instruction about the verb from an experimenter.

Other research on language learning has found evidence for a more specific prediction of structural alignment theory: namely, that progressive alignment benefits early learning. Progressive alignment is a way of addressing a bottleneck that arises in children's relational learning. Comparing two examples (such as two sentences involving the same novel verb) is a route to relational learning, but early in learning, children may lack sufficient relational knowledge to be able to align two disparate examples. In progressive alignment, learners are first given a close, overall similarity match that instantiates the desired relational structure. The high overall similarity makes it likely that children will spontaneously compare the two examples, and because the object matches are consistent with the relational alignment, young learners are likely to arrive at the correct alignment. Thus, progressive alignment can serve as "training wheels" for purely relational matches (Gentner, 2010; Gentner & Medina, 1998).

Childers et al. (2016) asked whether progressive alignment could aid children's verb learning. Indeed, the study found evidence that 3.5-year-old children benefit from progressive alignment. They presented children with two novel verbs under three conditions. In one condition, each verb was enacted four times with the same objects. In the progressive alignment condition, each verb was enacted first with highly similar objects playing similar roles in the events, followed by two events in which the objects were highly dissimilar across the enactments of the verb. In the all-far condition, each verb was enacted four times, with all enactments having highly dissimilar objects. After children witnessed these enactments, they were asked to enact the verb themselves, first using new objects similar to the ones used in the learning trials and one (the "far extension") using dissimilar objects. There were two results of note. First, children seeing multiple enactments of the same verb produced more correct extensions on the test than would children seeing a single enactment, consistent with prior findings (e.g., Childers, 2011; Childers & Paik, 2009). Second, on the critical far test, children who received progressive alignment from highly similar to less similar enactments performed best—significantly better than the single-enactment group.

Another study of progressive alignment in verb learning was done by Haryu et al. (2011). They taught 4-year-old children a verb for a novel event and asked whether the children could extend the verb to other events. They found that children were initially limited to close overall matches (i.e., literally similar events). That is, they extended the verb only when the new event shared similar objects as well as depicting the same action as the initial event; they failed when the objects were dissimilar, even when the new event shared its action with the initial event. In a second study, Haryu et al. found that progressive alignment from close to far matches enabled a new group of 4-year-olds to extend the verb based on sameness of action, without support from object similarity. Similarly Gentner et al. (2007) used high object similarity to help children to make the correct correspondences, thus supporting the correct alignment of relational structure. As in other work with progressive alignment, structural alignment resulted in heightening the common structure, which the children could then extend to an event that shared only that structure.

These findings are consistent with the general position that initial representations of verbs may be quite concrete and tied to the context in which they are learned (Lieven, Pine, & Baldwin, 1997; Tomasello, 1992, 2000) and that comparisons between current and stored utterances lead to more general, abstract representations of verb meaning. Initially, those comparisons will be between overall similar utterances, in which verbs appear in very similar frames. But via progressive alignment, these early concrete matches will potentiate future more abstract matches (Childers & Paik, 2009; Childers & Tomasello, 2001; Pruden, Shallcross, Hirsh-Pasek, & Golinkoff, 2008; see Gentner & Namy, 2006, for a more extensive discussion).

Structure mapping also has application to studies of artificial grammar learning—another arena in which infant researchers have investigated learning during the course of the experiment. Many artificial grammar tasks can be viewed as relational learning tasks (Aslin & Newport, 2012; Gerken, 2006; Gomez & Gerken, 2000; Johnson et al., 2009; Kuehne, Gentner, & Forbus, 2000; Marcus, Vijayan, Rao, & Vishton, 1999; Saffran, Pollak, Seibel, & Shkolnik, 2007). For example, in Marcus et al.'s (1999) study, after 7.5-month-olds heard 48 examples (16 patterns, three times each) of a syllable pattern such as AAB, they could then discriminate new instances of the AAB pattern from instances of an ABA pattern, even when all the specific syllables were new (see also Gomez & Gerken, 1999). Further, there is evidence that the ability to generalize across such patterns may operate across a broad range of stimuli, including tones and visual stimuli (Gomez & Gerken, 2000; Johnson et al., 2009; Saffran et al., 2007).

We suggest that structure mapping provides a natural mechanism for this process. Two key points supporting this claim are (1) by 7 months (and even earlier) infants are capable of structural alignment and abstraction and (2) our simulations reveal that the structural alignment process can capture key phenomena in artificial grammar learning. To take the first point, our studies of *same-different* learning show that infants can form a relational abstraction over a series of examples. More specifically, this process shows signatures of structural alignment and mapping, as discussed earlier.

Support for the second point—that the same process of structural alignment and abstraction can account for infants' artificial grammar learning—comes from simulation studies. Kuehne et al. (2000) showed that a computational model of analogical generalization called the sequential learning engine (SEQL) can capture the Marcus et al. (1999) findings. SEQL and its successor, SAGE,[3] use the structure-mapping engine (SME; Falkenhainer, Forbus, & Gentner, 1989; Forbus et al., 2017) to iteratively compare input examples, creating an ongoing generalization. If SAGE (or SEQL) is given an input example, it will store that example. If the example is followed by another, SAGE compares it to the first one, using SME. If there is sufficient overlap (i.e., if SME's score is above a preset threshold), the common structure is stored as a generalization. If the overlap is below threshold, the example will be stored separately. This process continues as new examples arrive; if new examples are sufficiently similar to the ongoing generalization, they are assimilated into it and the generalization is updated. New examples that cannot be assimilated into the main abstraction are compared to the set of examples; if a new example is very similar to a stored example, a new generalization is formed from their common structure. Thus, it naturally results in a generalization (or sometimes more than one generalization) plus exceptions.

SEQL was given the same input as the infants in Marcus et al.: three repetitions of each of the 16 three-syllable strings, for a total of 48 strings. Each syllable was encoded as having 12 phonemic features (following Elman, 1998). The relational pattern within each string (e.g., AAB) was encoded by Magi, which uses SME to encode symmetry and repetition within an item (Ferguson, 1994). As the strings were

[3] SAGE (McLure, Friedman, & Forbus, 2015) operates using the same basic iterative comparison process as SEQL but keeps track of frequency information about alignable structures, enabling it to produce probabilistic generalizations.

presented, SEQL computed a generalization by comparing the first two exemplars (via SME) and storing their common structure and then incrementally comparing subsequent exemplars to the ongoing abstraction. After all 48 exemplars were presented, SEQL was given two test strings with new syllables. Like the infants in Marcus et al., SEQL found the test string with the same relational structure (e.g., CCD) more similar to its generalization than the one with different structure (e.g., CDC).

Structure mapping has application to studies on physical reasoning too. For example, in a series of six experiments, Wang and Baillargeon (2008) describe teaching trials that helped and hindered infants' learn the variable of height in covering events 1 month earlier than usual. The authors describe their findings in the context of their explanation-based learning theory. However, understanding these studies in the context of relational learning illustrates the broad context in which this ability operates. Successful learning was demonstrated in Experiments 1 and 2 that allowed infants to compare the height of the object to the height of multiple tall and short covers. In a third experiment, they replicated the effect of learning after three comparison trials even when the test was delayed by 24 hours. Learning was hindered in low-alignment conditions. In Experiment 4, infants failed to learn when they could not compare between an object being fully and partially hidden. In Experiment 5, infants failed to learn when there was no direct comparison between the relative heights of the covers and the object. Given the key roles of visual alignment and comparison across these experiments, structure-mapping theory predicts the same pattern of results.

Conclusions

We began this chapter highlighting the amazing ability humans have for deriving relational abstractions. Like many other animals, we can learn by association and by perceptual generalization. However, unlike most other species, we also acquire new information by means of relational generalization and transfer. In this chapter, we explore the origins of a uniquely developed human capacity—our ability to learn relational abstractions through analogical comparison. We focus on whether and how infants can use analogical comparison to derive relational abstractions from examples. We frame our work in terms of structure-mapping theory, which has been fruitfully applied to analogical processing in children and adults. We find that young infants show two key signatures of structure mapping: first, relational abstraction is fostered by comparing alignable examples, and second, relational abstraction is hampered by the presence of highly salient objects. The studies we review make it clear that structure-mapping processes are evident in the first months of life, prior to much influence of language and culture. This finding suggests that infants are born with analogical processing mechanisms that allow them to learn relations through comparing examples.

Turning to very early learning, we augmented our account by considering the nature of young infants' encoding processes, leading to two counterintuitive predictions. First, we predicted that young infants (2–3 months old) would be better able to form a relational abstraction when given two alternating exemplars than when given six different exemplars. This is based on the assumption that young infants may initially focus on the individual objects and shift to noticing the relation between them after repetition of the exemplar (Casasola, 2005a). As predicted, this pattern was found for young but not older infants. Second, we predicted that younger, but not older, infants would be able to form a relational abstraction from one repeated exemplar; the prediction follows from the assumption that young infants have unstable encoding processes.

Next, we revisited Premack's insight from 1983 that the tasks used to measure analogical abilities (RMTS, MTS, and *same/different* discrimination) are vastly different from each other. The takeaway from this section is that while many species can learn through association and perceptual generalizations, there are relatively few species that can succeed in the *same/different* discrimination task. Of the species that can succeed in the *same/different* task, humans are unique in that they need fewer than 10 trials to learn such relations. In the final sections, we reviewed how structure mapping extends to language acquisition, artificial grammar learning, and physical reasoning. The value of investigating the origins of our analogical abilities is that we will be in a better position to understand how language and culture capitalize on cognitive abilities. More broadly, we can address whether essential differences between humans and other species are evident from the earliest points in development.

References

Anderson, E. M., Chang, Y. J., Hespos, S., & Gentner, D. (2018). Comparison within pairs promotes analogical abstraction in three-month-olds. *Cognition, 176*, 74–86.

Anderson, E. M., Hespos, S. J., & Gentner, D. (2019). *When one example behaves like many*. Manuscript in preparation.

Aslin, R. N., & Newport, E. L. (2012). Statistical learning: From acquiring specific items to forming general rules. *Current Directions in Psychological Science, 21*(3), 170–176. https://doi.org/10.1177/0963721412436806

Baillargeon, R. (1994). How do infants learn about the physical world? *Current Directions in Psychological Science, 3*(5), 133–140. https://doi.org/10.1111/1467-8721.ep10770614

Baillargeon, R., Li, J., Gertner, Y., & Wu, D. (2011). How do infants reason about physical events? In U. Goswami (Ed.), *The Wiley-Blackwell handbook of childhood cognitive development* (2nd ed., pp. 11–48). Oxford, UK: Blackwell.

Baillargeon, R., Li, J., Ng, W., & Yuan, S. (2009). An account of infants' physical reasoning. In A. Woodward & A. Needham (Eds.), *Learning and the infant mind* (pp. 66–116). New York: Oxford University Press.

Bates, E., Marchman, V., Thal, D., Fenson, L., Dale, P., Reznick, J. S., et al. (1994). Developmental and stylistic variation in the composition of early vocabulary. *Journal of Child Language, 21*(1), 85–123.

Bomba, P. C., & Siqueland, E. R. (1983). The nature and structure of infant form categories. *Journal of Experimental Child Psychology, 35*(2), 294–328.

Bornstein, G. (2004). A taxonomy of social situations. *International Journal of Psychology, 39*(5–6), 385–385.

Bulf, H., Johnson, S. P., & Valenza, E. (2011). Visual statistical learning in the newborn infant. *Cognition, 121*(1), 127–132.

Carey, S. (2010). Beyond fast mapping. *Language Learning and Development, 6*(3), 184–205.

Casasola, M. (2005a). When less is more: How infants learn to form an abstract categorical representation of support. *Child Development, 76*(1), 279–290. https://doi.org/10.1111/j.1467-8624.2005.00844.x

Casasola, M. (2005b). Can language do the driving? The effect of linguistic input on infants' categorization of support spatial relations. *Developmental Psychology, 41*, 183–192.

Casasola, M., & Cohen, L. B. (2002). Infant categorization of containment, support, and tight-fit spatial relationships. *Developmental Science, 5*, 247–264.

Casasola, M., & Park, Y. (2013). Developmental changes in infant spatial categorization: When more is best and when less is enough. *Child Development, 84*(3), 1004–1019. https://doi.org/10.1111/cdev.12010

Castro, L., Kennedy, P. L., & Wasserman, E. A. (2010). Conditional same-different discrimination by pigeons: Acquisition and generalization to novel and few-item displays. *Journal of Experimental Psychology: Animal Behavior Processes, 36*(1), 23.

Childers, J. B. (2011). Attention to multiple events helps two 1/2-year-olds extend new verbs. *First Language, 31*, 3–22.

Childers, J. B., Heard, M. E., Ring, K., Pai, A., & Sallquist, J. (2012). Children use different cues to guide noun and verb extensions. *Language Learning and Development, 8*(3), 233–254.

Childers, J. B., Hirshkowitz, A., & Benavides, K. (2014). Attention to explicit and implicit contrast in verb learning. *Journal of Cognition and Development, 15*(2), 213–237.

Childers, J. B., & Paik, J. H. (2009). Korean-and English-speaking children use cross-situational information to learn novel predicate terms. *Journal of Child Language, 36*(1), 201–224.

Childers, J. B., Parrish, R., Olson, C., Fung, G., & McIntyre, K. (2016). Experience comparing similar events helps children extend new verbs. *Journal of Cognition and Development, 17*, 41–66.

Childers, J. B., & Tomasello, M. (2001). The role of pronouns in young children's acquisition of the English transitive construction. *Developmental Psychology, 37*(6), 739.

Christie, S., & Gentner, D. (2007). Relational similarity in identity relation: The role of language. In S. Vosniadou & D. Kayser (Eds.), *Proceedings of the Second European Cognitive Science Conference*. London: Taylor & Francis.

Christie, S., & Gentner, D. (2010). Where hypotheses come from: Learning new relations by structural alignment. *Journal of Cognition and Development, 11*(3), 356–373. https://doi.org/10.1080/15248371003700015

Christie, S., & Gentner, D. (2014). Language helps children succeed on a classic analogy task. *Cognitive Science, 38*, 383–397.

Christie, S., Gentner, D., Call, J., & Haun, D. B. M. (2016). Sensitivity to relational similarity and object similarity in apes and children. *Current Biology, 26*(4), 531–535.

Cooper, J. O., Heron, T. E., & Heward, W. L. (2007). *Applied behavior analysis* (2nd ed.). Upper Saddle River, NJ: Pearson Education.

DeLoache, J. S. (1995). Early symbol understanding and use. In D. L. Medin (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 33, pp. 65–114). San Diego, CA: Academic Press.

Doumas, L. A. A., & Hummel, J. E. (2013). Comparison and mapping facilitate relation discovery and predication. *PLoS One, 8*(6), e63889. https://doi.org/10.1371/journal.pone.0063889

Doumas, L. A. A., Hummel, J. E., & Sandhofer, C. M. (2008). A theory of the discovery and predication of relational concepts. *Psychological Review, 115*(1), 1–43. https://doi.org/10.1037/0033-295x.115.1.1

Elman, J. L. (1998). Connectionism, artificial life, and dynamical systems: New approaches to old questions. A companion to cognitive science. Oxford: Basil Blackwood, 233-277.

Fagot, J., & Thompson, R. K. R. (2011). Generalized relational matching by guinea baboons (Papio papio) in two-by-two-item analogy problems. *Psychological Science, 22*(10), 1304–1309. https://doi.org/10.1177/0956797611422916

Fagot, J., Wasserman, E. A., & Young, M. E. (2001). Discriminating the relation between relations: The role of entropy in abstract conceptualization by baboons (Papio papio) and humans (Homo sapiens). *Journal of Experimental Psychology: Animal Behavior Processes, 27*(4), 316.

Falkenhainer, B., Forbus, K. D., & Gentner, D. (1989). The structure-mapping engine: Algorithm and examples. *Artificial Intelligence, 41*(1), 1–63.

Ferguson, R. W. (1994). MAGI: Analogy-based encoding using regularity and symmetry. In *Proceedings of the 16th annual conference of the Cognitive Science Society* (pp. 283–288). New York: Psychology Press.

Ferry, A. L., Hespos, S. J., & Gentner, D. (2015). Prelinguistic relational concepts: Investigating analogical processing in infants. *Child Development, 86*, 1386–1405.

Ferry, A. L., Hespos, S. J., & Waxman, S. R. (2010). Categorization in 3- and 4-month-old infants: An advantage of words over tones. *Child Development, 81*(2), 472–479.

Flemming, T. M., Beran, M. J., & Washburn, D. A. (2007). Disconnect in concept learning by rhesus monkeys (Macaca mulatta): Judgment of relations and relations-between-relations. *Journal of Experimental Psychology: Animal Behavior Processes, 33*(1), 55.

Forbes, J. N., & Poulin-Dubois, D. (1997). Representational change in young children's understanding of familiar verb meaning. *Journal of Child Language, 24*(2), 389–406.

Forbus, K. D., Ferguson, R. W., Lovett, A., & Gentner, D. (2017). Extending SME to handle large-scale cognitive modeling. *Cognitive Science, 41*, 1152–1201.

Gentner, D. (1982). Why nouns are learned before verbs: Linguistic relativity versus natural partitioning. *Center for the Study of Reading Technical Report*; no. 257.

Gentner, D. (1983). Structure-mapping: A theoretical framework for analogy. *Cognitive Science, 7*, 155–170.

Gentner, D. (1988). Metaphor as structure mapping: The relational shift. *Child Development, 59*, 47–59.

Gentner, D. (1989). The mechanisms of analogical learning. In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical reasoning* (pp. 199–241). London: Cambridge University Press.

Gentner, D. (2003). Why we're so smart. In D. Gentner & S. Goldin-Meadow (Eds.), *Language in mind: Advances in the study of language and thought* (pp. 195–235). Cambridge, MA: MIT Press.

Gentner, D. (2005). The development of relational category knowledge. In L. Gershkoff-Stowe & D. H. Rakison (Eds.), *Building object categories in developmental time* (pp. 245–275). Hillsdale, NJ: Erlbaum.

Gentner, D. (2010). Bootstrapping the mind: Analogical processes and symbol systems. *Cognitive Science, 34*(5), 752–775. https://doi.org/10.1111/j.1551-6709.2010.01114.x

Gentner, D., Anggoro, F. K., & Klibanoff, R. S. (2011). Structure mapping and relational language support children's learning of relational categories. *Child Development, 82*(4), 1173–1188. https://doi.org/10.1111/j.1467-8624.2011.01599.x

Gentner, D., & Boroditsky, L. (2001). Individuation, relativity, and early word learning. *Language Acquisition and Conceptual Development, 3*, 215–256.

Gentner, D., & Christie, S. (2010). Mutual bootstrapping between language and analogical processing. *Language and Cognition, 2*(2), 261–283.

Gentner, D., & Forbus, K. (2011). Computational models of analogy. *WIREs Cognitive Science, 2*, 266–276.

Gentner, D., Holyoak, K. J., & Kokinov, B. N. (2001). *The analogical mind: Perspectives from cognitive science*. Cambridge, MA: The MIT Press.

Gentner, D., & Hoyos, C. (2017). Analogy & abstraction. *Topics in Cognitive Science, 9*(3), 672–693.

Gentner, D., Loewenstein, J., & Hung, B. (2007). Comparison facilitates children's learning of names for parts. *Journal of Cognition and Development, 8*, 285–307.

Gentner, D., & Markman, A. B. (1997). Structure mapping in analogy and similarity. *American Psychologist, 52*, 45–56.
Gentner, D., & Markman, A. B. (2006). Defining structural similarity. *The Journal of Cognitive Science, 6*, 1–20.
Gentner, D., & Medina, J. (1998). Similarity and the development of rules. *Cognition, 65*(2–3), 263–297. https://doi.org/10.1016/s0010-0277(98)00002-x
Gentner, D., & Namy, L. (1999). Comparison in the development of categories. *Cognitive Development, 14*, 487–513.
Gentner, D., & Namy, L. L. (2006). Analogical processes in language learning. *Current Directions in Psychological Science, 15*(6), 297–301.
Gentner, D., & Rattermann, M. J. (1991). Language and the career of similarity. In S. A. Gelman & J. P. Byrnes (Eds.), *Perspectives on language and thought: Interrelations in development* (pp. 225–277). New York: Cambridge University Press.
Gentner, D., & Toupin, C. (1986). Systematicity and surface similarity in the development of analogy. *Cognitive Science, 10*(3), 277–300. https://doi.org/10.1207/s15516709cog1003_2
Gerken, L. (2006). Decisions, decisions: Infant language learning when multiple generalizations are possible. *Cognition, 98*(3), B67–B74. https://doi.org/10.1016/j.cognition.2005.03.003
Gerken, L., & Bollt, A. (2008). Three exemplars allow at least some linguistic generalizations: Implications for generalization mechanisms and constraints. *Language Learning and Development, 4*(3), 228–248.
Gerken, L., & Quam, C. (2017). Infant learning is influenced by local spurious generalizations. *Developmental Science, 20*(3), e12410.
Gick, M. L., & Holyoak, K. J. (1983). Schema induction and analogical transfer. *Cognitive Psychology, 15*(1), 1–38. https://doi.org/10.1016/0010-0285(83)90002-6
Giurfa, M., Zhang, S., Jenett, A., Menzel, R., & Srinivasan, M. V. (2001). The concepts of 'sameness' and 'difference' in an insect. *Nature, 410*(6831), 930.
Gleitman, L. R., Cassidy, K., Nappa, R., Papafragou, A., & Trueswell, J. C. (2005). Hard words. *Language Learning and Development, 1*(1), 23–64.
Gomez, R. L. (2002). Variability and detection of invariant structure. *Psychological Science, 13*(5), 431–436.
Gomez, R. L., & Gerken, L. (1999). Artificial grammar learning by 1-year-olds leads to specific and abstract knowledge. *Cognition, 70*(2), 109–135.
Gomez, R. L., & Gerken, L. (2000). Infant artificial language learning and language acquisition. *Trends in Cognitive Sciences, 4*(5), 178–186. https://doi.org/10.1016/s1364-6613(00)01467-4
Halford, G. S. (1992). Analogical reasoning and conceptual complexity in cognitive development. *Human Development, 35*(4), 193–217. https://doi.org/10.1159/000277167
Haryu, E., Imai, M., & Okada, H. (2011). Object similarity bootstraps young children to action-based verb extension. *Child Development, 82*(2), 674–686.
Haun, D. B. M., & Call, J. (2009). Great apes' capacities to recognize relational similarity. *Cognition, 110*, 147–159.
Hermer, L., & Spelke, E. S. (1994). A geometric process for spatial reorientation in young children. *Nature, 370*, 57–59.
Hespos, S. J., & Baillargeon, R. (2001). Infants' knowledge about occlusion and containment events: A surprising discrepancy. *Psychological Science, 12*(2), 141–147. https://doi.org/10.1111/1467-9280.00324
Hespos, S. J., & Baillargeon, R. (2006). Decalage in infants' knowledge about occlusion and containment events: Converging evidence from action tasks. *Cognition, 99*(2), B31–B41. https://doi.org/10.1016/J.Cognition.2005.01.010
Hespos, S. J., & Baillargeon, R. (2008). Young infants' actions reveal their developing knowledge of support variables: Converging evidence for violation-of-expectation findings. *Cognition, 107*(1), 304–316. https://doi.org/10.1016/J.Cognition.2007.07.009
Hespos, S. J., Grossman, S. R., & Saylor, M. M. (2010). Infants' ability to parse continuous actions: Further evidence. *Neural Networks, 23*(8–9), 1026–1032. https://doi.org/10.1016/J.Neunet.2010.07.010

Hespos, S. J., & Piccin, T. B. (2009). To generalize or not to generalize: spatial categories are influenced by physical attributes and language. *Developmental Science, 12*(1), 88–95. https://doi.org/10.1111/J.1467-7687.2008.00749.X

Hespos, S. J., Saylor, M. M., & Grossman, S. R. (2009). Infants' ability to parse continuous actions. *Developmental Psychology, 45*(2), 575–585. https://doi.org/10.1037/A0014145

Hespos, S. J., & Spelke, E. S. (2004). Conceptual precursors to language. *Nature, 430*(6998), 453–456. https://doi.org/10.1038/Nature02634

Hochmann, J. R., Mody, S., & Carey, S. (2016). Infants' representations of same and different in match-and non-match-to-sample. *Cognitive Psychology, 86*, 87–111.

Hochmann, J. R., Tuerk, A. S., Sandborn, S., Zhu, R., Long, R., Dempster, M., et al. (2017). Children's representation of abstract relations in relational/array match-to-sample tasks. *Cognitive Psychology, 99*, 17–43.

Hoyos, C., Horton, W. S., & Gentner, D. (under review). *Analogical comparison promotes theory-of-mind development*.

Huttenlocher, J., Smiley, P., & Charney, R. (1983). Emergence of action categories in the child: Evidence from verb meanings. *Psychological Review, 90*(1), 72.

Imai, M., Haryu, E., & Okada, H. (2005). Mapping novel nouns and verbs onto dynamic action events: Are verb meanings easier to learn than noun meanings for Japanese children? *Child Development, 76*, 340–355.

Jamrozik, A., & Gentner, D. (under review). *Labeling unlocks inert knowledge*.

Johnson, S. P., Fernandes, K. J., Frank, M. C., Kirkham, N., Marcus, G., Rabagliati, H., et al. (2009). Abstract rule learning for visual sequences in 8- and 11-month-olds. *Infancy, 14*(1), 2–18. https://doi.org/10.1080/15250000802569611

Kibbe, M., & Feigenson, L. (2015). Infants use temporal regularities to chunk objects in memory. *Cognition, 146*, 251–263.

Kotovsky, L., & Gentner, D. (1996). Comparison and categorization in the development of relational similarity. *Child Development, 67*, 2797–2822.

Kuehne, S. E., Gentner, D., & Forbus, K. D. (2000). Modeling infant learning via symbolic structural alignment. In L. R. Gleitman & A. K. Joshi (Eds.), *Proceedings of the Twenty-Second Annual Conference of the Cognitive Science Society* (pp. 286–291). Mahwah, NJ: Lawrence Erlbaum.

Lieven, E. V. M., Pine, J. M., & Baldwin, G. (1997). Lexically-based learning and early grammatical development. *Journal of Child Language, 24*, 187–219.

Liu, J., Golinkoff, R. M., & Sak, K. (2001). One cow does not an animal make!: Children can extend novel words at the superordinate level. *Child Development, 72*, 1674–1694.

Loewenstein, J., & Gentner, D. (2001). Spatial mapping in preschoolers: Close comparisons facilitate far mappings. *Journal of Cognition and Development, 2*(2), 189–219.

Loewenstein, J., & Gentner, D. (2005). Relational language and the development of relational mapping. *Cognitive Psychology, 50*(4), 315–353. https://doi.org/10.1016/j.cogpsych.2004.09.004

Lupyan, G. (2012). Linguistically modulated perception and cognition: The label-feedback hypothesis. *Frontiers in Psychology, 2*, 114–126. https://doi.org/10.3389/fpsyg.2012.00054

Macnamara, J. (1972). Cognitive basis of language learning in infants. *Psychological Review, 79*(1), 1.

Maguire, M. J., Hirsh-Pasek, K., Golinkoff, R. M., & Brandone, A. C. (2008). Focusing on the relation: Fewer exemplars facilitate children's initial verb learning and extension. *Developmental Science, 11*(4), 628–634. https://doi.org/10.1111/j.1467-7687.2008.00707.x

Marcus, G. F., Vijayan, S., Rao, S. B., & Vishton, P. M. (1999). Rule learning by seven-month-old infants. *Science, 283*(5398), 77–80. https://doi.org/10.1126/science.283.5398.77

Markman, A. B., & Gentner, D. (1993). Structural alignment during similarity comparisons. *Cognitive Psychology, 25*(4), 431–467. https://doi.org/10.1006/cogp.1993.1011

Markman, A. B., & Gentner, D. (1997). The effects of alignability on memory storage. *Psychological Science, 8*(5), 363–367.

Markman, A. B., & Wisniewski, E. J. (1997). Similar and different: The differentiation of basic-level categories. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23*(1), 54.

McDonough, L., Choi, S., & Mandler, J. M. (2003). Understanding spatial relations: Flexible infants, lexical adults. *Cognitive Psychology, 46*(3), 229–259. https://doi.org/10.1016/S0010-0285(02)00514-5

McLure, M. D., Friedman, S. E., & Forbus, K. D. (2015, February). Extending analogical generalization with near-misses. In *Twenty-Ninth AAAI Conference on Artificial Intelligence*. Palo Alto, CA: AAAI Press.

Moher, M., Tuerk, A., & Feigenson, L. (2012). Severn-month-old infants chuck items in memory. *Journal of Experimental Child Psychology, 112*(4), 361–377.

Namy, L. L., & Clepper, L. (2010). The differing roles of comparison and contrast in categorization. *Journal of Experimental Child Psychology, 107*, 291–305.

Namy, L. L., & Gentner, D. (2002). Making a silk purse out of two sow's ears: Young children's use of comparison in category learning. *Journal of Experimental Psychology: General, 131*(1), 5–15. https://doi.org/10.1037/0096-3445.131.1.5

Needham, A., & Baillargeon, R. (1993). Intuitions about support in 4.5-month-old infants. *Cognition, 47*, 121–148.

Needham, A., Dueker, G., & Lockhead, G. (2005). Infants' formation and use of categories to segregate objects. *Cognition, 94*, 215–240.

Paik, J. H., & Mix, K. S. (2006). Preschoolers' use of surface similarity in object comparisons: Taking context into account. *Journal of Experimental Child Psychology, 95*(3), 194–214. https://doi.org/10.1016/j.jecp.2006.06.002

Paik, J. H., & Mix, K. S. (2008). It's all relative: Different levels of relational similarity used in children's comparisons. *British Journal of Developmental Psychology, 26*(4), 499–505.

Penn, D. C., Holyoak, K. J., & Povinelli, D. J. (2008). Darwin's mistake: Explaining the discontinuity between human and nonhuman minds. *Behavioral and Brain Sciences, 31*(2), 109–130. https://doi.org/10.1017/s0140525x08003543

Premack, D. (1983). The codes of man and beasts. *Behavioral and Brain Sciences, 6*(1), 125–167. https://doi.org/10.1017/s0140525x00015077

Pruden, S. M., Shallcross, W. L., Hirsh-Pasek, K., & Golinkoff, R. M. (2008). Foundations of verb learning: Comparison helps infants abstract event components. In *Proceedings of the 32nd annual Boston University Conference on Language Development* (pp. 402–414). Somerville, MA: Cascadilla.

Pyers, J. E., & Senghas, A. (2007). Reported action in Nicaraguan and American sign languages: Emerging versus established systems. In P. Perniss, R. Pfau, & M. Steinbach (Eds.), *Visible variation: Comparative studies on sign language structure*. Berlin, Germany: Mouton de Gruyter.

Pyers, J. E., Shusterman, A., Senghas, A., Spelke, E. S., & Emmorey, K. (2010). Evidence from an emerging sign language reveals that language supports spatial cognition. *Proceedings of the National Academy of Sciences, 107*(27), 12116–12120.

Quinn, P. C., & Bhatt, R. S. (2005). Learning perceptual organization in infancy. *Psychological Science, 16*(7), 511–515.

Quinn, P. C., Cummins, M., Kase, J., Martin, E., & Weissman, S. (1996). Development of categorical representations for above and below spatial relations in 3- to 7-month-old infants. *Developmental Psychology, 32*(5), 942–950. https://doi.org/10.1037/0012-1649.32.5.942

Richland, L. E., Morrison, R. G., & Holyoak, K. J. (2006). Children's development of analogical reasoning: Insights from scene analogy problems. *Journal of Experimental Child Psychology, 94*(3), 249–273. https://doi.org/10.1016/j.jecp.2006.02.002

Rogers, T. T., & McClelland, J. L. (2005). A parallel distributed processing approach to semantic cognition: Applications to conceptual development. In L. Gershkoff-Stowe & D. H. Rakison (Eds.), *Building object categories in developmental time* (pp. 335–387). Mahwah, NJ: Lawrence Earlbaum.

Ross, M. (1989). Relation of implicit theories to the construction of personal histories. *Psychological Review, 96*(2), 341.

Saffran, J. R., Pollak, S. D., Seibel, R. L., & Shkolnik, A. (2007). Dog is a dog is a dog: Infant rule learning is not specific to language. *Cognition, 105*(3), 669–680. https://doi.org/10.1016/j.cognition.2006.11.004

Sagi, E., Gentner, D., & Lovett, A. (2012). What difference reveals about similarity. *Cognitive Science, 36*(6), 1019–1050.

Shields, W. E., Smith, J. D., & Washburn, D. A. (1997). Uncertain responses by humans and Rhesus monkeys (Macaca mulatta) in a psychophysical same–different task. *Journal of Experimental Psychology: General, 126*(2), 147.

Smirnova, A., Zorina, Z., Obozova, T., & Wasserman, E. (2015). Crows spontaneously exhibit analogical reasoning. *Current Biology, 25*(2), 256–260.

Son, J. Y., Doumas, L. A. A., & Goldstone, R. L. (2010). When do words promote analogical transfer? *The Journal of Problem Solving, 3*(1), 4.

Spelke, E. S., Breinlinger, K., Macomber, J., & Jacobson, K. (1992). Origins of knowledge. *Psychological Review, 99*(4), 605.

Thibaut, J.-P., French, R., & Vezneva, M. (2010). The development of analogy making in children: Cognitive load and executive functions. *Journal of Experimental Child Psychology, 106*(1), 1–19. https://doi.org/10.1016/j.jecp.2010.01.001

Thompson, R. K. R., & Oden, D. L. (2000). Categorical perception and conceptual judgments by nonhuman primates: The paleological monkey and the analogical ape. *Cognitive Science, 24*(3), 363–396. https://doi.org/10.1016/s0364-0213(00)00029-x

Thompson, R. K. R., Oden, D. L., & Boysen, S. T. (1997). Language-naive chimpanzees (Pan troglodytes) judge relations between relations in a conceptual matching-to-sample task. *Journal of Experimental Psychology: Animal Behavior Processes, 23*(1), 31–43. https://doi.org/10.1037/0097-7403.23.1.31

Tomasello, M. (1992). *First verbs: A case study of early grammatical development*. New York: Cambridge University Press. https://doi.org/10.1017/CBO9780511527678

Tomasello, M. (2000). Do young children have adult syntactic competence? *Cognition, 74*(3), 209–253.

Truppa, V., Mortari, E. P., Garofoli, D., Privitera, S., & Visalberghi, E. (2011). Same/different concept learning by capuchin monkeys in matching-to-sample tasks. *PLoS One, 6*(8), e23809.

Tyrrell, D. J., Stauffer, L. B., & Snowman, L. G. (1991). Perception of abstract identity/difference relationships by infants. *Infant Behavior & Development, 14*(1), 125–129. https://doi.org/10.1016/0163-6383(91)90059-2

Wang, S.-H., & Baillargeon, R. (2008). Can infants be 'taught' to attend to a new physical variable in an event category? The case of height in covering events. *Cognitive Psychology, 56*(4), 284–326. https://doi.org/10.1016/j.cogpsych.2007.06.003

Wasserman, E. A., Castro, L., & Fagot, J. (2017). Relational thinking in animals and humans: From percepts to concepts. In J. Call (Ed.),. (Editor-in-Chief) *American Psychological Association Handbook of Comparative Cognition* (Vol. 2, pp. 359–384). Washington, DC: American Psychological Association.

Wasserman, E. A., & Young, M. E. (2010). Same–different discrimination: The keel and backbone of thought and reasoning. *Journal of Experimental Psychology: Animal Behavior Processes, 36*(1), 3.

Wasserman, E. A., Young, M. E., & Fagot, J. (2001). Effects of number of items on the baboon's discrimination of same from different visual displays. *Animal Cognition, 4*, 163–170.

Waxman, S., Fu, X., Arunachalam, S., Leddon, E., Geraghty, K., & Song, H. J. (2013). Are nouns learned before verbs? Infants provide insight into a long-standing debate. *Child Development Perspectives, 7*(3), 155–159.

Waxman, S. R., & Klibanoff, R. S. (2000). The role of comparison in the extension of novel adjectives. *Developmental Psychology, 36*(5), 571–581.

Wright, A. A., & Katz, J. S. (2006). Mechanisms of same/different concept learning in primates and avians. *Behavioural Processes, 72*(3), 234–254. https://doi.org/10.1016/j.beproc.2006.03.009

Xu, F., & Tenenbaum, J. B. (2007). Word learning as Bayesian inference. *Psychological Review, 114*(2), 245.

Young, M. E., & Wasserman, E. A. (1997). Entropy detection by pigeons: Response to mixed visual displays after same-different discrimination training. *Journal of Experimental Psychology: Animal Behavior Processes, 23*(2), 157–170. https://doi.org/10.1037/0097-7403.23.2.157

Young, M. E., & Wasserman, E. A. (2002). Detecting variety: What's so special about uniformity? *Journal of Experimental Psychology: General, 131*(1), 131–143. https://doi.org/10.1037/0096-3445.131.1.131

Zentall, T. R., Singer, R. A., & Miller, H. C. (2008). Matching-to-sample by pigeons: The dissociation of comparison choice frequency from the probability of reinforcement. *Behavioural Processes, 78*(2), 185–190. https://doi.org/10.1016/j.beproc.2008.01.015

Chapter 6
The Emergence of Inductive Reasoning During Infancy: Learning from Single and Multiple Exemplars

Susan A. Graham, Michelle S. Zepeda, and Ena Vukatana

Abstract In this chapter, we describe the emergence of category-based inductive reasoning during the infancy and preschool years, with focus on the adherence to a fundamental induction principle, *premise-conclusion similarity.* We review evidence demonstrating that 13- to 22-month-old infants and preschoolers use both category information and perceptual similarity to guide their inductive inferences about nonobvious properties under various conditions. Next, we describe recent studies from our lab focusing on 9- and 11-month-olds' tendency to associate properties with familiar and unfamiliar animal categories. These studies highlight the following: (1) infants as young as 9 months can link sound properties with animal categories and (2) the tendency to map properties to categories varies by the type of category (familiar vs. unfamiliar animals) and whether infants are familiarized with a single category member or multiple category members.

Inductive reasoning is a central aspect of human cognition, allowing individuals to generalize from the known to new instances and situations. Consider the following situation: a young child sees a brown four-legged animal walking toward them. If the child recognizes this animal belongs to the category *dog*, they can generate predictions about this particular animal and its characteristics and behaviors. That is, drawing upon their previous knowledge about dogs, the child might expect this

This research was supported by a Discovery Grant from the Natural Sciences and Engineering Research Council of Canada [194530-2011] and funding from the Canada Research Chairs Program, the Alberta Children's Hospital Foundation, and the Canada Foundation for Innovation awarded to SG. MZ was supported by the Canada Graduate Scholarship, and EV was supported by the Vanier Canada Graduate Scholarship. We thank the parents and infants who participated in the studies, the research assistants who helped us, and our collaborators, Suzanne Curtin and Nina Anderson. We are also grateful to Ka Wing Lai for her assistance in preparing this chapter and to Scot Parker for his help creating the stimuli.

S. A. Graham (✉) · M. S. Zepeda · E. Vukatana
University of Calgary, Calgary, AB, Canada
e-mail: grahams@ucalgary.ca

J. B. Childers (ed.), *Language and Concept Acquisition from Infancy Through Childhood*, https://doi.org/10.1007/978-3-030-35594-4_6

animal to be friendly, to bark, and to chase balls. In contrast, if the child identifies the animal as belonging to the category *bear*, their inferences about the animal and its characteristics and behaviors would likely be quite different! This example illustrates a type of induction that is pervasive in our everyday reasoning, namely, *category-based inductive reasoning*. Put simply, category-based inductive reasoning involves invoking the premise that properties of a category will likely hold true for other members of the same category. As the example also illustrates, category-based inductive reasoning allows individuals to move beyond representing specific entities (i.e., this particular dog or bear) to reasoning about these entities as instances of categories (e.g., *dogs*, *bears*), resulting in increased cognitive efficiency and the opportunity to benefit from past experiences.

In this chapter, we review research on the emergence of category-based inductive reasoning during early childhood. We begin with a discussion of preschoolers' inductive abilities to set the stage but then focus more on infancy and the developmental origins of inductive reasoning during the first year of life. In keeping with the overall theme of this volume, we weave through the chapter discussion of the role of multiple examples in inductive reasoning.

Inductive Reasoning During Early Childhood

Much is known about the developmental emergence of category-based inductive reasoning during the preschool and early childhood years (see Gelman, 2003; Hayes, 2007; Hayes & Heit, 2013; Kalish & Thevenow-Harrison, 2014, for reviews). To assess inductive generalizations in preschoolers, preschoolers are typically tested using a variant of the following paradigm: they are presented with a target object or objects (e.g., a bird) and learn about a nonobvious property of that object (e.g., "lives in a nest"). They are then asked whether that property extends to test objects that vary in some way from the target object. The resulting generalization patterns provide insight into the nature of children's categorical representations as well as the factors that influence their inferential decisions.

Seminal research has demonstrated that preschoolers have sophisticated inductive reasoning skills (e.g., Davidson & Gelman, 1990; Diesendruck & Peretz, 2013; Gelman, 1988; Gelman & Coley, 1990; Gelman & Davidson, 2013; Gelman & Markman, 1986; Gelman & O'Reilly, 1988; Rhodes & Gelman, 2008). They readily engage in category-based reasoning after exposure to only one exemplar of a given category and will flexibly adjust their property extensions as a function of the type of information provided. That is, in the absence of other cues to category membership, 2- to 5-year-olds will generalize properties to objects that are highly perceptually similar (e.g., Gelman & Coley, 1990; Graham, Booth, & Waxman, 2012; Noles & Gelman, 2012; Sloutsky & Fisher, 2004; Sloutsky, Kloos, & Fisher, 2007). When categorical information is provided in the form of shared category labels, however, preschoolers will overlook perceptual similarity and use the shared label information to guide their inferences (e.g., Booth, 2014; Gelman & Coley, 1990;

Gelman & Markman, 1987; Jaswal, 2004; Jaswal & Markman, 2002; Sweller & Hayes, 2014). For example, when target and test objects are labeled with the same count noun (a conventional marker of category membership; e.g., *bird*), children as young as 2½ years of age generalize properties from the target to new category members, even when those members differ in perceptual similarity from the target object (Gelman & Coley, 1990; Gelman & Davidson, 2013). Importantly, preschoolers recognize that not all labels signal shared category membership and default to reasoning on the basis of perceptual similarity when making their inferences in such cases (Gelman & Coley, 1990; Graham et al., 2012). For example, when target and test objects preschoolers are marked with adjectives (e.g., is sleepy), rather than count noun labels, 2.5-year-olds ignore these shared labels and revert to reasoning on the basis of shared perceptual similarity.

Not only will preschoolers overlook perceptual similarity when other category information is available, but they will also modulate their inferences based on factors such property generalizability, category homogeneity, and previous knowledge about a particular category (Gelman, 1988). For example, preschoolers recognize that some properties should be generalized, while others should not; and do not generalize those properties that are arbitrary (e.g., "fell on the floor this morning") or that reference transient properties (e.g., "hungry"; Brandone, 2017; Gelman, 1988; Graham, Cameron, & Welder, 2005; Graham, Welder, & McCrimmon, 2003; Waxman, Lynch, Casey, & Baer, 1997). Furthermore, preschoolers attend to category homogeneity, in that they prefer to generalize properties from a target to more homogeneous categories than to more heterogeneous categories (e.g., they are more likely to make property inferences from bears to mammals than to the more general category of animals; e.g., Brandone, 2017; Gelman, 1988; Lawson & Kalish, 2006). Finally, entities that are more typical or representative of a category are more likely to promote inductive inferences on the part of preschoolers (e.g., López, Gelman, Gutheil, & Smith, 1992; Rhodes, Brickman, & Gelman, 2008).

In summary, the literature reviewed briefly here suggests that many fundamental inductive phenomena are in place by early childhood (see Hayes, 2007, for a more detailed review). This does not imply, however, that inductive reasoning does not continue to develop over the childhood years. Consider, for example, the influence of evidence diversity on inductive generalizations. Adults tend to generalize more broadly from diverse samples (e.g., dogs and whales) than from less diverse sample (e.g., dogs and wolves; Feeney & Heit, 2011; Osherson, Wilkie, Smith, & Lopez, 1990). Although children as young as 5 years of age show evidence of attending to the diversity of the sample when reasoning about the properties of unfamiliar categories, they do not do so when reasoning about familiar natural kinds (Rhodes & Liebenson, 2015). In a similar vein, attention to sample size in inductive reasoning emerges by 3 years of age but only when there is a significant disparity in the sample sizes presented and items in a set are presented sequentially and not when presented simultaneously (Lawson, 2014). As these two examples illustrate, there is both stability in fundamental inductive processes from early childhood to adulthood and developmental changes that modulate when and how these processes are engaged.

Inductive Reasoning in Infancy

Given preschoolers' well-developed inductive reasoning skills, researchers have sought to characterize the emergence of inductive reasoning capacities during the infancy period. Due to infants' limited language abilities, inductive inference tasks relying on verbal responses are not feasible. Thus, researchers have drawn upon paradigms which capitalize on infants' tendency to imitate another's actions. Seminal research, using imitation paradigms, suggests that both perceptual similarity and category knowledge guide infants' inductive generalizations (e.g., Baldwin, Markman, & Melartin, 1993; Mandler & McDonough, 1996, 1998). For example, Baldwin et al. (1993) presented 9- to 16-month-olds with unfamiliar target objects that possessed a nonobvious property (e.g., a horn that honked when squeezed), demonstrated the action that elicited the property, and observed whether infants would perform the same action on test objects of varying degrees of similarity. Even after this brief exposure to the object, infants extended properties on the basis of perceptual features, performing the target action on those objects that were highly similar to the target object. Although no significant age effects were observed, Baldwin et al. noted that inductive inferences were more clearly exhibited by infants 11 months of age and older.

Building on this foundational work, we have conducted a number of studies examining infants' reasoning about the nonobvious properties of unfamiliar kinds, focusing on infants between the ages of 13 and 22 months. In our research, we have used a variant of the imitation paradigm (based on that of Baldwin et al., 1993; see Fig. 6.1). In this paradigm, infants are presented with unfamiliar target objects that possess a nonobvious sound property that is elicited by a particular action (e.g., rings when tapped on the top) across a series of trials. On each trial, the experimenter introduces the target object and demonstrates how the property can be evoked. Infants are then given the opportunity to explore the target object and elicit the property. The experimenter then presents the infant with a test object – depending on the particular trial, that test object may be highly perceptually similar to the target object (i.e., the high similarity object) or may be perceptually dissimilar in appearance (i.e., the low similarity object), and the infant is allowed to interact with that test object.

Within each object type (high or low similarity), infants are presented with trials in three within-subject conditions: the *violated-expectation* condition, the *baseline* condition, and the *predicted* condition. The condition of greatest interest is the *violated-expectation* condition; here, the target object possesses the nonobvious sound property, but the test objects do not. If infants judge the test object to belong to the same category as the target object, they will repeatedly attempt to elicit the nonobvious sound property on the disabled test object. In other words, infants' performance of actions on the test objects in condition provides evidence of inductive reasoning (i.e., they expect the test objects to possess the same nonobvious property as the target object) (Baldwin et al., 1993; Welder & Graham, 2001).

The *baseline* condition, in which the target and test object's nonobvious properties are disabled, acts as a control condition, providing a baseline measure of infants'

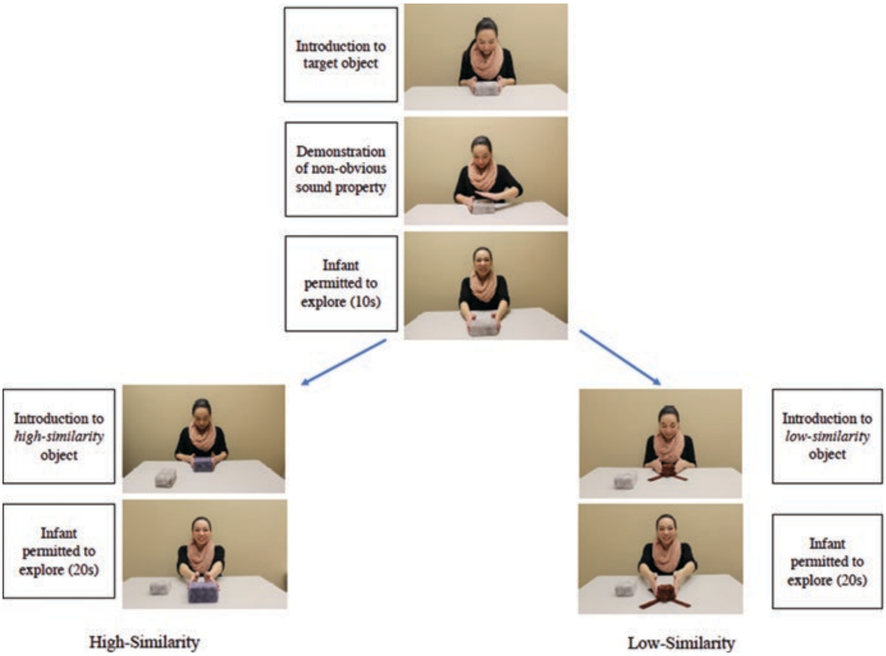


Fig. 6.1 Overview of imitation paradigm

exploratory actions. Comparing infants' performance of actions in the baseline condition to that of the violated-expectation condition indicates whether the property of the target object is indeed nonobvious. That is, if infants attempt to elicit the property from the test objects in the violated-expectation condition, but not the baseline condition, this provides evidence that the objects do not suggest the target action through their appearances alone. Results from the experiments described below confirm this prediction; infants perform few, if any, actions on the objects in the baseline condition.

The *predicted* condition, in which the target and test object both possess the nonobvious property, also acts as a control condition to ensure that infants do not develop the expectation that all test objects are disabled. We typically do not analyze infants' actions in this condition as it is difficult to judge which actions are due to the infants' expectations about shared properties and which actions are due to the reinforcing nature of the sound property itself (see Welder & Graham, 2001, for a more detailed discussion of these control conditions).

Across several studies using this imitation paradigm (e.g., Graham & Kilbreath, 2007; Keates & Graham, 2008; Switzer & Graham, 2017; Welder & Graham, 2001), we have focused on whether infants as young as 13 months of age, like adults and preschoolers, adhere to the fundamental induction principle of *premise-conclusion similarity* (Hayes, Heit, & Swendsen, 2010; Osherson et al., 1990). As noted earlier, according to this principle, the likelihood of generalizing from a premise or target category to a conclusion category varies as a function of the perceived similarity between the two. We focused our investigations on this principle as it has been identified as one of the "touchstone" phenomena in inductive reasoning, depending more on computations of similarity and less on background knowledge (Hayes & Heit, 2018). Given the young age of our participants, we reasoned that this principle may be the first to emerge in development. Across a series of studies examining this phenomenon, we have explicated the types of similarity infants will rely upon to guide their category-based inductive inferences, with particular focus on the role of shape similarity and category labels, as we review below. Note in the results reported below, we focus on findings from the violated-expectation condition, as this condition provides evidence of infants' inductive reasoning, as described above.

Shape similarity In a number of studies, we have documented that infants between 13 and 22 months of age rely on shape similarity across exemplars to guide their inductive inferences, in the absence of other information about object category (Graham, Keates, Vukatana, & Khu, 2013; Graham, Kilbreath, & Welder, 2004; Welder & Graham, 2001). That is, infants will privilege object shape over other types of perceptual properties (i.e., color) when engaging in inductive reasoning, assuming that similarly shaped objects share nonobvious properties. These results indicate that shape is taken as a reliable perceptual cue to the category membership of objects early in development, even prior to having acquired substantial vocabulary (Graham & Diesendruck, 2010; but see Colunga & Smith, 2008, for an alternative interpretation of the role of shape in early categories).

Why is shape privileged over other perceptual features in inductive reasoning tasks? First, shape (or form similarity) is easily and quickly perceived, even by the developing visual system (e.g., Quinn & Bhatt, 2015). Second, shape is often, but not perfectly, correlated with object category (i.e., bird-shaped things are often birds) and tends to not vary across category members to the same degree as other perceptual properties such as size and color (e.g., different-colored dogs and different-sized dogs). Third, infants' tendency to privilege shape similarity over color and texture similarity aligns with research demonstrating that shape is central to object segregation, object individuation, and object construals. Around 4 months of age, infants are more likely to use shape differences (vs. color or pattern differences) between objects to discover object boundaries (Needham, 1999) and to individuate objects (e.g., Wilcox, 1999). Around 9 months of age, infants use shape to recognize objects in occlusion events (e.g., Káldy & Leslie, 2003; Tremoulet, Leslie, & Hall, 2000; Xu, Carey, & Quint, 2004). This early-emerging primacy of shape may assist infants in later categorization and inductive reasoning tasks by directing their attention to shape, over other perceptual properties, when searching for commonalities across exemplars.

Category labels Naming objects with shared count nouns plays a critical role in infants' inductive reasoning, helping infants to unite objects into categories and guiding their inferences about the shared properties of category members (e.g., Graham et al., 2004; Graham & Kilbreath, 2007; Keates & Graham, 2008; Welder & Graham, 2001). That is, when the target and test objects are labeled with the same count noun, infants will reason that even dissimilar-shaped objects share a nonobvious property (see Chap. 11 for discussion of how relational labels can promote comparisons). Furthermore, by 16 months, infants have a refined understanding of nouns as category labels – only labels that are presented referentially (i.e., by a person vs. by a tape recorder), embedded within an intentional naming phrase (vs. presented alone), and marked as count nouns (e.g., "Look at this *blick.*" vs. marked as adjectives "Look. This is blickish.") guide infants' inferences about nonobvious properties (Keates & Graham, 2008).

More recently, we have focused on another role of nouns, that is, the use of nouns to sort highly similar objects into distinct categories. Consider, for example, a situation in which an infant sees both a blackbird and a bat. Both animals have similar shapes and sizes, both have wings, and both can fly – how do infants determine that these two animals belong to two different categories? One means by which accurate category membership is arrived at in cases like this is through category labels (i.e., providing different category labels). In two studies, we have shown that the ability to use distinct labels to carve out two different categories emerges between 14 and 16 months of age (Graham et al., 2013; Switzer & Graham, 2017; see Chap. 7 for discussion of how learning different verbs leads to a differentiation of verb meanings). That is, when target and test objects were labeled with different count nouns (e.g., "This is a blick. This is a wug."), 14- to 16-month-olds were significantly less likely to generalize the nonobvious property to the high-similarity test object in the violated-expectation condition, suggesting that they appreciated that the objects belonged to distinct categories and did not share nonobvious properties.

Together, these studies demonstrate infants' reliance on shared category labels to guide their inferences which reflects the expectation that count noun labels signal shared category membership and shared category membership promotes inductive inferences. Furthermore, when considered in conjunction with studies with preschoolers (e.g., Noles, 2019; Noles & Gelman, 2012), these results challenge accounts that children's reliance on shared names to license their inferences are solely the result of attentional or associative mechanisms (e.g., Sloutsky & Fisher, 2004). For example, if infants were relying only on an associative process, then they would have formed a unit that included the object category, label, and nonobvious property and used that unit to generalize to new category members. This unit presumably would include any type of label, be it a count noun, adjective, or isolated word (i.e., a word not presented in a naming phrase). Instead, our results indicate that infants only used the label when it was embedded in a naming phrase, marked as a count noun, and presented by the experimenter.

From pictures to objects We also have examined whether infants can draw inferences from symbolic artifacts, namely pictures, and generalize properties to real-world objects. The ability to draw inferences from symbolic artifacts allows infants to learn about objects without directly interacting with those objects. Our findings demonstrated that infants form expectations about nonobvious properties of objects after being exposed to a picture book that depicted the property of the object. That is, 13-, 15-, and 18-month-olds were read a picture book showing an adult evoking a nonobvious property of an unfamiliar object and then presented with the real-world objects depicted in the book. Infants in all three age groups imitated the actions of the depicted adult on the real-world objects, indicating that they expected the real-world objects to have same property as depicted in the book (Keates, Graham, & Ganea, 2014). This ability to transfer nonobvious properties from pictures to their real referents is present as early as 13 months of age and continues to be refined during the late infancy and preschool years (Keates et al., 2014; Khu, Graham, & Ganea, 2014).

Summary Together, this research highlights the flexible nature of infants' inductive reasoning strategies during the second year of life, demonstrating that infants will readily generalize properties to form a target object to category members based on both shape similarity and shared count noun labels. Moreover, infants have a refined understanding of the role of count nouns in inductive reasoning. That is, they appreciate that shared count nouns, but not shared adjectives, shared isolated word forms (e.g., "blick," or distinct nouns), license inductive inferences. Finally, infants can form expectations about shared object properties even when the objects are presented within different symbolic modalities, reasoning that real-world objects share the same properties as objects depicted in picture books.

Reasoning from single exemplars Infants' inductive reasoning abilities are even more remarkable when one considers that they are reasoning based upon exposure to a single exemplar of an unfamiliar category. That is, in all our experiments and in that of Baldwin et al. (1993), infants infer that two unfamiliar objects share the same nonobvious property after seeing the experimenter evoke the property on a single target object (in contrast, see Chap. 7 for a verb study showing difficulty at test when seeing only a single event). This indicates that infants between 13 and 22 months have robust inductive reasoning abilities and, like preschoolers, will generalize properties from one object to another after minimal experience with these unfamiliar object categories. Infants' ready willingness to generalize properties to other members of the same category following minimal exposure signals considerable developmental continuity in adherence to premise-conclusion similarity and raises the intriguing possibility that early in development, children assume, as a default, that properties can be generalized to other category members.

In considering these findings, we also note that the imitation task used provides a supportive context to explore infants' inferences. That is, infants are presented with relatively simple, yet engaging, objects in a highly interactive task that provides infants with support to identify the critical features of the objects and link the properties with the categories. During the demonstration phase, the experimenter engages the infant, directing them toward the objects using rich intentional cues. In interacting with the target exemplar themselves, infants gain direct experience in evoking the property from the object, and our observations of the infants in our studies suggest that they find producing the sound properties highly reinforcing. Perhaps more importantly, the target object remains in view during the test trials, thereby allowing infants to directly compare the given test object to the target object and assess the relation between the target and test objects. Thus, in the context of this highly interactive imitation task, infants between 13 and 22 months require only minimal support to generalize properties, findings that stand somewhat in contrast to the more fragile abilities of infants under the age of 12 months that we will review in the next section.

Developmental Origins of Inductive Reasoning

Given the remarkable developmental continuity in category-based inductive reasoning between the preschool and late infancy years, the question of when, and how, this ability emerges in development arises. One early study provides some evidence that inductive reasoning skills emerge during the first year of life (McDonough & Mandler, 1998). Using an imitation paradigm, McDonough and Mandler (1998) demonstrated that 9- and 11-month-olds generalized the properties of animals and vehicles to within-category members (e.g., infants extended the property of drinking to members of the animal category, such as a cat and a bird). Beyond this study (and that of Baldwin et al., 1993, who included infants ranging in age from 9 to 16 months), however, comparatively little is known about the emergence of inductive reasoning in infants younger than 12 months. Clearly, though, the ability to engage in even a rudimentary form of inductive reasoning would assist young infants in organizing the new information they encounter during their first year and help them make effective predictions about the properties of new entities.

In recent research, we have begun to address this developmental gap by examining a fundamental precursor of inductive reasoning – namely, the ability to create associations between categories and their respective properties. That is, we have examined infants' tendency to link object properties with object categories – an ability which precedes the inductive reasoning abilities observed in later developmental periods. This ability to link properties with categories, rather than individuals, can allow for the extension of properties to new members of a category. Note that we distinguish this process from inductive reasoning per se as the paradigm we use does not allow us to distinguish whether infants' generalizations of object properties reflect primarily associative processes (see Chaps. 2 and 4 for discussion of infant statistical learning) or whether they support a more sophisticated understanding of object categories on the part of the infant. In this line of research, we have sought to gain new insights into how infants begin to organize the new information they encounter during their first year and to understand the conditions that enable infants to make effective predictions about the properties of new entities that they may encounter.

In our studies, we have used a looking time paradigm rather than an imitation paradigm to allow us to examine a broader age range during early infancy (Baldwin et al., 1993; Welder & Graham, 2001). Relative to imitation paradigms, looking time paradigms place fewer demands on infants' motor and cognitive skills (i.e., by using infants' looking time as the primary outcome rather than motor skills). An overview of paradigm is illustrated in Fig. 6.2. In this paradigm, infants are presented with dynamic videos of two animals each paired with a novel distinctive sound across a

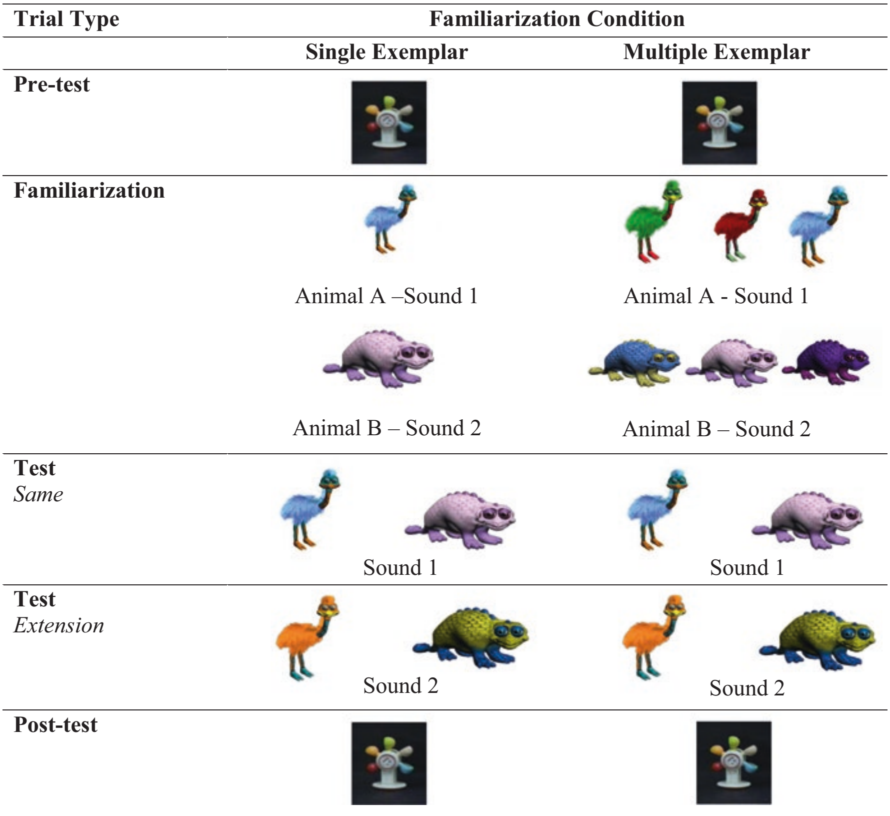


Fig. 6.2 Overview of looking time paradigm

series of familiarization trials, followed by test trials. As research has demonstrated that synchronous presentation of the movement of an object and a sound helps infants learn arbitrary relations between visual and auditory stimuli (e.g., Bahrick, Hernandez-Reif, & Flom, 2005; Gogate & Bahrick, 1998; Slater, Quinn, Brown, & Hayes, 1999), the onset of the sound is contingent upon the animal's mouth movement (i.e., the sound and mouth movement are synchronous). This synchrony also highlights for infants that the sound is an intrinsic property of the animal.

During the familiarization phase, infants are presented with an exemplar from each animal category (e.g., Animal A [blue] – Sound 1 – and Animal B [pink] – Sound 2) across a series of trials. On each trial, infants are presented with one exemplar from one of the two unfamiliar animal categories. The trial begins with an animal in profile, then turning its head to face the front toward the infant, with no sound. Next, the animal, with its head continuing to face toward the infant, opens its mouth and produces a sound, alternating between sound and silence at 1-second intervals. Then a new trial begins. The order of presentation of the two animal exemplars is randomized across this phase.

Following familiarization, we test infants' learning and extension of the properties: To evaluate the acquisition of the animal-sound mappings, infants are presented with side-by-side presentation of the two animals observed during familiarization, accompanied by one of the characteristic sounds (*same trials*). *Extension* trials assess infants' ability to extend the sound property to new category members that vary in similarity to the originally presented animals. Following from other research using preferential looking testing paradigms (e.g., Yoshida, Fennell, Swingley, & Werker, 2009), if infants learned the original animal-sound association, they will look toward the animal that matched the sound (i.e., the target animal) at rates greater than chance during the *same* trials. If infants generalized the sound property to new category members, their looking to the target animal will be significantly greater than chance on *extension* trials. Looking times in the pre- and posttest trials are compared to ensure infants maintain attention across the duration of the experiment.

Using this paradigm, we have made significant progress in documenting the early emergence of the foundations of inductive reasoning during the first year of life. That is, we have shown that 9- and 11-month-olds will learn and generalize the properties of both unfamiliar and familiar categories (Vukatana, Graham, Curtin, & Zepeda, 2015; Vukatana, Zepeda, Anderson, Curtin, & Graham, 2019). This ability to generalize properties, however, is modulated by a number of interacting factors, including the familiarity of the category, whether single or multiple exemplars of each category are presented during familiarization, and infants' age. We first review findings from this series of studies, highlighting the conditions under which 9- and 11-month-old infants learn and generalize object properties. We organize this discussion by contrasting our findings from unfamiliar versus familiar categories, followed by a broader discussion of when infants learn from single exemplars versus multiple exemplars.

Unfamiliar Animal Categories

In our first examination of infants' category-property links, we focused our investigations on whether infants would associate properties with unfamiliar "basic-level" animal categories and then extend those properties to new members of the categories (Vukatana et al., 2015; Zepeda & Graham, 2019). For these experiments, we developed novel animal categories that resembled basic-level categories (see Fig. 6.2). That is, following from the key attributes of basic-level categories noted by Rosch, Mervis, Gray, Johnson, and Boyes-Braem (1976), members of our unfamiliar animal categories shared salient features (i.e., texture, size, parts, shape) and were clearly discriminable (differed from one another in color). We focused on unfamiliar animal categories to provide insight into infants' ability to link properties with categories "online," following from our earlier research demonstrating that older infants can quickly make these links. That is, we used unfamiliar categories to ensure infants did not have any preexisting knowledge about the objects themselves that would influence their category-property mappings. Furthermore, we focused on "basic-like" categories, as previous research had only examined 9- and 11-month-olds' generalizations of the shared properties of familiar global categories (McDonough & Mandler, 1998; Pauen, 2002). In these studies, we first examined how the number of familiarization exemplars presented influenced whether infants will learn a novel animal-novel sound pairing and extend that sound property to new exemplars of a familiarized category, followed by consideration of whether providing category training would assist infants in learning from a single exemplar.

In our first set of experiments, we varied whether 11-month-olds were tested in one of two familiarization conditions: (a) single exemplar of each animal category or (b) multiple exemplars of each category (Vukatana et al., 2015). In the *single exemplar condition*, infants were presented with one exemplar of each of two animal categories (e.g., Animal A [blue] – Sound 1; Animal B [pink] – Sound 2) during familiarization. In the *multiple exemplar condition*, infants were familiarized with three different-colored exemplars from each animal category, with the members of the same category always producing the same sound. Infants in both conditions were then tested with *same* and *extension* trials, as outlined in the overview of our looking time paradigm in Fig. 6.2.

Infants' performance differed as a function of familiarization condition. That is, 11-month-old infants in the *single exemplar* condition did not show evidence of either (a) learning the original animal-sound mapping or (b) extending the sound property to a new member of the familiarized category. In contrast, infants in the *multiple exemplar condition* both learnt the animal-sound pairings and extended the property to new members of the relevant category. Together, these findings indicate that familiarization with one exemplar of a novel category is not sufficient to promote 11-month-olds' learning and extension of a sound property. Familiarization with multiple exemplars, however, facilitated both infants' tendency to form an animal-sound mapping and their tendency to extend the sound property to a new member of the category.

In a subsequent experiment in this series, we tested 9-month-olds in the same multiple exemplar condition presented to 11-month-olds (Vukatana et al., 2015). This experiment yielded interesting developmental differences with 9-month-old infants acquiring the unfamiliar animal-sound mappings but not extending the sound property to new exemplars of familiarized categories. Thus, at 9 months of age, multiple exemplars facilitated learning but did not promote extension of the newly learnt property. We will discuss the facilitative role of the presentation of multiple exemplars in later sections of the chapter.

In the next set of experiments, we continued our examination of the conditions that might facilitate infants' association of sound properties with unfamiliar animal categories. Here, we focused specifically on examining why generalizing from a single exemplar was so challenging for 11-month-olds, as described above (Vukatana et al., 2015). This question is particularly relevant as studies using imitation paradigms have demonstrated that infants of this same age will extend properties to a new category member when presented with one exemplar of a familiar category (i.e., McDonough & Mandler, 1998). One possible explanation for this discrepancy in findings may be related to the challenges associated with learning about an *unfamiliar* category (we will discuss this possibility further in later sections of the chapter).

To potentially increase infants' tendency to learn from a single category exemplar, we asked whether incorporating a training task at the beginning of the experiment would facilitate 11-month-olds' acquisition of unfamiliar animal-property associations based on exposure to a single category exemplar (Zepeda & Graham, 2019). In designing our training task, we drew upon studies demonstrating that presenting infants with familiar label-object pairs will shift their performance on a subsequent word-learning task (e.g., Fennell & Waxman, 2010; MacKenzie, Graham, Curtin, & Archer, 2014; May & Werker, 2014; Namy & Waxman, 2000; Vukatana, Curtin, & Graham, 2016). For example, MacKenzie et al. (2014) demonstrated that English-learning 12-month-olds will only map phonotactically illegal words to objects when they are provided with the referential-word training at the beginning of the task. Specifically, infants who were first presented with a series of familiar objects paired with their familiar word label subsequently mapped phonotactically illegal word forms to novel objects. In contrast, infants who were presented with familiar objects paired with exclamations (i.e., Ooh) did not map the illegal word forms to objects.

Drawing upon this research, we added a training phase to our previously developed looking time paradigm (see Fig. 6.3). First, we presented infants with stimuli of familiar animals, paired with their characteristic sounds (i.e., orange cat – meow; black dog – bark), at the beginning of the task. We chose cats and dogs as infants as young as 4 months reliably categorize stimuli these animals (Kovack-Lesh, Horst, & Oakes, 2008; Kovack-Lesh, Oakes, & McMurray, 2012; Quinn & Eimas, 1996). Our goal in presenting infants first with familiar animal-sound pairings was to move them into "a categorical mode," which would then assist them in learning information about new categories. Next, to orient infants to the nature of the *test* trials, infants were presented with two contrast-teaching trials. On these trials, infants saw the two familiar animals, side-by-side, accompanied by one of the sounds

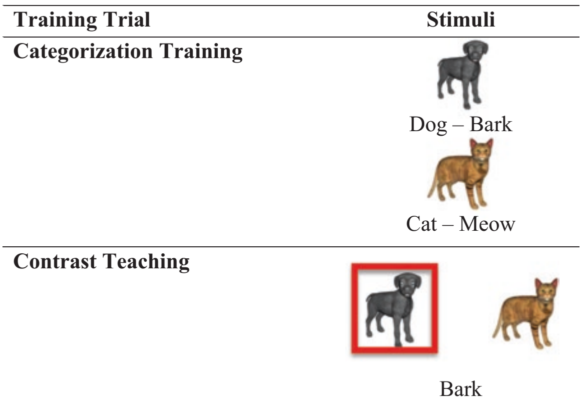


Fig. 6.3 Training trials in Zepeda and Graham (2019)

(e.g., orange cat, black dog; meow). On these trials, the target animal (e.g., orange cat) opened and closed its mouth and was enclosed by a colored frame, while the nontarget animal (e.g., black dog) kept its mouth closed. Following this training phase, infants were presented with two unfamiliar animal categories as in the single exemplar condition as described above (Vukatana et al., 2015).

Contrary to our expectations, 11-month-olds neither learnt the unfamiliar animal-sound associations nor generalized the property to a new member of the same category. Thus, the inclusion of training trials did not help infants develop animal-sound associations, when presented with a single exemplar of each category. In a second experiment, we sought to further reduce the demands of the task by integrating training throughout the familiarization phase, rather than placing it at the beginning of the task. That is, infants were presented with one familiar animal (i.e., a dog – barking) and one unfamiliar animal (i.e., an unfamiliar bird – an unfamiliar sound) during familiarization. We reasoned that integrating the familiar animal throughout familiarization and reducing the number of unfamiliar animal-sound pairings to be learnt would decrease the cognitive demands of the task and increase exposure to the familiar animal-sound pairings. In doing so, the number of training trials was increased (i.e., from 4 to 12), and thus, infants were given more opportunities to learn from by comparing the familiar and unfamiliar animals on sequential trials. The results from this experiment paralleled those from the first experiment – infants neither learnt nor generalized the unfamiliar animal-sound pairings.

Summary These findings demonstrate that 11-month-olds do not learn category-property mappings when familiarized with a single exemplar of an unfamiliar categories, even with specific training, a finding that stands in contrast to results from imitation-based studies with infants 13 months and older. These findings also highlight the facilitative role of multiple exemplars in learning and generalizing category-property links, although this effect interacted with age. That is, multiple exemplars facilitated both learning and generalization of unfamiliar animal properties for 11-month-olds but only facilitated the learning of the animal-sound mapping for 9-month-olds. In the next series of studies, we turn next to consideration of familiar categories to see if drawing upon naturally occurring categories would facilitate 9- and 11-month-olds' category-property links.

Familiar Animal Categories

In this next series of experiments, we investigated whether 9- and 11-month-olds will establish category-property links when presented with naturally occurring animal kinds (i.e., cats and dogs and their characteristic sounds). Our reasoning was as follows: First, infants readily distinguish and categorize cats and dogs (e.g., Behl-Chadha, 1996; Eimas & Quinn, 1994; Oakes & Ribar, 2005; Quinn, Eimas, & Rosenkrantz, 1993). Second, as noted above, McDonough and Mandler (1998) found that infants of this same age generalize the properties of familiar animals.

Finally, we reasoned that infants in our studies were likely to have some experience with cats and dogs, through direct exposure to pets or through toys, books, or other media. This intuition was confirmed; in all the experiments we report below, the majority of infants had some familiarity with cats and dogs through direct or indirect experience. Thus, we hypothesized that infants may have a preexisting representation of cats and dogs that they can recruit in our task to extend properties to new category members.

In this series, we again examined whether familiarization with single versus multiple exemplars influenced 9- and 11-month-olds' category-property generalizations (Vukatana et al., 2019; Vukatana, 2017). Here, we organize our presentation of these findings as a function of age group to highlight the developmental differences that emerged across these experiments.

Eleven-month-olds Here, we familiarized infants with single exemplars of two animal-sound pairings (one cat and one dog making their respective characteristic sounds) and examined their category-property mappings in one of two conditions – the *same-breed* condition and the *different-breed* condition (Vukatana et al., 2019; see Fig. 6.4 for the animals presented in each condition). On extension trials in the *same-breed* condition, infants saw new cat and dog exemplars differing only in color from exemplars presented during familiarization, allowing us to examine whether infants generalized properties to highly similar category members. On extension

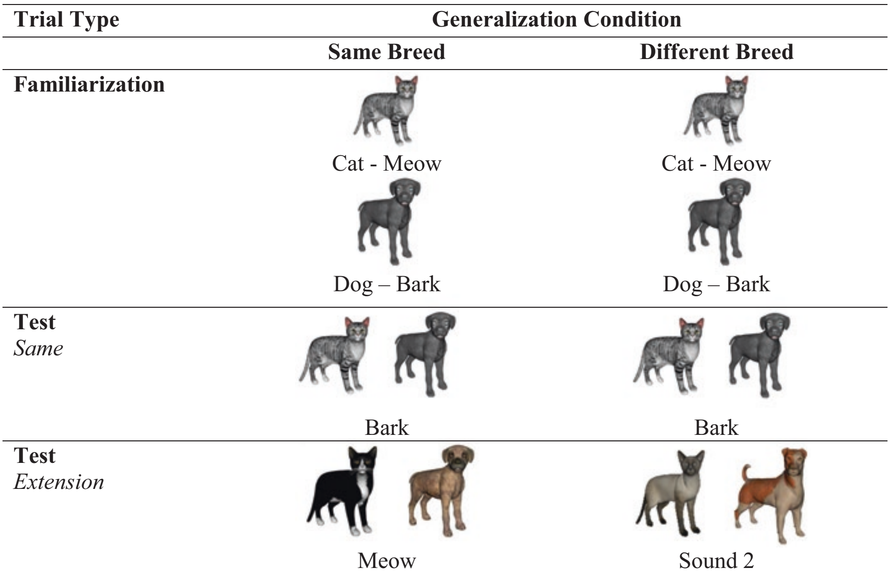


Fig. 6.4 Example of stimuli used in same- and different-breed conditions

trials in the *different-breed* condition, infants saw new, less perceptually similar category members – this condition allowed us to examine infants' ability to establish broader category-property links, as opposed to solely shaped-based associations.

When familiarized with single exemplars of familiar categories, 11-month-olds readily generalized characteristic sound properties of dogs and cats to both same- and different-breed category members. These findings contrast sharply with results from the unfamiliar animal studies reviewed above (Vukatana et al., 2015), suggesting that infants' *preexisting categorical representations* for dogs and cats may have facilitated their performance. In keeping with this proposal, 11-month-olds did not learn or generalize sound properties when familiarized with mismatched animal-sound pairings (i.e., dog meowing and cat barking) in a subsequent experiment. Together, this suggests that 11-month-olds have preexisting expectations about the links between the characteristic sound properties and the animal categories that facilitate their category-property mappings when presented with a single exemplar of each category during familiarization.

Nine-month-olds When presented with a single member of each familiar category during familiarization, 9-month-olds learnt the animal-sound link (Vukatana, 2017). Their generalization of this link, however, varied as a function of similarity between the familiarized animals and the new category member. That is, 9-month-olds generalized the sound property to highly similar category members (i.e., the same-breed animals), but not to the new category members of a different breed. In contrast, when presented with multiple exemplars of a category, 9-month-olds learnt and generalized the sound property to cats and dogs of a different breed.

Summary Nine- and eleven-month-olds readily generalized category-property mappings when familiarized with a single exemplar of a familiar category to highly similar category members. In contrast, when asked to generalize to new category members of a different breed, 9-month-olds, but not 11-month-olds, required familiarization with multiple exemplars. These findings also highlight the facilitative role of multiple exemplars in learning and generalizing category-property links, although this effect interacted with age.

Learning from One Versus Many: Integrating Findings Across Studies

We now turn to integrating our findings on the developmental origins of inductive reasoning, incorporating the interacting effects of number of exemplars, category type, and age. For ease of reference, we summarize the results from this series of studies on infants' tendency to form category-property links in Table 6.1.

In weaving these results together, we propose that the basic processes by which infants link object properties with object categories likely change little across the infancy and perhaps even the childhood years. That is, the research reviewed in this

Table 6.1 Summary of Vukatana (2017), Vukatana et al. (2015, 2019), and Zepeda and Graham (2019)

	Nine-month-olds		Eleven-month-olds	
	Learn	Generalize	Learn	Generalize
Unfamiliar animals				
Single exemplar	✓	✗	✗	✗
Multiple exemplars	✓	✗	✓	✓
Familiar animals				
Single exemplar	✓	✓[a]	✓	✓[b]
Multiple exemplars	✓	✓	—	—

[a]To same breed only
[b]To same and different breeds

chapter thus far suggests considerable developmental continuity in the fundamental ability to associate properties with categories. Instead, following from frameworks that have advocated for process-oriented approaches to developmental abilities, such as infant categorization (e.g., Madole & Oakes, 1999, 2003; Oakes & Madole, 2000) and speech perception (e.g., Werker & Curtin, 2005), we focus on the *conditions* that facilitate or hinder 9- to 11-month-old infants' tendency to form category-property associations, organizing the discussion around learning from single versus multiple exemplars.

Learning from single exemplars When presented with single exemplars of categories, 9- and 11-month-olds generalize properties to new category members both in our looking paradigms, as reviewed above, and in McDonough and Mandler's (1998) imitation studies but only when presented with familiar categories. Why can infants so readily generalize from familiar animals, after seeing a single representative of the category, and yet fail to even establish one-to-one mappings when the categories are unfamiliar? We consider both the demands of the learning task and changes in categorical representations with experience.

In the case of familiar animal categories, it is likely that infants activated their preexisting categorical representations for these stimuli (due to either direct or indirect exposure to these animals). As we noted earlier, the infants in our studies all had some direct or indirect experience with cats and dogs. The activation of such representations assists infants in making property extensions when presented with a single category exemplar and when asked to generalize to less perceptually similar category members of a different breed, in the case of 11-month-olds. Thus, by having broad exposure to these animals, infants may have been able to more easily able to attend to the category and to the relevant sound property, supporting their ability to establish category-property links. In keeping with the notion, research has found that infants can integrate information from their exposure to objects outside the lab to support categorization in early infancy (Bar-Haim, Ziv, Lamy, & Hodes, 2006; Bornstein & Mash, 2010; Hurley & Oakes, 2015).

In contrast, in the case of unfamiliar animal categories, infants were required to form category-property links *online* without drawing upon preexisting information or prior experiences with the stimuli. During familiarization, infants had to encode the perceptual characteristics of each animal and two unique sounds, track the relation between each animal and its' particular sound, and then retrieve this correct pairing, during the test trials in order to match the sound with the appropriate animal. Other research suggests that encoding these relations online in dynamic events may be difficult for infants under 12 months of age (Baumgartner & Oakes, 2011, 2013; Perone, Madole, & Oakes, 2011; Perone, Madole, Ross-Sheehy, Carey, & Oakes, 2008; Perone & Oakes, 2006). In conjunction with this research, our findings signal that infants' associative capacities may be challenged when learning about unfamiliar animals in dynamic events.

The contrast between infants' performance with familiar versus unfamiliar categories can also reflect changes in infants' categorization processes as a function of experience with a category. In particular, Quinn (2002) noted that infants' representations shift as they gain more experience with members of a category, moving from representing individual exemplars to developing more summary-type representations, to including representations of both category prototypes and individual exemplars. Thus, in the case of the unfamiliar categories, the need to represent each animal would have led to an increase in memory demands, resulting in infants' failing to form an animal-sound mapping in the first place. In contrast, with the familiar categories, infants likely activated their existing representations, allowing them to quickly link the sound property with the animal.

This account can also explain the differences between 9- and 11-month-olds' performance with the familiar categories – recall that 9-month-olds generalized the sound properties to the highly similar, but not dissimilar, animals, following familiarization to a single exemplar. That is, due to greater exposure to cats and dogs, 11-month-olds likely have more robust categorical representations for these animals and, more importantly, appear to access those representations during the experimental task. In contrast, 9-month-olds' ability to access their categorical representations may be in its emergent stages – thus, while a change in one feature (i.e., color) did not appear to disrupt their generalizations, greater perceptual variability between the familiarization and test exemplars did, leading infants to restrict their property extensions, when they had only experienced a single exemplar during and before the test trials.

Learning from multiple exemplars Presenting infants with multiple exemplars facilitated infant category-property mappings and generalizations under three conditions: (1) leading 11-month-olds to learn and generalize the properties of unfamiliar animals categories, (2) helping 9-month-olds to establish unfamiliar animal-sound mappings, and (3) leading 9-month-olds to generalize to less similar members of familiar animal categories. We consider these findings from two, not necessarily mutually exclusive perspectives.

The findings that familiarization with multiple exemplars promotes infants' property generalizations are consistent with an inductive reasoning phenomenon observed in adults and preschool-aged children, namely, that sample size promotes induction (*the sample size principle*; Gutheil & Gelman, 1997; Hayes & Kahn, 2005; Lo, Sides, Rozelle, & Osherson, 2002; Osherson et al., 1990). For example, Lawson (2014) demonstrated that 3-year-olds were more likely to generalize from larger sample than from smaller samples when items were presented sequentially (vs. simultaneously). In keeping with this notion, 9-month-olds' tendency to generalize properties of dogs and cats to less similar category members when presented with three exemplars, but not with one exemplar, during familiarization may reflect adherence to this principle. That is, familiarization with multiple exemplars may have highlighted the sound as a property of the broader category. This explanation is particularly fitting as 9-month-olds did both learn the animal-sound mappings and generalize to highly similar category members without the support of multiple exemplars. Thus, in this case, more exemplars of familiar categories, presented sequentially, led to broader generalizations.

An explanation based on adherence to the sample size principle alone, however, is less fitting when one considers the role of multiple exemplars in infants' unfamiliar category-property links. Recall that in these cases, 11-month-olds moved from neither learning animal-sound mappings nor generalizing properties with single exemplars to both learning and generalizing. Thus, it was not a matter of moving infants from a more restricted generalization pattern to a broader generalization pattern. Instead, we propose that presenting multiple exemplars engaged a comparative process, allowing infants to attend to the shared commonalities within a given category (see Chap. 5 for discussion of how comparisons can highlight deeper commonalities between exemplars). That is, multiple exemplars led infants to detect which features were most relevant to and shared within the category (i.e., shape and sound) and devote less attention to exemplar-specific features (i.e., color), facilitating the formation of a category and the generalization of the sound property to new category members.

An explanation based on comparison as a general learning mechanism is a consistent research in a number of domains in cognitive development (as illustrated beautifully in several chapters in this volume). With specific focus on categorization, presenting more than one exemplar leads preschoolers to detect commonalities among objects that would likely not otherwise be considered as a basis for categorical decisions (Gentner & Namy, 1999, 2004; Graham, Namy, Gentner, & Meagher, 2010; Namy & Gentner, 2002; Namy, Gentner, & Clepper, 2007). This process is evident in infants' categorical decisions as well – for example, Oakes and Ribar (2005) found that 4-month-olds differentiated between the categories of cats and dogs when familiarization items were presented in pairs, but not when presented sequentially, indicating that the opportunity to compare items simultaneously facilitated category formation and differentiation (Oakes & Ribar, 2005). By 6 months of age, however, infants formed exclusive categories of cats and dogs when the familiarization stimuli were presented sequentially. Quinn and Bhatt (2010) also demonstrated that exemplar variability across trials, but not within trials, facilitated 6- and 7-month-olds' formation of shapes. Akin the explanation we propose for our findings, Quinn and Bhatt propose that exemplar variability highlighted shared commonalities and drew infants' attention away from individual exemplar features.

Why, however, did multiple exemplars not facilitate 9-month-olds' generalization of unfamiliar animal properties? We posit that this finding reflects an interaction between the comparison process and 9-month-olds' cognitive processes, specifically memory. Research has demonstrated that visual short-term memory increases in memory capacity between 6.5 and 12 months of age (Rose, Feldman & Jankowski, 2001; Ross-Sheehy, Oakes, & Luck, 2003). Thus, it is possible that the 9-month-olds had difficulty attending to and/or encoding all the relevant features of the unfamiliar animals (e.g., shape, sound) when presented with multiple category exemplars. Consistent with this interpretation, it has been proposed that whether a cue available during learning "facilitates" categorization depends on the information-processing capabilities of infants at a given developmental stage (i.e., the demands of the task must match the capabilities of infants; Madole & Oakes, 1999, 2003; Oakes & Madole, 2000). Thus, changes in information-processing abilities (e.g., attention, memory) can account for older infants' ability to benefit from the presentation of multiple exemplars in property generalization tasks and 9-month-olds' failure to do so. Results from the familiar animal categories demonstrate that 9-month-olds can indeed establish category-property links. Thus, it is possible that with more support (e.g., longer familiarization phase, more exemplars), 9-month-olds would benefit from comparing multiple exemplars when presented with unfamiliar categories.

Other research has also shown that the facilitative role of comparison on infants' categorization is context-dependent (e.g., Kovack-Lesh & Oakes, 2007; Oakes & Ribar, 2005; Reznick & Kagan, 1983). For example, Kovack-Lesh and Oakes (2007) found that 10-month-olds distinguished categories of horses and dogs when familiarized with different pairs of items from the same category (e.g., golden retriever and black lab), but not when familiarized with pairs of identical items (e.g., two golden retrievers). In this study, infants were afforded the opportunity to compare members of a category; however, the facilitative effect of comparison only emerged when infants compared different members of the same category. Thus, these studies suggest that infants' ability to take advantage of information that can facilitate categorization (e.g., the opportunity to compare) varies with the context in which exemplars are presented.

Taken together, the findings from this series of studies document the conditions that promote learning and generalization of properties to category members between 9 and 11 months of age. In conjunction with other research, our results highlight that early category-property inferences engage highly dynamic and context-dependent processes.

Conclusions

We focused this chapter on a fundamental cognitive process that is endemic in human cognition, namely, category-based inductive reasoning. We began by tracing the development of this ability in early childhood through infancy, noting the developmental continuity across the preschool and late infancy years, particularly in the ability to privilege category information over perceptual similarity in guiding inductive reasoning. We next turned to a discussion on the origins of inductive reasoning, highlighting both 9- and 11-month-olds' abilities and limitations in generalizing the properties of both familiar and novel animal categories. The results of the studies reviewed in that section document that adherence to a fundamental inductive principle, *premise-conclusion similarity*, emerges as early as 9 months of age. Infants' adherence to this principle, however, varies as a function of age, category type, and the size of the sample (single vs. multiple exemplars), indicating that this ability reflects highly dynamic and context-dependent processes. Finally, returning to the focus of this volume on learning from multiple exemplars, we conclude that infants will generalize from a single exemplar, when they can engage an underlying categorical representation. When asked to form a novel category online and generalize properties to new category members, seeing more than one exemplar of a category paves the way.

References

Bahrick, L. E., Hernandez-Reif, M., & Flom, R. (2005). The development of infant learning about specific face-voice relations. *Developmental Psychology, 41*(3), 541–552. https://doi.org/10.1037/0012-1649.41.3.541

Baldwin, D. A., Markman, E. M., & Melartin, R. L. (1993). Infants' ability to draw inferences about nonobvious object properties: Evidence from exploratory play. Child Development, 64(3), 711–728. https://doi.org/10.1111/j.1467-8624.1993.tb02938.x

Bar-Haim, Y., Ziv, T., Lamy, D., & Hodes, R. M. (2006). Nature and nurture in own-race face processing. *Psychological Science, 17*(2), 159–163. https://doi.org/10.1111/j.1467-9280.2006.01679.x

Baumgartner, H. A., & Oakes, L. M. (2011). Infants' developing sensitivity to object function: Attention to features and feature correlations. *Journal of Cognition and Development, 12*(3), 275–298. https://doi.org/10.1080/15248372.2010.542217

Baumgartner, H. A., & Oakes, L. M. (2013). Investigating the relation between infants' manual activity with objects and their perception of dynamic events. *Infancy, 18*(6), 983–1006. https://doi.org/10.1111/infa.12009

Behl-Chadha, G. (1996). Basic-level and superordinate-like categorical representations in early infancy. *Cognition, 60*(2), 105–141. https://doi.org/10.1016/0010-0277(96)00706-8

Booth, A. E. (2014). Conceptually coherent categories support label-based inductive generalizations in preschoolers. *Journal of Experimental Child Psychology, 123*, 1–14. https://doi.org/10.1016/j.jecp.2014.01.007

Bornstein, M. H., & Mash, C. (2010). Experience-based and on-line categorization of objects in early infancy. *Child Development, 81*(3), 884–897. https://doi.org/10.1111/j.1467-8624.2010.01440.x

Brandone, A. C. (2017). Changes in beliefs about category homogeneity and variability across childhood. *Child Development, 88*(3), 846–866. https://doi.org/10.1111/cdev.1261

Colunga, E., & Smith, L. B. (2008). Knowledge embedded in process: The self-organization of skilled noun learning. *Developmental Science, 11*, 195–203.

Davidson, N. S., & Gelman, S. A. (1990). Inductions from novel categories: The role of language and conceptual structure. *Cognitive Development, 5*(2), 151–176. https://doi.org/10.1016/0885-2014(90)90024-N

Diesendruck, G., & Peretz, S. (2013). Domain differences in the weights of perceptual and conceptual information in children's categorization. *Developmental Psychology, 49*(12), 2383–2395. https://doi.org/10.1037/a0032049

Eimas, P. D., & Quinn, P. C. (1994). Studies on the formation of perceptually based basic-level categories in young infants. *Child Development, 65*(3), 903–917. https://doi.org/10.1111/j.1467-8624.1994.tb00792.x

Feeney, A., & Heit, E. (2011). Properties of the diversity effect in category-based inductive reasoning. *Thinking & Reasoning, 17*(2), 156–181. https://doi.org/10.1080/13546783.2011.566703

Fennell, C. T., & Waxman, S. R. (2010). What paradox? Referential cues allow for infant use of phonetic detail in word learning. *Child Development, 81*(5), 1376–1383. https://doi.org/10.1111/j.1467-8624.2010.01479.x

Gelman, S., & O'Reilly, A. W. (1988). Children's inductive inferences within superordinate categories: The role of language and category structure. *Child Development, 59*(4), 876–887. https://doi.org/10.2307/1130255

Gelman, S. A. (1988). The development of induction within natural kind and artifact categories. *Cognitive Psychology, 20*(1), 65–95. https://doi.org/10.1016/0010-0285(88)90025-4

Gelman, S. A. (2003). *The essential child: Origins of essentialism in everyday thought*. New York, NY: Oxford University Press.

Gelman, S. A., & Coley, J. D. (1990). The importance of knowing a dodo is a bird: Categories and inferences in 2-year-old children. *Developmental Psychology, 26*(5), 795–804. https://doi.org/10.1037/0012-1649.26.5.796

Gelman, S. A., & Davidson, N. S. (2013). Conceptual influences on category-based induction. *Cognitive Psychology, 66*(3), 327–353. https://doi.org/10.1016/j.cogpsych.2013.02.001

Gelman, S. A., & Markman, E. M. (1986). Categories and induction in young children. *Cognition, 23*(3), 183–209. https://doi.org/10.1016/0010-0277(86)90034-X

Gelman, S. A., & Markman, E. M. (1987). Young children's inductions from natural kinds: The role of categories and appearances. *Child Development, 58*(6), 1532–1541. https://doi.org/10.2307/1130693

Gentner, D., & Namy, L. L. (1999). Comparison in the development of categories. *Cognitive Development, 14*(4), 487–513. https://doi.org/10.1016/S0885-2014(99)00016-7

Gentner, D., & Namy, L. L. (2004). The role of comparison in children's early word learning. In D. G. Hall & S. R. Waxman (Eds.), *Weaving a lexicon* (pp. 533–568). Cambridge, MA: MIT Press.

Gogate, L. J., & Bahrick, L. E. (1998). Intersensory redundancy facilitates learning of arbitrary relations between vowel sounds and objects in seven-month-old infants. *Journal of Experimental Child Psychology, 69*(2), 133–149. https://doi.org/10.1006/jecp.1998.2438

Graham, S. A., Booth, A. E., & Waxman, S. R. (2012). Words are not merely features: Only consistently applied nouns guide 4-year-olds' inferences about object categories. *Language Learning and Development, 8*(2), 136–145. https://doi.org/10.1080/15475441.2011.599304

Graham, S. A., Cameron, C. L., & Welder, A. N. (2005). Preschoolers' extension of familiar adjectives. *Journal of Experimental Child Psychology, 91*(3), 205–226. https://doi.org/10.1016/j.jecp.2005.03.001

Graham, S. A., & Diesendruck, G. (2010). Fifteen-month-old infants attend to shape over other perceptual properties in an induction task. *Cognitive Development, 25*(2), 111–123. https://doi.org/10.1016/j.cogdev.2009.06.002

Graham, S. A., Keates, J., Vukatana, E., & Khu, M. (2013). Distinct labels attenuate 15-month-olds' attention to shape in an inductive inference task. *Frontiers in Psychology, 3*, 586. https://doi.org/10.3389/fpsyg.2012.00586

Graham, S. A., & Kilbreath, C. S. (2007). It's a sign of the kind: Gestures and words guide infants' inductive inferences. *Developmental Psychology, 43*(5), 1111–1123. https://doi.org/10.1037/0012-1649.43.5.1111

Graham, S. A., Kilbreath, C. S., & Welder, A. N. (2004). Thirteen-month-olds rely on shared labels and shape similarity for inductive inferences. *Child Development, 75*(2), 409–427. https://doi.org/10.1111/j.1467-8624.2004.00683.x

Graham, S. A., Namy, L. L., Gentner, D., & Meagher, K. (2010). The role of comparison in preschoolers' novel object categorization. *Journal of Experimental Child Psychology, 107*(3), 280–290. https://doi.org/10.1016/j.jecp.2010.04.017

Graham, S. A., Welder, A. N., & McCrimmon, A. W. (2003). Hot dogs and zavy cats: Preschoolers' and adults' expectations about familiar and novel adjectives. *Brain and Language, 84*(1), 16–37. https://doi.org/10.1016/S0093-934X(02)00518-7

Gutheil, G., & Gelman, S. A. (1997). Children's use of sample size and diversity information within basic-level categories. *Journal of Experimental Child Psychology, 64*(2), 159–174. https://doi.org/10.1006/jecp.1996.2344

Hayes, B. K. & Kahn, T. (2005, July). *Children's sensitivity to sample size in inductive reasoning*. Paper presented at the 14th Biennial Conference of the Australasian Human Development Association, Perth.

Hayes, B. K. (2007). The development of inductive reasoning. In A. Feeney & E. Heit (Eds.), *Inductive reasoning: Experimental, developmental, and computational approaches* (pp. 25–54). New York, NY: Cambridge University Press. https://doi.org/10.1017/CBO9780511619304.003

Hayes, B. K., & Heit, E. (2013). Induction. In E. Reisberg (Ed.), *The oxford handbook of cognitive psychology* (pp. 618–634). New York, NY: Oxford University Press. https://doi.org/10.1093/oxfordhb/9780195376746.013.0039

Hayes, B. K., & Heit, E. (2018). Inductive reasoning 2.0. *Wiley Interdisciplinary Reviews: Cognitive Science, 9*(3), e1459. https://doi.org/10.1002/wcs.1459

Hayes, B. K., Heit, E., & Swendsen, H. (2010). Inductive reasoning. *Wiley Interdisciplinary Reviews: Cognitive Science, 1*, 278–292. https://doi.org/10.1002/wcs.44

Hurley, K. B., & Oakes, L. M. (2015). Experience and distribution of attention: Pet exposure and infants' scanning of animal images. *Journal of Cognition and Development, 16*, 11–30. https://doi.org/10.1080/15248372.2013.833922

Jaswal, V. K. (2004). Don't believe everything you hear: Preschoolers' sensitivity to speaker intent in category induction. *Child Development, 75*(6), 1871–1885. https://doi.org/10.1111/j.1467-8624.2004.00822.x

Jaswal, V. K, & Markman, E. M. (2002). *Children's acceptance and use of unexpected category labels to draw non-obvious inferences*. Proceedings of the Annual Meeting of the Cognitive Science Society (Vol. 24).

Káldy, Z., & Leslie, A. M. (2003). Identification of objects in 9-month-old infants: Integrating 'what' and 'where' information. *Developmental Science, 6*(3), 360–373. https://doi.org/10.1111/1467-7687.00290

Kalish, C. W., & Thevenow-Harrison, J. T. (2014). Descriptive and inferential problems of induction: Toward a common framework. *Psychology of Learning and Motivation, 61*, 1–39. https://doi.org/10.1016/B978-0-12-800283-4.00001-0

Keates, J., & Graham, S. A. (2008). Category markers or attributes why do labels guide infants' inductive inferences? *Psychological Science, 19*(12), 1287–1293. https://doi.org/10.1111/j.1467-9280.2008.02237.x

Keates, J., Graham, S. A., & Ganea, P. A. (2014). Infants transfer nonobvious properties from pictures to real-world objects. *Journal of Experimental Child Psychology, 1*(125), 35–47. https://doi.org/10.1016/j.jecp.2014.02.003

Khu, M., Graham, S. A., & Ganea, P. A. (2014). Learning from picture books: Infants' use of naming information. *Frontiers in Psychology, 5*, 144. https://doi.org/10.3389/fpsyg.2014.00144

Kovack-Lesh, K. A., Horst, J. S., & Oakes, L. M. (2008). The cat is out of the bag: The joint influence of previous experience and looking behavior on infant categorization. *Infancy, 13*(4), 285–307. https://doi.org/10.1080/15250000802189428

Kovack-Lesh, K. A., & Oakes, L. M. (2007). Hold your horses: How exposure to different items influences infant categorization. *Journal of Experimental Child Psychology, 98*(2), 69–93. https://doi.org/10.1016/j.jecp.2007.05.001

Kovack-Lesh, K. A., Oakes, L. M., & McMurray, B. (2012). Contributions of attentional style and previous experience to 4-month-old infants' categorization. *Infancy, 17*(3), 324–338. https://doi.org/10.1111/j.1532-7078.2011.00073.x

Lawson, C. A. (2014). Three-year-olds obey the sample size principle of induction: The influence of evidence presentation and sample size disparity on young children's generalizations. *Journal of Experimental Child Psychology, 123*, 147–154. https://doi.org/10.1016/j.jecp.2013.12.004

Lawson, C. A., & Kalish, C. W. (2006). Inductive inferences across time and identity: Are category members more alike than single individuals? *Journal of Cognition and Development, 7*(2), 233–252. https://doi.org/10.1207/s15327647jcd0702_5

Lo, Y., Sides, A., Rozelle, J., & Osherson, D. (2002). Evidential diversity and premise probability in young children's inductive judgment. *Cognitive Science, 26*(2), 181–206. https://doi.org/10.1207/s15516709cog2602_2

López, A., Gelman, S. A., Gutheil, G., & Smith, E. E. (1992). The development of category-based induction. *Child Development, 63*(5), 1070–1090. https://doi.org/10.1111/j.1467-8624.1992.tb01681.x

MacKenzie, H. K., Graham, S. A., Curtin, S., & Archer, S. L. (2014). The flexibility of 12-month-olds' preferences for phonologically appropriate object labels. *Developmental Psychology, 50*(2), 422.

Madole, K. L., & Oakes, L. M. (1999). Making sense of infant categorization: Stable processes and changing representations. *Developmental Review, 19*, 263–296. https://doi.org/10.1006/drev.1998.0481

Madole, K. L., & Oakes, L. M. (2003). Principles of developmental change in infants' category formation. In D. H. Rakison & L. M. Oakes (Eds.), *Early concept and category development: Making sense of the blooming, buzzing confusion* (pp. 132–158). New York, NY: Oxford University Press.

Mandler, J. M., & McDonough, L. (1996). Drinking and driving don't mix: Inductive generalization in infancy. *Cognition, 59*, 307–335. https://doi.org/10.1016/0010-0277(95)00696-6

Mandler, J. M., & McDonough, L. (1998). Studies in inductive inference in infancy. *Cognitive Psychology, 37*, 60–69. https://doi.org/10.1006/cogp.1998.0691

May, L., & Werker, J. F. (2014). Can a click be a word? Infants' learning of non-native words. *Infancy, 19*(3), 281–300. https://doi.org/10.1111/infa.12048

McDonough, L., & Mandler, J. M. (1998). Inductive generalization in 9- and 11-month-olds. *Developmental Science, 1*, 227–232. https://doi.org/10.1111/1467-7687.00035

Namy, L. L., & Gentner, D. (2002). Making a silk purse out of sow's ears: Young children's use of comparison in category learning. *Journal of Experimental Psychology: General, 131*, 5–15. https://doi.org/10.1037/0096-3445.131.1.5

Namy, L. L., Gentner, D., & Clepper, L. E. (2007). How close is too close? Alignment and perceptual similarity in children's categorization. *Cognition, Brain, and Behavior, 11*, 647–659.

Namy, L. L., & Waxman, S. R. (2000). Naming and exclaiming: Infants' sensitivity to naming contexts. *Journal of Cognition and Development, 1*(4), 405–428. https://doi.org/10.1207/S15327647JCD0104_03

Needham, A. (1999). The role of shape in 4-month-old infants' object segregation. *Infant Behavior and Development, 22*(2), 161–178. https://doi.org/10.1016/S0163-6383(99)00008-9

Noles, N. S. (2019). Salience or centrality: Why do some features influence inductive generalization more than others? *Developmental Psychology, 55*(3), 612. https://doi.org/10.1037/dev0000652

Noles, N. S., & Gelman, S. A. (2012). Effects of categorical labels on similarity judgments: A critical analysis of similarity-based approaches. *Developmental Psychology, 48*(3), 890–896. https://doi.org/10.1037/a0026075

Oakes, L. M., & Madole, K. L. (2000). The future of infant categorization research: A process-oriented approach. *Child Development, 71*, 119–126. https://doi.org/10.1111/1467-8624.00125

Oakes, L. M., & Ribar, R. J. (2005). A comparison of infants' categorization in paired and successive presentation tasks. *Infancy, 7*, 85–98. https://doi.org/10.1207/s15327078in0701_7

Osherson, D. N., Wilkie, O., Smith, E. E., & Lopez, A. (1990). Category-based induction. *Psychological Review, 97*(2), 185–200.

Pauen, S. (2002). Evidence for knowledge–based category discrimination in infancy. *Child Development, 73*(4), 1016–1033.

Perone, S., Madole, K. L., & Oakes, L. M. (2011). Learning how actions function: The role of outcomes in infants' representation of events. *Infant Behavior and Development, 34*(2), 351–362. https://doi.org/10.1016/j.infbeh.2011.02.006

Perone, S., Madole, K. L., Ross-Sheehy, S., Carey, M., & Oakes, L. M. (2008). The relation between infants' activity with objects and attention to object appearance. *Developmental Psychology, 44*(5), 1242. https://doi.org/10.1037/0012-1649.44.5.1242

Perone, S., & Oakes, L. M. (2006). It clicks when it is rolled and it squeaks when it is squeezed: What 10-month-old infants learn about object function. *Child Development, 77*(6), 1608–1622. https://doi.org/10.1111/j.1467-8624.2006.00962.x

Quinn, P. C. (2002). Beyond prototypes: Asymmetries in infant categorization and what they teach us about the mechanisms guiding early knowledge acquisition. In R. V. Kail & H. W. Reese (Eds.), *Advances in child development and behavior* (pp. 161–193). San Diego, CA: Academic Press. https://doi.org/10.1016/S0065-2407(02)80054-4

Quinn, P. C., & Bhatt, R. S. (2010). Learning perceptual organization in infancy: The effect of simultaneous versus sequential variability experience. *Perception, 39*(6), 795–806. https://doi.org/10.1068/p6639

Quinn, P. C., & Bhatt, R. S. (2015). Development of perceptual organization in infancy. In *The Oxford handbook of perceptual organization* (pp. 691–712). Oxford, UK: Oxford University Press.

Quinn, P. C., & Eimas, P. D. (1996). Perceptual cues that permit categorical differentiation of animal species by infants. *Journal of Experimental Child Psychology, 63*(1), 189–211. https://doi.org/10.1006/jecp.1996.0047

Quinn, P. C., Eimas, P. D., & Rosenkrantz, S. L. (1993). Evidence for representations of perceptually similar natural categories by 3-month-old and 4-month-old infants. *Perception, 22*(4), 463–475. https://doi.org/10.1068/p220463

Reznick, J. S., & Kagan, J. (1983). Category detection in infancy. *Advances in Infancy Research, 2*, 79–111.

Rhodes, M., Brickman, D., & Gelman, S. A. (2008). Sample diversity and premise typicality in inductive reasoning: Evidence for developmental change. *Cognition, 108*(2), 543–556. https://doi.org/10.1016/j.cognition.2008.03.002

Rhodes, M., & Gelman, S. A. (2008). Categories influence predictions about individual consistency. *Child Development, 79*(5), 1270–1287. https://doi.org/10.1111/j.14678624.2008.01188.x

Rhodes, M., & Liebenson, P. (2015). Continuity and change in the development of category-based induction: The test case of diversity-based reasoning. *Cognitive Psychology, 82*, 74–95. https://doi.org/10.1016/j.cogpsych.2015.07.003

Rosch, E., Mervis, C. B., Gray, W. D., Johnson, D. M., & Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive Psychology, 8*(3), 382–439. https://doi.org/10.1016/0010-0285(76)90013-X

Rose, S. A., Feldman, J. F., & Jankowski, J. J. (2001). Visual short-term memory in the first year of life: Capacity and recency effects. *Developmental Psychology, 37*(4), 539.

Ross-Sheehy, S., Oakes, L. M., & Luck, S. J. (2003). The development of visual short-term memory capacity in infants. *Child Development, 74*(6), 1807–1822. https://doi.org/10.1046/j.1467-8624.2003.00639.x

Slater, A., Quinn, P. C., Brown, E., & Hayes, R. (1999). Intermodal perception at birth: Intersensory redundancy guides newborn infants' learning of arbitrary auditory-visual pairings. *Developmental Science, 2*(3), 333–338. https://doi.org/10.1111/1467-7687.00079

Sloutsky, V. M., & Fisher, A. V. (2004). Induction and categorization in young children: A similarity-based model. *Journal of Experimental Psychology: General, 133*(2), 166–188. https://doi.org/10.1037/0096-3445.133.2.166

Sloutsky, V. M., Kloos, H., & Fisher, A. V. (2007). When looks are everything: Appearance similarity versus kind information in early induction. *Psychological Science, 18*(2), 179–185. https://doi.org/10.1111/j.1467-9280.2007.01869.x

Sweller, N., & Hayes, B. K. (2014). The development of induction based on noun and feature labels. *Psychonomic Bulletin & Review, 21*, 1048–1055. https://doi.org/10.3758/s13423-013-0576-z

Switzer, J., & Graham, S. A. (2017). 14-to 16-month-olds attend to distinct labels in an inductive reasoning task. *Frontiers in Psychology, 8*, 609. https://doi.org/10.3389/fpsyg.2017.00609

Tremoulet, P. D., Leslie, A. M., & Hall, D. G. (2000). Infant individuation and identification of objects. *Cognitive Development, 15*(4), 499–522. https://doi.org/10.1016/S0885-2014(01)00038-7

Vukatana, E. (2017). *Infants' acquisition of category-property links: Exploring 7-to 11-month-olds' abilities* (Unpublished doctoral dissertation). University of Calgary, Calgary, Canada.

Vukatana, E., Curtin, S., & Graham, S. A. (2016). Infants' acceptance of phonotactically illegal word forms as object labels. *Journal of Child Language, 43*(6), 1400–1411. https://doi.org/10.1017/S0305000915000707

Vukatana, E., Graham, S. A., Curtin, S., & Zepeda, M. S. (2015). One is not enough: Multiple exemplars facilitate infants' generalizations of novel properties. *Infancy, 20*(5), 548–575. https://doi.org/10.1111/infa.12092

Vukatana, E., Zepeda, M. S., Anderson, N., Curtin, S., & Graham, S. A. (2019). *11-month-olds link sound properties with animal categories*. Manuscript in preparation.

Waxman, S. R., Lynch, E. B., Casey, K. L., & Baer, L. (1997). Setters and samoyeds: The emergence of subordinate level categories as a basis for inductive inference in preschool-age children. *Developmental Psychology, 33*(6), 1074–1090. https://doi.org/10.1037/0012-1649.33.6.1074

Welder, A. N., & Graham, S. A. (2001). The influence of shape similarity and shared labels on infants' inductive inferences about nonobvious object properties. *Child Development, 72*(6), 1653–1673. https://doi.org/10.1111/1467-8624.00371

Werker, J. F., & Curtin, S. (2005). PRIMIR: A developmental framework of infant speech processing. *Language Learning and Development, 1*(2), 197–234. https://doi.org/10.1080/15475441.2005.9684216

Wilcox, T. (1999). Object individuation: Infants' use of shape, size, pattern, and color. *Cognition, 72*(2), 125–166. https://doi.org/10.1016/S0010-0277(99)00035-9

Xu, F., Carey, S., & Quint, N. (2004). The emergence of kind-based object individuation in infancy. *Cognitive Psychology, 49*(2), 155–190. https://doi.org/10.1016/j.cogpsych.2004.01.001

Yoshida, K. A., Fennell, C. T., Swingley, D., & Werker, J. F. (2009). Fourteen-month-old infants learn similar-sounding words. *Developmental Science, 12*(3), 412–418. https://doi.org/10.1111/j.1467-7687.2008.00789.x

Zepeda, M. S., & Graham, S. A. (2019). Does category-training facilitate 11-month-olds' acquisition of unfamiliar category-property associations? *Infant Behavior and Development, 57*, 101380. https://doi.org/10.1016/j.infbeh.2019.101380

Chapter 7
Learning Individual Verbs and the Verb System: When Are Multiple Examples Helpful?

Mutsumi Imai and Jane B. Childers

Abstract This chapter focuses on the problem of verb learning, including learning the meaning of a single new verb and learning the verb system in a language. Verb learning occurs in three phases: finding the core of meaning, discovering dominant patterns in a language, and delineating boundaries between individual verbs. In the first phase, two types of perceptual similarity are shown to be useful—sound symbolism and object similarity. Children benefit from seeing high-similarity examples before low-similarity ones (progressive alignment), as well as from contrast. After describing how children may discover patterns within a language, we focus on how children learn a verb within an overall system by describing verbs for carrying/holding in Chinese. Children between 3 and 7 years produced fewer verbs than their mothers, better approximated adult verb meanings with age. MDS and INDSCAL analyses reveal they attended to the objects in the events and reveal three semantic islands of verb meaning. An entropy analysis shows that there is an early stage of verb learning in which input frequency is important and a later stage in which the degree of boundary overlap with other verbs affects their ease of acquisition. In sum, the chapter shows children's use of multiple exemplars for verb learning, using structure mapping as a theoretical framework, and addressing the whole of verb learning in development.

We thank our research labs, the children and families who gave of their time—without whom we would be unable to learn about the world—our departments, and our funding sources (NICHD 2R15 HD044447; MEXT/JSPS 26580078, 16H01928 and 18H05084).

M. Imai (✉)
Keio University, Tokyo, Japan
e-mail: imai@sfc.keio.ac.jp

J. B. Childers
Department of Psychology, Trinity University, San Antonio, TX, USA

J. B. Childers (ed.), *Language and Concept Acquisition from Infancy Through Childhood*, https://doi.org/10.1007/978-3-030-35594-4_7

Children's Challenge in Acquiring the Lexical System of Verbs

What does it take for children to learn language? Children need to learn individual words of course. Inspired by the philosopher W. V. O. Quine who characterized the process of word learning as a problem of induction, pioneering developmental psychologists set out to create a new research paradigm which rightly pointed out that children need to learn words by inferring their meanings on their own, and that their inferences need to be constrained because there are too many possible inductive generalizations one can make from a single instance (Carey, 1978; Gleitman, 1990; Markman, 1989; 1990).

Decades of work have been devoted to identify how children make inferences about the meanings of words they encounter for the first time, mostly focusing on strategies children use to learn nouns. Specifically, some researchers have proposed that biases or constraints help children identify the potential referent in the scene, such as the whole object or the mutual exclusivity bias (Markman & Hutchinson, 1984; Markman & Wachtel, 1988). Others have examined how children determine the *range* of generalizations they should make, proposing word learning biases including the taxonomic bias and the shape bias (Imai, Gentner, & Uchida, 1994; Landau, Smith, & Jones, 1988; Markman, 1989). The principle of contrast is an additional constraint proposed for word learning, though it is based on knowledge about the pragmatics of language (Clark, 1990). Still other researchers have highlighted extralinguistic information including children's use of the speaker's facial expressions, eye gaze, and the informativeness of the speaker's utterance (e.g., see Chap. 9).

When compared to these many cues that inform noun learning, cues that could be used for inferring a new verb's meaning are less abundant. For toddlers, inferring the meanings of verbs is much more challenging than it is for nouns for a number of reasons, and children need to overcome a plethora of problems to get to adult-like meanings of verbs. First, verbs are temporally dynamic and ephemeral. In real life, children observe various actions in continuous sequence rather than a single, isolated action. It is thus difficult to determine when the action denoted by the verb begins and when it ends (Gentner, 1982; Gentner & Boroditsky, 2001; Imai, Haryu, & Okada, 2005).

Children face even further challenges in determining the appropriate range of generalizations for a newly learned verb. That is, how should that verb be used in new situational contexts, sentence frames, or with new entities? The visual information available to children when they hear a verb is likely to be very rich. To be able to generalize the verb, children need to extract just those parts of the event that link to that verb's core semantic features which serve as a basis for generalization. For that purpose, they first need to find out which objects and relations in the scene are relevant for guiding their later verb generalizations. When children see events while hearing a verb, they often see multiple objects, actions, and relations which are unsegregated. They need to deduce which aspects of an action link to the core meaning of a specific verb, including which objects may be required as a part of that verb's meaning (e.g., a stapler for the verb "staple" in English) and which are optional.

Separating the invariant elements in a scene from the variable ones is necessary but not sufficient to get to the meaning of a verb. Some verbs denote a movement done in a particular manner (e.g., swagger), while other verbs denote a direction of motion (e.g., rise, enter), and for other verbs, the resulting state of the object is important (e.g., crush). Thus, children need to determine whether they should pay attention to manner, path/direction, or result of the action to extract the meaning of the new verb. There are substantial crosslinguistic differences in how these features are incorporated into verb meanings (Gentner, 1982; Talmy, 1975, 1985), and children who attend to common patterns in their language would benefit. For example, manner, but not path, tends to be incorporated in verb meanings in English, whereas the reverse pattern is true for Spanish. In Japanese, in addition to path, ground information (the property of the object or landscape through which the figure object moves) is likely to be included (Muehleisen & Imai, 1997). This means that understanding how likely it is that a particular semantic feature is a key element in verb meanings could be important for making inferences about the meanings of new verbs.

If the child is fortunate enough to find those semantic features that are likely to be incorporated in her ambient language, this is a great accomplishment, but it is not the end of her challenges. The lexicon is not merely an assembly of words each standing on its own; rather, it is a complexly structured system in which words are contrasted one another along multiple dimensions at multiple levels (Saji et al., 2011; cf. de Saussure, 1916/1983). Thus, in order to be able to use words according to the conventions of the adult speakers of the ambient language community, children need to know how a particular word differs from the other words that surround it in the same semantic domain. For example, in English, the meaning of the verb "walk" is understood in relation to the verbs like "run," "jump," and "crawl." Likewise, the understanding of the meaning of the verb "tear" requires learning how this word differs from the similar-meaning verbs such as "cut," "break," "rip," and "split".

Importantly, the way a particular semantic domain is divided into a set of verbs can also differ across languages (e.g., Bowerman, 1982; Gentner, 1982; Levin, 1993). For example, English divides the domain of human locomotion finely, contrasting motions with different manners. In contrast, Japanese distinguishes locomotion broadly with only several verbs. In another semantic domain, that is in the domain of carrying, Mandarin Chinese distinguishes the range of motions that English speakers simply call by the one verb "carrying" or "holding" with more than 20 different verbs, but it does not make the distinction that is critical for English, that is, whether the agent is moving or not moving. Because the boundary of a verb's meaning requires delineation with all neighboring verbs, children eventually need to learn all of the verbs in a given semantic domain and how the entire semantic domain is carved up by them to acquire the precise meaning of a single verb.

In this chapter, we will argue that the acquisition of a verb lexicon is grounded in several fundamental cognitive capacities that have been noted to be important for almost all other types of learning. These abilities include (but are not limited to) the ability to detect perceptual similarity, the ability to form categories, the ability to track statistical distributions, the ability to segment events, and the ability to read the intention of the conversational partner. However, we will also argue that these basic cognitive functions alone are not sufficient for learning of meanings of individual verbs, let alone for acquiring a mature verb lexicon in which verbs in the same semantic domain are meaningfully related and woven into a system. What is critical is the ability to use these core cognitive faculties to form some initial verb meanings, and then the willingness to continuously bootstrap and reorganize these meanings to build more abstract and adult-like representations.

This chapter explores how a bootstrapping process could help children move toward an adult-like representation of individual verbs, as well as how children build a connected system of verb representations. In so doing, we will highlight three phases that are particularly important for verb meaning acquisition: (1) extraction of the verbal core from other elements of an action event, (2) finding the dominant lexicalization pattern for the ambient language, and (3) delineating boundaries between similar-meaning verbs and constructing the lexical system for a semantic domain.

We begin in the next section (section "Bootstrapping from Perceptual to Relational Similarity in Extracting the Core of Verb Meanings") by exploring how children could use two types of perceptual similarity to bootstrap themselves into an abstract representation of verb meanings, including how they extract a verbal core from other elements of a scene using similarity they can perceive without much experience with verb learning. The first kind of similarity is similarity between sounds and meaning, which is called *sound symbolism* and is a form of resemblance between properties of a linguistic form and the sensorimotor and/or affective properties of referents (Imai & Kita, 2014; Perniss & Vigliocco, 2014). The other kind of similarity is the similarity of objects between the original action event and a new event to which the verb should be generalized; this type of similarity influences children's ability to compare and learn from multiple events.

In the section "An Additional Mechanism for Verb Learning: Contrast", we will explore how children find patterns dominant in their ambient language, moving from reliance on universally shared cues to language-specific linguistic features in their inference of meanings of verbs. In the section "Verb Meaning Acquisition Within the Constraints of the Lexical System", we will outline how children construct a semantic domain by delineating boundaries between verbs in a domain and explore the role of comparison and contrast for this process. In summary, although the problem of acquiring verbs is complex, we will propose ways children tackle this difficult problem using varied (mostly domain-general) strategies, including strategies linked to the comparison of examples (when forming initial verb meanings) and the comparison and contrast of individual verbs (when forming the verb lexicon).

Bootstrapping from Perceptual to Relational Similarity in Extracting the Core of Verb Meanings

Young children recruit constellations of cues—conceptual, social, pragmatic, and distributional—to constrain their inferences about word meanings (e.g., Clark, 1990; Hollich, Golinkoff, & Hirsh-Pasek, 2007; Imai & Gentner, 1997; Imai & Haryu, 2001; Tomasello & Barton, 1994; Tomasello & Kruger, 1992), but not all cues are available from the earliest stages of lexical development. For verb learning, it has been well established that children use the argument structure of the sentence when inferring verb meanings (Fisher, Gleitman, & Gleitman, 1991; Gleitman, 1990; Naigles, 1990), but it is not clear whether this knowledge is available for children to learn the *first* set of verbs or whether this knowledge is learned from experience with a language (e.g., through statistical learning; see, Chap. 4). Furthermore, although syntactic cues are helpful for mapping a new verb to a rough, macro level concept (e.g., whether it should be mapped to a caused motion or a spontaneous motion; Fisher, 1996; Lidz, Gleitman, & Gleitman, 2003; Naigles, 1990), they are not as useful for helping children to find the differences among words that appear in the same set of argument structures (e.g., walking vs. running).

Use of Multimodal Similarity (Iconicity)

What other cues are available for very young children? It is reasonable to think that a biologically endowed ability to map multimodal information—especially the ability to detect similarity between sound and vision, or sound symbolism—provides one such cue. For example, Ngas (an African language) includes the following verbs: su (“to run”)/su rututu (“to run making this sound”), jə (“to come”)/jə ɓulm (“to come moving like a python”), melp (“to shine”)/melp nkar-nkar (“to shine brightly”), and pye (“white”)/pye pwak-pwak (“really white”) (Don Burquest, personal communication).

To establish whether children attend to sound symbolism, Imai et al. (2008) and Kantartzis, Imai, and Kita (2011) tested whether Japanese- and English-speaking 3-year-olds could find a core meaning for a newly taught verb and generalize it to a new situation in which an actor is doing the same motion in three different conditions. In the experimental condition, the word was inherently similar in its sound and the type of motion to which it referred (i.e., it was a mimetic-based word). For example, in one study, a word, presented as a verb “choka-choka-shiteru” in Japanese and as “doing choka-choka” in English, was paired with a light and fast walking motion with small steps. In the first control condition, the mimetic-based nonsense word was paired with a different motion that did not match in a sound-symbolic way. In the second control condition, a nonsense word that resembled a typical monosyllabic verb in Japanese and English (e.g., *neke-tteiru* or *fepping*, respectively) was paired with the same motion from the experimental condition; as a nonmimetic word, it did not provide a sound-symbolic cue.

Consistent with results from previous studies (Golinkoff et al., 2002; Imai et al., 2005; Imai et al., 2008), in the two control conditions, both Japanese- and English-speaking 3-year-olds failed to generalize the newly taught verb to the identical action performed by a different actor (see also Maguire, Hirsh-Pasek, Golinkoff, and Brandone (2008) for related findings). However, in the experimental condition in which the novel verb sound-symbolically matched the action, not only Japanese 3-year-olds but also English-speaking 3-year-olds (who were not familiar with the sound-symbolic system of Japanese mimetics) were able to generalize the verb to a new event (Kantartzis et al., 2011; see also Yoshida (2012) for similar findings).

Thus, regardless of the language they were acquiring, these results show that children have the ability to detect inherent similarities between a word form and its meaning and use this information to extract the core of verb meanings from other elements involved in the scene, particularly objects. This differs from instances in which sound symbolism is not available. In these cases, children tend to have difficulty separating the object—either the agent or a patient object—from the action seen in a single scene (e.g., see Imai et al., 2005). They are able to overcome this tendency either by using sound symbolism cues (if available in their language), which could help children understand that objects are not as central the meaning of a particular verb, or through the comparison of multiple instances (a topic we will return to in sections "An Additional Mechanism for Verb Learning: Contrast" and "Verb Meaning Acquisition Within the Constraints of the Lexical System").

Use of Object Similarity

As discussed, children can use sound symbolism to identify the core meaning of the verb from the event scene (Imai et al., 2008; Kantartzis et al., 2011; Yoshida, 2012), which usually includes a set of objects that are important to the meaning of that new verb (i.e., either agent, patient, instrument, or other object); however, this cue may not be always available. As noted earlier, extracting the core meaning of a verb involves understanding relational similarities between different events, but this does not come naturally for young toddlers (Gentner & Kurtz, 2005; Gentner & Rattermann, 1991), and this is a central reason why young children have difficulty generalizing a newly learned verb to a scene with the same action but different objects (including agents). Providing evidence for this assertion, Imai and colleagues (Imai et al., 2005, 2008) tested 3- and 5-year-old children from three different language groups—Japanese, Chinese, and English. Six sets of video action events served as stimulus materials. Each set consisted of a standard event and two test events. In each standard event, a young woman was shown doing a novel action with a novel object. The two test events were variants of the standard event. In one, the same person was doing the same action with a different object (Action-Same-Object-Change, henceforth AS). In the other, the person was doing a different action with the same object (Action-Change-Object-Same, henceforth OS). While watching the standard event, a child heard a novel verb. The child was then shown the two test videos and was asked to which event the target word should be extended. In this study, if children understand that a verb maps to an action, and that the agent and the object of the action event can be changed across different instances of the event referred to by the verb, they should select the AS event to extend the novel verb.

Across languages, 3-year-olds performed at chance, while 5-year-olds could reliably extend a novel verb to an event involving the same action but a different object. Thus, by 5 years, children in each of the three languages were successful in generalizing verbs by action following a single exposure of the verb in the learning phase. Importantly, structural alignment theory predicts the comparison of multiple examples leads observers to attend to relations or highlight relations (see Chap. 5), and this is the process younger children may need to overcome an attention to objects during verb learning. Additionally, having objects which are higher in similarity between a base event and a target event can help children to successfully compare examples to each other (e.g., Gentner, Loewenstein, & Hung, 2007) or to notice the relational similarity between the two events (Gentner & Rattermann, 1991). This theory maintains that children (and adults) acquire abstract relational concepts by aligning elements across the source and target domains (e.g., events, scenes) hierarchically starting with concrete, directly perceptible elements (objects). The alignment of these elements across instances leads to the alignment of similar abstract relations between elements in the two instances, which drives attention to common relations (see Gentner, 2010). Similarity between the source and the target makes the alignment of the structures easier, and experience with high-similarity comparisons helps naïve observers learn how to align more varied examples (which is called "progressive alignment" in this theory). Building on this prediction that high similarity is useful, we expected that, when overall perceptual similarity between the source event (to which the verb is given) and the target event (to which the verb should be generalized) is increased, children would be better able to align a single example of an event linked to a verb with a new event and that this would help children grasp the core meaning of a new verb.

To examine this possibility, Haryu, Imai, and Okada (2011) manipulated object similarity and tested Japanese-speaking 3- and 4-year-olds on a verb generalization task. In Study 1, the structure of the stimuli was exactly the same as the previous studies by Imai et al. (2005, 2008), but here there were two conditions—the similar object condition and the dissimilar object condition. Consistent with the prediction from structural alignment theory, children were successful in generalizing the verb to the event containing the same action in the similar object condition (e.g., objects were similar in shape and size) but not in the dissimilar object condition, although the benefit of object similarity was less strong for 3-year-olds as compared to 4-year-olds because the attention to objects was very strong in 3-year-olds.

To investigate whether experience with high-similarity comparisons could help children with low-similarity mappings (as predicted by the theory), in Study 2, some children were given experience with high-similarity pairs. Specifically, in this study, Japanese-speaking 4-year-olds were assigned to either a similar-dissimilar (SD) condition or a dissimilar-dissimilar (DD) condition. Children in both conditions received exactly the same sets of single learning trial + test trial for the last four trials with exactly the same instructions and the procedure. However, the two conditions differed in the first four trials. In those first four trials, children in the SD condition experienced pairs of learning trials and test trials with similar objects, while those in the DD conditions experienced trials with dissimilar objects. Consistent with the object similarity bootstrapping hypothesis (or "progressive alignment" in structural alignment theory; e.g., Gentner, 2010), after having successfully generalized verbs to the same action event, the children in the SD condition were able to generalize the verbs even when the object in the target scene was perceptually dissimilar from that in the original scene.

A related set of studies by Childers et al. (2016) also shows 2½- and 3½-year-old children benefit from progressive alignment (or high object similarity) experience when learning new verbs. In their first study, children were shown live events in which the experimenter used objects to demonstrate a pair of similar events and then a pair of more varied ones (progressive alignment condition). This condition was compared to a condition in which children saw two pairs of varied events and to a control condition in which children saw a single repeated event during the learning phase. At test, all children were asked to enact the event using new objects that were similar to the prior ones (i.e., close extension trial) and were more varied (far extension trial). Results showed that the influence of the high-similarity comparisons was most prominent for the test trials with varied objects (far extension trials). In especially these trials, children with the high-similarity experience extended the verbs and differed from the control group whereas children in the all varied condition did not.

A second study used video events and an eye tracker. Children saw split screen pairs of events with some children initially seeing a similar pair of events than a more varied one (progressive alignment or PA condition) while others seeing two pairs of varied events. Again, children saw the same pair of events before test, and their looking during this pair was tracked using the eye tracker. Eye tracking results showed increases in looking to two important objects (agents and affected objects) in the events in the PA condition only in the 2½-year-old age group. At test, there were successful verb extensions by 3½-year-olds only in the PA condition. Thus, this second study showed that children's looking patterns during learning were influenced only when there was an initial high-similarity pair and that benefits of this similarity first experience at test were seen by 3½ years. Importantly, these two studies with different procedures, as well as the studies by Haryu et al. (2011), provide converging support for structural alignment theory. More specifically, Haryu et al. (2011) show that experience of high object similarity between one learning event and test helps to boost 4-year-olds' ability to be successful when there is low object similarity between a learning event and a new test event. Childers et al. (2016) add to this finding by showing that experience with a pair of events during learning that are more similar to each other helps children produce more creative enactments with more varied objects (as young as 2½ years) and extend the new verb to new video events by 3½ years. Thus, as noted, these results from two different labs, using different types of events and stimuli with different methods, provide converging results that object similarity is an important cue for children to bootstrap their way into more successful comparisons of varied events and/or more varied verb extensions.

Preliminary data from a new study from the Childers' lab also is beginning to show that children in this age range benefit from progressive alignment experience even when events to be compared are separated in time. In this study, using video stimuli shown on an iPad, children in a PA condition with 1-minute delays between events in the learning phase are successfully extending new verbs to new events at test, while children in an all varied condition are not. This new study is important because, in everyday life, relevant events do not usually follow each other in time.

Summary and Implications

In this section, we have discussed children's use of similarity to extract a core of verb meaning, reviewing two kinds of similarity: iconic similarity carried in the sound of the word and object similarity. Both types of similarity are perceptible and easily accessible, but each works somewhat differently. The iconicity between word form and meaning helps children realize that it is the manner of the motion that the novel word denotes because the sound of the word is inherently "similar" (iconic) to the manner of the motion. In contrast, in the second case, object similarity invites children to align events, helps them successfully perform the correct alignments across the examples, and helps them deduce how a new event fits with an existing verb meaning (from learning to test trials). In both cases, children show an ability to use perceptual similarity across instances as a way to get to relational sameness. When this experience is repeated, it can help children discover that a critical core of verb meanings is actions or relations, apart from specific objects. This insight then can help them move toward the ability to generalize verbs without needing to rely on similarity. These two processes help children *learn how to learn a verb*. In this sense, they should be considered driving forces for the acquisition of abstract meanings of individual verbs.

An Additional Mechanism for Verb Learning: Contrast

One more example-based mechanism to consider that differs from the two mechanisms above is children's attention to contrastive information. Contrast provides another way for children to get past object similarity. Children learning verbs can have access to two types of contrast information: explicit contrast (e.g., "That's not x-ing. It's y-ing") and implicit contrast (e.g., "She's x-ing and they're y-ing."); in everyday life, implicit contrast information is likely more common than are explicit statements. Contrast could help children learn a verb because it helps children narrow the range of meaning of one verb based on the scope of meaning of another verb or helps children delineate individual verb meanings before they have a complete understanding of the entire "verb neighborhood." Very few prior studies have included contrastive statements in a verb learning task, and thus more studies are needed in this area (see Au & Markman (1987), Au & Laframboise (1990), Haryu & Imai (2002), Waxman & Klibanoff (2000), Booth & Waxman (2009) for studies of contrast in noun and adjective learning).

One study that did include contrast information in a verb-learning task is Waxman, Lidz, Braun, and Lavin (2009). In their study, 24-month-olds were shown dynamic video events with an agent performing an action on an inanimate patient (e.g., a man waving a balloon). During the learning phase, four events were shown that could be compared (i.e., the man was shown waving four different balloons with varying shapes), and children in a verb condition heard sentences containing new verbs (e.g., "Look! The man is larping a balloon"). (There was also a noun condition and a no-word control.) Next, two more events were shown, providing contrastive information ("Uh-oh, he's not larping that"). Then children saw a randomly chosen repeated event from the first four trials one more time and heard a positive sentence (e.g., "Yay, he's larping that").

At test, children saw two scenes simultaneously. In Study 1, the test trial pairs showed a familiar action (the man waving a balloon) vs. a novel action with the same object (e.g., the man was tapping the balloon). Children in the verb condition preferred the familiar test trial even though at baseline, they preferred the novel test trial. A weakness in these data is that this looking did not exceed 50%, though the researchers argue the switch from baseline pattern is important (and was found only in the verb condition).

In a second study, the test trial was changed to show the same familiar test scene (e.g., man waving a balloon) vs. a new test scene of the same action but with a different object (e.g., man waving a rake). In some ways, this is a more awkward verb learning study as the action is the same in both test trials; however, the authors predicted that children in the verb condition could then maintain their looking to the novel test event they prefer at baseline (if they can overlook the object seen in the familiar test scene) while children in the noun condition looked longer at the familiar test event. The results were overall consistent with the authors' predictions.

As a whole, focusing on the verb learning condition in these two studies in this paper, results suggest that 2-year-olds benefit from four comparison trials, an explicit contrast trial, and a repeated comparison trial when learning a new verb. In other words, we can make 2-year-olds map a novel verb to an action and generalize it to a scene that involves the same action with a different object if we provide a maximum scaffolding by giving cues comparison and contrast. However, although the results are impressive at this young age, the exact mechanism by which it is effective is unclear. Could children perform similarly without the contrast trial? Do they need the reminder before test with the familiar event? Possibly not, though it is unclear.

Roseberry, Hirsh-Pasek, Parish-Morris, and Golinkoff (2009) also conducted a verb learning study that included implicit contrast in one of the test trials and used looking time as the dependent variable. In that study, 2- and 3-year-olds who heard a novel verb (e.g., "Look at Cookie Monster wezzling!"), in one of the four test trials, heard a different novel verb (implicit contrast statement) ("Where is glorping?"). On that trial, they looked equally to the new event depicting the action that went with the newly learned verb and a new event depicting an action they had not seen. This suggests that they were beginning to understand that hearing a different new verb should lead them to seek out a new event, but if they had looked longest at the newest event, their looking behavior would have been more convincing.

Saji et al. (2011) also included implicit contrast into their study of verbs linked to the actions of holding and carrying in Chinese (we provide more details about this study in the next section). One important finding was that the understanding and use of varied verbs in Chinese for these actions develops between 3 and 7 years and becomes more adult-like with age. One way the organization of verb meanings changes is children learn or refine the meaning of related verbs over time, or through a process of implicit contrast.

Childers, Hirshkowitz, and Benavides (2014) also conducted a set of studies that tested children's verb learning when given explicit and implicit contrast statements. In Study 1, 3½-year-olds participated in an explicit contrast condition, implicit contrast condition, or control. In the experimental session, the experimenter showed children two different actions using the same objects. In the explicit contrast condition, they heard a positive statement for one (e.g., "Look! I'm meeking it!") and a negative statement for the other ("Look! I'm not meeking it!"). In the implicit contrast condition, they heard a different verb for each event (e.g., "Look! I'm meeking it" and "Look! I'm koobing it"), and in the control, they heard general sentences while seeing both events (e.g., "Look what I can do!" and "Now look what I can do"). Each action and sentence was repeated once, and at test, children were given the objects and asked to perform the action ("Can you show me meeking?" or, in the control, "Can you do it?"). Events were counterbalanced so that a particular event was correct for only half of the participants.

Results showed that both contrast conditions differed from the control condition and that they did not differ from each other. Children in the implicit contrast group differed significantly from the control group responses and from chance, while children in the explicit contrast group tended to differ from the control group but also differed significantly from chance. Thus, this first study showed that 3½-year-olds benefited from implicit and explicit contrast statements, and perhaps from comparison as they had access to repeated events, in initially learning a new verb.

Study 2 extended these results by including a new set of objects during the learning phase. The same test objects were included (to be able to compare results with Study 1), but in this study, there were new objects included in the learning sets, and thus children needed to extend the verbs at test. Children participated in an explicit contrast condition or a control condition; in the control, children heard a single verb and only saw that event with that verb repeated (so had no contrastive information). Children in the explicit contrast condition were more successful at extending the new verbs than were children in the control condition, but children in both conditions struggled, performing at rates lower than chance. Thus, they benefitted from contrast but had trouble extending a new verb to new objects under these conditions. (This fits with a prior study by Namy & Clepper (2010), which showed that comparison may be better than contrast for learning novel superordinate terms.)

In a third study, 2½-, 3½-, and 4½-year-olds were shown events that could be compared and contrasted. Events were shown in pairs, but pairs differed over the three trials in the learning phase. Thus, children heard positive and negative statements about actions within a pair with a single novel verb (e.g., "koobing" and "not koobing") and then saw a new pair of actions that could be compared with the prior pair and that also included positive and negative explicit contrast statements. At test, 2½-year-olds performed at rates that were significantly below chance, 3½-year-olds performed at chance—but also performed significantly better than did 3½-year-olds in Study 2 who only had contrastive information—and 4½-year-olds succeeded. Additionally, although 3½-year-olds as a group were performing at chance, as individuals, the number of children who were able to be consistently successful exceeded chance.

In comparing this result to the prior result in the Waxman, Lidz et al. study, here children had to process interleaved trials of comparison and contrast. Namy and Clepper (2010) showed that contrast was most useful in a categorization task when it followed comparison trials, as in the Waxman, Lidz et al. study. This seems to provide us with important insights when considering how children learn verbs in everyday contexts. It may be that children need to use comparison to learn a new verb first, including linking that new verb to actions they have seen without extending it to new events, and then later, as other verbs are learned, they are able to use more contrastive information to build their lexicon (as in Saji et al. (2011) described later in this chapter).

Verb Meaning Acquisition Within the Constraints of the Lexical System

Although children's use of contrast begins to address the question of how children infer the meanings of verbs with some consideration of other verbs, in this section, we will consider how children build up a lexicon more deeply. Segregating the action from other elements of the event scene and learning individual verbs are important achievements, but even when children can do this, they still have a long way to go to develop the mature lexicon adults possess.

Complexity of the Semantic Structures in Lexical Domains in the Real World

In the real world, the lexicon is an extremely complex system, in which every word form in a language expresses a unique meaning (Bolinger, 1977; Clark, 1990; Lyons, 1963), and the meaning of any particular word depends on how the word is related to other similar words (Aichison, 1987; de Saussure, 1916/1983; Lyons, 1977; Miller & Johnson-Laird, 1976). Furthermore, the lexicon consists of structured subsets (Cruse, 1986; Fillmore, 1982; Fillmore & Atkins, 1992; Levin, 1993; Pustejovsky, 1995), whereby words are contrasted with one another along different semantic dimensions (cf. de Saussure, 1916/1983). Each language has a typical conflation pattern, but there is always a set of verbs that do not follow the typical pattern. Thus, although it is still helpful to consider language-specific conflation patterns, when appropriate, children need to be flexible in using it to accommodate verbs that do not conform that pattern; they need to figure out how a particular word differs from the other similar-meaning words.

Let us return to our discussion of the verbs that we call carrying/holding in Mandarin Chinese, for example. This language makes fine distinctions in terms of the manner in which the object is carried or held. However, importantly, the key distinction the English language makes (i.e., whether the agent is moving or not while holding the object) is not relevant in Chinese. There are roughly 20 verbs that can be translated into English as "to hold" or "to carry," and each of them, in Chinese, refers to a different event of a person holding/carrying an object in a distinct manner on a particular body part. When the object is held in the hand, different verbs are applied depending on the shape of the hand and the arm. Here, although certain objects typically appear with specific verbs (e.g., bowls with "ding" [carry on head], trays with "tuo" [carry on palm], children and backpacks with "bao" [carry in two arms]), the verbs can be used for other objects as long as the object can be held in the manner denoted by the verb. For example, carrying/holding an object on one's head is denoted by *ding* (顶),[1] while carrying/holding an object on one's shoulder is *kang* (扛). Carrying/holding an object with two arms is denoted by *bao* (抱), but if the object is held with one arm at the side of the body, the action is called *jia* (夹). Several verbs like *na* (拿), *ti* (提), and *lin* (拎) refer to carrying/holding actions with one hand, and verb choice depends largely on the shape of the hand holding the object. However, these verbs are not necessarily all contrastive with clear gaps among them, nor are they configured with each separated evenly from each other. Instead, one part of the semantic space is densely covered by several close synonyms with overlapping boundaries, while other parts of the space are only sparsely covered with clear gaps with other verbs.

What Do Children Need to Discover to Acquire Verbs in a Complex-Structured Lexical Domain?

How do Chinese-learning children acquire individual verbs within such a complex system? To be able to use words as adults do in their language community, children first need to discover the semantic features underlying the given semantic domain in the language. They also need to discover the relations among these words and learn where the boundaries are drawn between different word meanings. To understand how children discover and acquire extremely complex lexical systems, we need to understand how the representations of word meanings from the same domain start out.

For this purpose, Imai and colleagues examined how children learning Mandarin Chinese as their native language understand the meanings of Chinese carrying/holding verbs and how the learners' meanings of these verbs evolve as they accumulate learning experiences (Saji et al., 2011). The authors employed production data as the index of learners' knowledge of word meanings rather than comprehension data. In the field of vocabulary development, using comprehension data tends to be more common than production data (Ellis, 1995; Gleitman, 1990; Imai et al., 1994; Markman, 1989). However, as discussed earlier, a deep understanding of word

[1] We had originally conducted separate analyses both for the "carrying" and "holding" matrices, but the results were very similar. Thus, we only report the results of the analysis using the matrix for the "carrying" events to avoid redundancy.

meaning should be manifested by how precisely one can *apply* the word in various contexts, and this is difficult to assess using comprehension tasks.

Thirteen verbs in the domain were selected. These verbs are commonly used in everyday contexts and hence should be familiar to young children. Two video clips were prepared for each verb, one showing a carrying action with the actor moving with an object and the other showing a holding action with the actor holding an object while standing still. This manipulation was incorporated in the experimental design to see whether Chinese children would apply the same verb across the moving and nonmoving versions of the same action clips. In each trial, the participant was asked to name the action using a single verb. Trials were presented in a random order so that, if they understood the meaning of the verb, they should give the same verb across moving and nonmoving versions of the clips.

To understand the semantic structure of the domain in more detail, adult native speakers of Chinese (students of Beijing Normal University) were also tested on the production task with the same stimuli and procedure. Also tested were mothers of 2-year-olds or 5-year-olds to see if they adjusted their verb use when speaking to their child. Each mother is to describe each clip with a single verb to the child sitting next to her. After that, the mother was asked to describe each stimulus to an adult experimenter. Additionally, to understand the complexity of the lexical structure more closely, university students who had not participated in the production experiment were tested on a comprehension task. Every combination of the 13 target verbs and 13 stimuli was presented one at a time in random order, and the participants were asked to answer (with yes or no) whether the verb could be used to describe the given video clip.

Findings from the "Carry" Verb Acquisition Study

The production data were analyzed first in light of the following two points: (1) How many verb types children and adults produced across the 26 videos in each age group; (2) how closely the pattern of children's uses of verbs agreed with that of adults, and how patterns changed with age.

How Many Verb Types Did Children Know?

Traditionally, the most commonly used measure for vocabulary growth is the number of word types children produce or understand, for example, as reported in the MacArthur Communicative Development Inventories (CDI) (Fenson et al., 1994). Saji et al. (2011) counted the number of verb types each individual produced across the 26 carrying/holding videos, which were ideally denoted by the 13 target verb types. The adults (undergraduates) on average produced 11.2 verb types. Importantly, the mothers of 2-year-olds and 5-year-olds did not differ from the undergraduates in the number of verbs they produced.

Children produced a smaller number of verb types (7.25, 6.25, and 8.57 for 3-, 5-, and 7-year-olds, respectively), but within the three age groups, there was no developmental difference in the number of verb types children produced. The results suggest that adult native speakers of Chinese mostly used different verbs for each of the 13 carrying/holding actions. The 3- to 7-year-old children used fewer verbs than adults, but the number of verb types produced by the 3-, 5-, and 7-year-olds was approximately the same.

Does Children's Representation of the Lexical Domain Stay the Same Between 3 and 7?

Even if the number of verb types children produce spontaneously does not differ among the three age groups, how they use the verbs could still differ. To investigate this, Saji et al. (2011) compared the pattern of children's verb use with that of adults' following an algorithm used by Ameel, Malt, and Storms (2008). Here, we calculated the correlation between each age group and the undergraduate group, using the "carry" production matrix.[2]

First, whether children and adults produced the same verb for the 13 pairs of the video clips (i.e., one moving and the other standing still with the same manner of holding an object) was examined. The correlations were high for all four groups, although the agreement (or saying the same verb for the dynamic and static events) increased with age: 3-year-olds, $r = 0.71$; 5-year-olds, $r = 0.85$; 7-year-olds, $r = 0.84$; adults, $r = 0.94$. Thus, from early on, Chinese children understand that the distinction between "carrying" and "holding" (i.e., whether the event involves movement of the actor) is not relevant, and they applied the same verb consistently for both moving and nonmoving actions performed in the same manner on the same object; thus, the data were collapsed for further analyses.

Next, Fig. 7.1 shows the correlations between specific responses on each trial in each of the three age groups of children and adults. The correlation between 3-year-olds and adults was low ($r = 0.17$). The degree of convergence with adult verb use increased linearly from age 3 to 7 years (i.e., 5-year-olds, $r = 0.43$; 7-year-olds, $r = 0.58$). Thus, even though the 3-, 5-, and 7-year-olds did not differ in the number of verbs they produced, they did differ in how they applied the verbs. With age, children gradually converged on the adult pattern of use. However, the degree of convergence was not very high even for the 7-year-olds ($r = 0.58$), considering the high correlations (0.84 in average) among the three adult groups (undergraduates, mothers of 2-year-olds, and mothers of 5-year-olds). The results suggest that it takes a long time for children to learn how to use these words in the same way as do adults but that they make steady strides toward adult-like representations of the verbs.

[2] In Chinese, the distinction between a morpheme and a word is difficult to make. The 13 verbs were words consisting of a single morpheme. To make sure that the frequency count for each verb does not contain cases in which the same morpheme is a part of a different word (e.g., "ti" [dangle around the arm] used in "ti-gao" [to raise, to improve]), we went through the examples manually and excluded the latter cases from the counts.

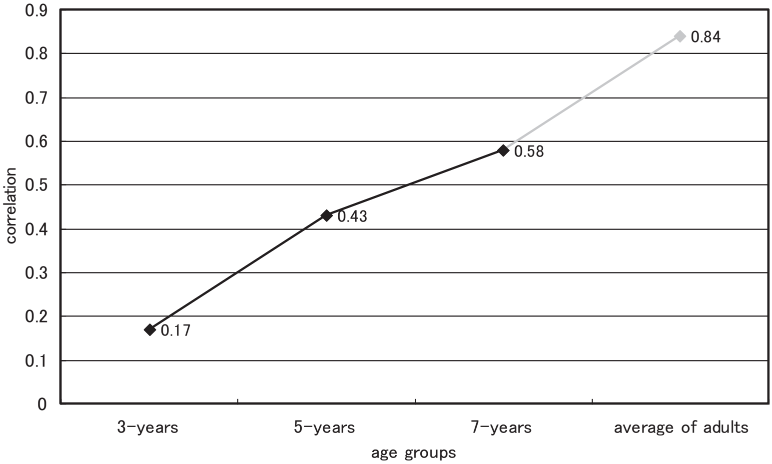


Fig. 7.1 Correlation between each of the three child groups and the adult group. Note: The value for the adult group (0.84) represents the average of correlation values between the undergraduate group and the two groups of mothers

Did mothers adjust their input to young children? In other words, did they use only the most frequent verbs when talking to them? Recall that mothers of 2-year-old and 5-year-old children were invited to name the same stimuli to their child to test this possibility. The correlations between the adult group (college students) and each of the two mother groups were higher than $r = 0.80$, suggesting that even mothers of 2-year-olds used the same set of verbs as they did with the experimenter. These results suggest that the major difference in the patterns of verb use for adults and children cannot be attributed to the input from caretakers; instead, it must be attributed to internal factors at work in the children.

Reliance of Object Similarity to Structure the Semantic Domain

In the next analysis, Saji et al. (2011) examined the similarity structure of this semantic domain for each age group to see how the structure of the lexicon evolves with development. A similarity matrix was created separately for each age group and was submitted to multidimensional scaling (MDS) analyses. Figure 7.2 shows MDS solutions for each age group including the adult group. In the adult data, the 13 verbs are separated from one another and configured into a circle, suggesting that different verbs were differentially applied to the 13 actions. In contrast, in 3-year-old Chinese children's plots (Fig. 7.2, upper left), the "na," "jia," and "peng" (捧) actions were completely overlapping, which indicates that they did not apply different verbs to these events while they applied verbs differentially to the other actions to some extent. This tendency was also observed but to a decreased degree in the 5- and 7-year-olds.

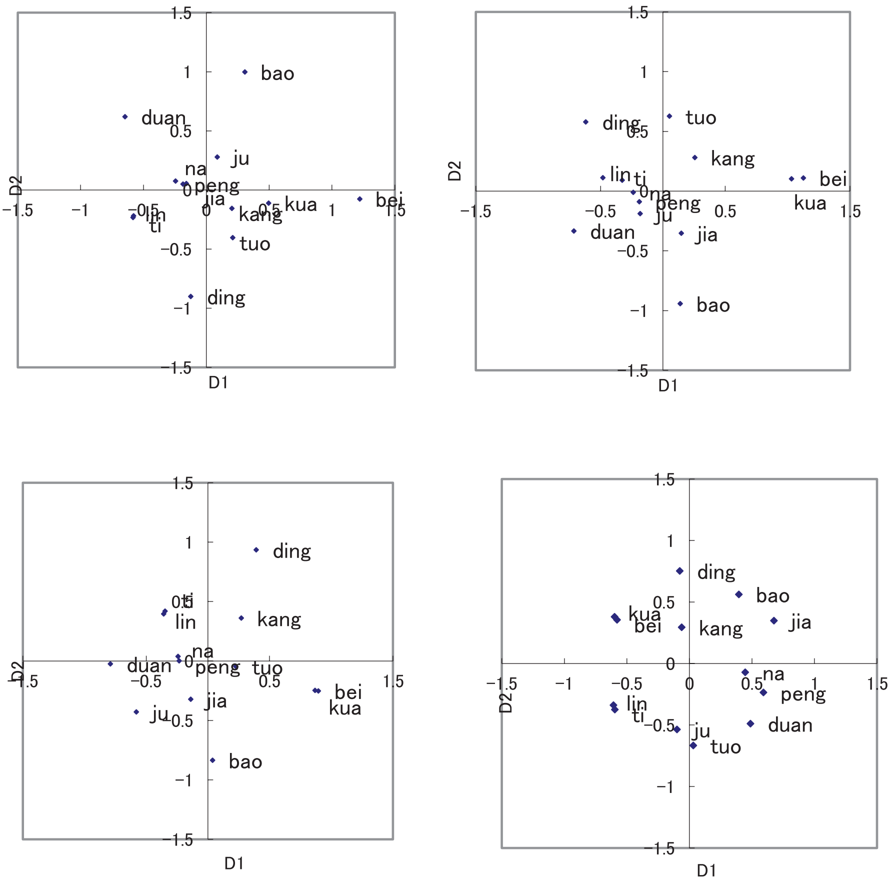


Fig. 7.2 MDS solution for the production data of native speakers of Chinese: Chinese 3-year-old children (upper left); Chinese 5-year-old children (upper right); Chinese 7-year-old children (lower left); Native speakers of Chinese (lower right)

Why was the children's pattern of verb use so different from that of adults? To learn word meanings, children have to detect which semantic features are critical for dividing up a given semantic domain. Perhaps the semantic features young children first rely on to organize the semantic domain are different from those used by adults (e.g., Bowerman, 2005; MacWhinney, 1987; Schaefer, 1979). As already discussed, the critical semantic features that differentiate carrying/holding verbs in Chinese concern specific body parts and the manner in which the actor supports the object (e.g., on the head, on the back, or on the shoulder). However, young children may not yet be aware that manner of support is more important than are other features involved in the events, such as the objects. In the novel verb generalization studies reviewed earlier in this chapter (Imai et al., 2005, 2008), 3-year-olds had difficulty segregating objects from the action but were helped when there was higher object similarity between the source and target events; in that case, children were able to understand the abstract "verbal core" of the verb meaning. We also described other studies of progressive alignment showing that having more similar comparisons helps children learn how to compare, and object similarity across examples is one form of similarity that would fit. It thus seems reasonable that young children could first organize a semantic domain for a verb around the type of the objects used in the verbs in the carrying/holding action.

To examine the possibility that children may use types of objects to organize their verb meanings, Saji et al. further conducted INDSCAL (INDividual SCALing MDS), a version of the MDS technique developed for evaluating individual/group differences in a multidimensional space common across groups (Carroll & Chang, 1970). While MDS can provide a visual representation of patterns of similarity or distance by detecting underlying dimensions from all of the input groups, INDSCAL allows us to capture the weights each input group assigned to the dimensions detected from all input groups. The four correlation matrices from the different age groups calculated in Analysis 2 were fed into INDSCAL as the input data. INDSCAL provided two kinds of output: first, it identified the dimensions underlying the verb production patterns, along which all the age groups categorized the videos (Common Space), and second, it identified the weights each group placed on each of the common dimensions when they named an event (Individual Space).

Figure 7.3a and b show the Common Space; the location of each event point was calculated using the data from all four age groups. Each point thus represents 13 videos of carrying, and distances between the points reflect the similarity among the videos based on the naming pattern produced (Fig. 7.3a for Dimension 1 × Dimension 2, Fig. 7.3b for Dimension 1 × Dimension 3). In the Common Space, if participants tended to apply the same verb to any two given videos, the distance between the two videos is small, and each of the dimensions extracted reflects a criterion by which the naming of the videos is distinguished. The Individual Space shows how the different age groups weighted the semantic features represented by each dimension (Fig. 7.4a and b).

For the configurations along the three dimensions, the videos plotted in the positive direction on Dimension 1 include carry actions where the object was supported by body parts other than hands (e.g., "ding" [carry on head], "kua" [hang on the shoulder], and "bei" [carry on back]), whereas the videos plotted in the negative direction were generally carrying actions where the object was carried with the hand (e.g., "lin" [dangle with one hand], "ti" [dangle around the arm], "na" [carry with one hand]). Thus, we can interpret Dimension 1 as representing the semantic feature distinguishing events via the manner of holding for the object.

The interpretation of Dimension 2 is less transparent, but it appears to distinguish the "bao" (carry in two arms) event (hugging a stuffed cat in two arms) from all other events. Given that the stuffed animal being carried in the "bao" event stands out from the other objects for children, Dimension 2 may be related to "salience of the object." Dimension 3 could be interpreted as the dimension differentiating the events according to "objects to be held." The events plotted along the positive direction of Dimension 3 included the "bei" [carry on back] events with a rucksack, "lin" [dangle with one hand] with a plastic shopping bag, "ti" [dangle around arm] with

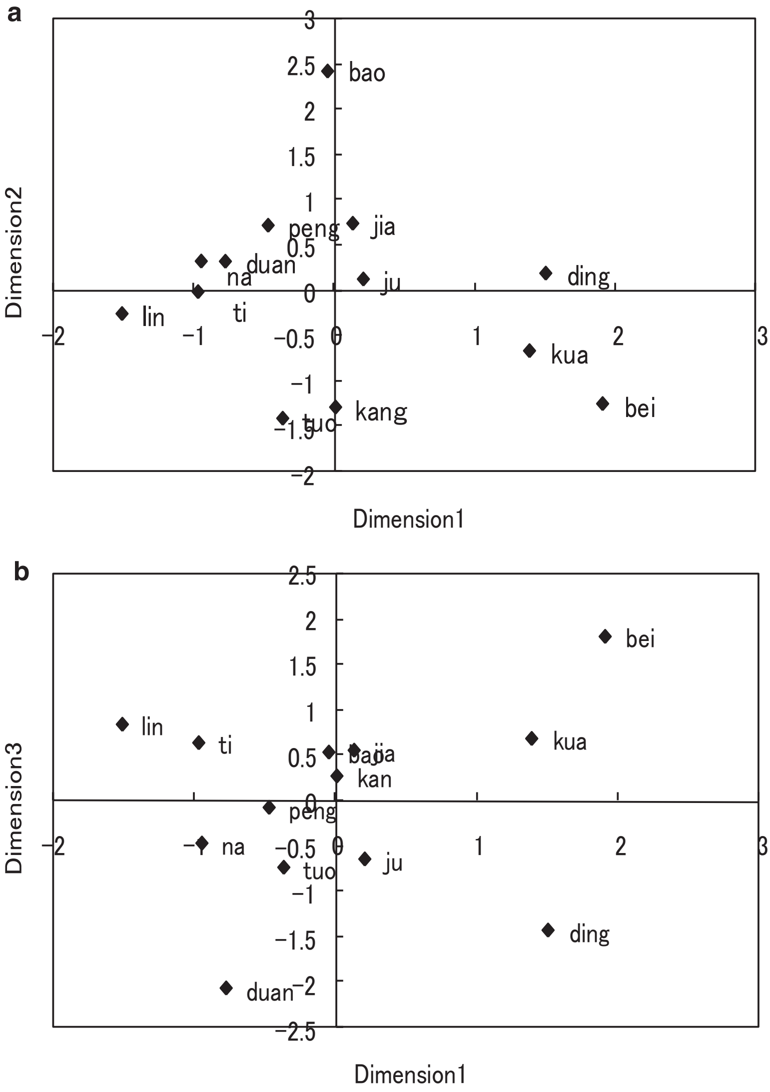


Fig. 7.3 A Common Space extracted in a INDSCAL model: (**a**) Dimension 1 × Dimension 2; (**b**) Dimension 1 × Dimension 3. Each plot represents 13 videos, and distances between plots represent the similarity of the verb production pattern

a tote bag, "kua" [hang on the shoulder] with a shoulder bag, and "jia" [carry under one arm], with a square business bag—where all the objects were bags of some kind. In contrast, the objects in the videos plotted along the negative direction were "bowls"—the "duan" [carry with two hands with caution] event with a glass bowl with water and the "ding" [carry on head] event with a wooden bowl. In summary,

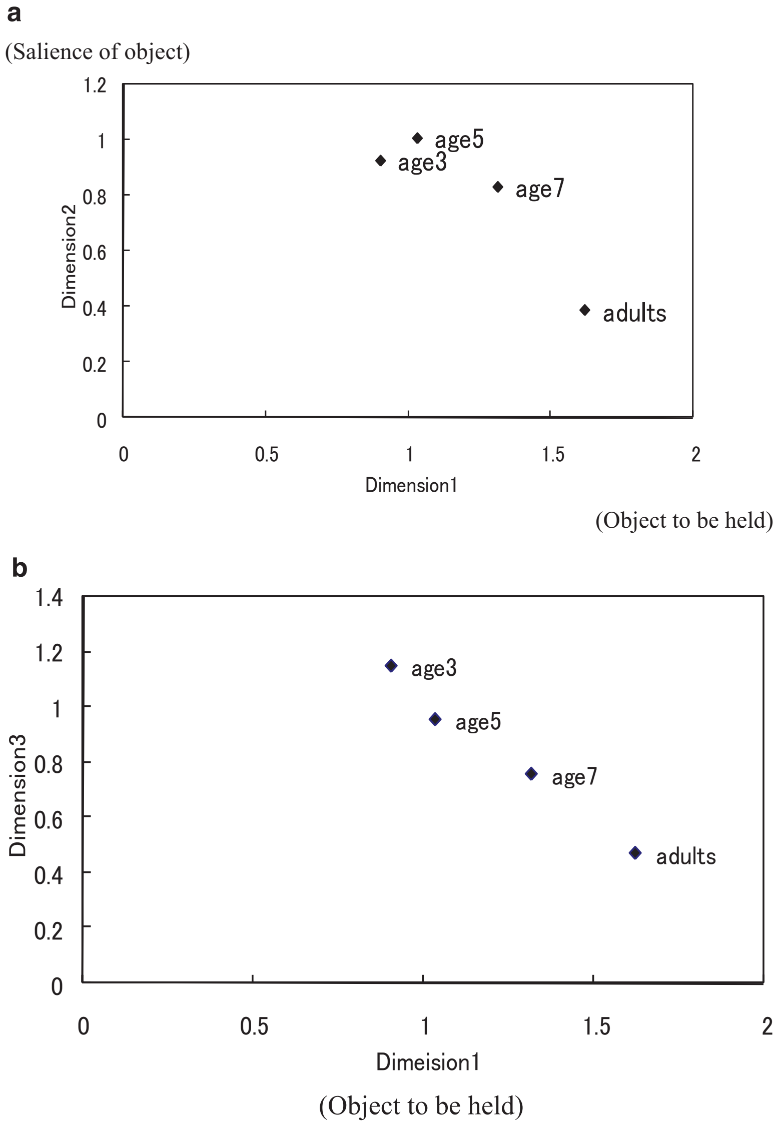


Fig. 7.4 An Individual Space extracted in an INDSCAL model: (**a**) Dimension 1 × Dimension 2; (**b**) Dimension 1 × Dimension 3

the body support and the manner of carrying and holding showed up as Dimension 1, and the object properties are represented in Dimensions 2 and 3.

How were the three dimensions weighted by the four different age groups? Figure 7.4a and b show the weight plots for the four groups on the three dimensions in Common Space. As expected, there were major differences between children and adults in the weights for each dimension. While Dimension 1 (salience of body parts) was more important for adults than Dimension 2 (salience of object) or Dimension 3 (object properties), Dimension 2 and Dimension 3 were more prominent than Dimension 1 for the three child groups.

Summarizing the results of the two MDS analyses (MDS for each age group and INDSCAL), the following developmental process may take place in Chinese learning children. They first form three semantic islands: an island around *na* (actions using one hand), *bei* (actions in which objects were supported by the back), and *bao* (actions in which objects were held by two arms), perhaps by perceptual saliency of body parts with which the object is held. They suggest children noticed that different types of objects were used for each type of action such that smallish objects that can be held up by one hand tended to appear with the verb *na,* bigger objects (often small animals or children) with *bao*, and backpacks with *bei*. As children observe a new verb used to refer to a "similar" action with a similar object, they tentatively locate it in one of the three islands, and differentiation of the verbs gradually takes place within each cluster (perhaps using attention to implicit contrast, discussed in the prior section).

Factors Determining the Ease of Learning

Saji et al. (2011) further analyzed which factors affect the ease of learning. Here, two criteria for determining the ease (or difficulty) of learning were considered: (1) how frequently a given word is used by children and (2) how closely children can apply the verb to videos that fit the range used by adults. In the literature of lexical development, "early learned words" are usually considered "easy words to learn." However, examining which verbs are used in an adult-like way from an early age, and which ones are used differently, may provide insight into understanding the factors that affect the acquisition of verb meanings. For this purpose, Saji et al. adopted entropy (H) as a quantitative index to represent how children and adults differentially use verbs. Entropy has been used as a measure for showing the proximity for a thermodynamic system to equilibrium, but it is now widely used in information theory and statistics (Mori & Yoshida, 1990). The notion of entropy in statistics is often used in descriptive statistics as an index to represent the degree of dispersion of responses for a categorical variable. If the responses are concentrated in one or a small number of response categories, the entropy value becomes low; if they are widely dispersed across different response categories, the value becomes high (see Saji et al. for the formula). In the current context, if the range of application of the verb is restricted to one or a small number of videos, the entropy value will be low, whereas if participants produce a single verb for a range of videos, the entropy value will be high. Low entropy values for the adults could suggest that for a given action, they used the verb that we had originally intended (e.g., used the verb "ding" [carry on head] for the video of what we intended to be a "ding" action) most of the time; in this case, the degree of dispersion of the verb use would be small. In contrast, if children tended to apply each verb to a broader range of videos, the entropy value would be high.

Two important trends emerged from entropy analyses. First, overall, young children tended to use various verbs for a given action, while adults tended to use a specific verb for a specific event with high agreement. However, some verbs (e.g., "ding" [carry on head]) converged on the adult pattern almost from the beginning. Children used "ding" only for the action adults also described with "ding," without applying other verbs such as "na" [carry with one hand] to this action, nor did they overuse "ding" for other hold/carry actions in the domain. This suggests that, in a semantic domain, if words are first overextended and then gradually restricted with development, this process may happen more quickly for some verbs than others. (Of course, the alternative is that verbs are used conservatively at first (e.g., see Tomasello, 2000, which could also explain these results)). As discussed in the previous section, "na" [carry with one hand], "bao" [carry with two arms], and "bei" [carry on the back] form the first three semantic islands and are considered as "early learned verbs" in previous studies (e.g., Hao, Shu, Xing, & Li, 2008). Interestingly, we found that the words showing the higher entropies were the ones children used broadly and more frequently than other words. The verb "ding" [carry on the head] had the lowest entropy, meaning that it was not confused with other verbs. This verb seems to differ from "na," "bei," and "bao" in that it has no close neighbors that share boundaries with it. Perhaps the convergence to the range of adults' use was influenced by the presence of similar-meaning words sharing boundaries in the same semantic domain. If a word does not have close neighbors with overlapping boundaries, the degree of convergence may be high even from very early stages of lexical development. In contrast, if a word has overlapping boundaries with many other words, the degree of convergence between children's meaning and adults' meaning may be low at early ages, and it may take a long time for children to arrive at a meaning equivalent to that possessed by adults because then children must delineate many boundaries with many similar-meaning words.

Second, "na" and "ding" also differ greatly in the range of instances adults accept as referents. In the production task, adult Chinese speakers used "na" for the video we assumed to be the "na" action and did not use it for other actions as they preferred to use verbs that specifically designated those actions. However, the comprehension data indicated that adults would also accept actions denoted by other hand-holding actions such as "ti" [dangle around the arm] and "lin" [dangle with one hand] as referents of "na," although to a lesser degree. The reverse direction was not observed: Adults did not judge the verbs "ti" or "lin" to be acceptable to refer to the "na" video. Thus, "na" has a broader range of applicability than the neighboring verbs "lin" and "ti." Importantly, children used "na" not only for the actions adults accepted as referents of this verb but also for those adults did not accept. The adult comprehension data revealed that "bei" [carry on back] and "bao" also cover broader ranges than other verbs. It may be the case that children overextend a word that covers the broadest range of referents in the semantic domain, which might result in late convergence with adults' meanings, and this is why we see the three islands around "na," "bei," and "bao."

Additionally, Saji et al. used regression analyses to examine whether two semantic properties of the verbs—the degree of boundary overlap with neighboring words (*boundary overlap*) and the range to which the verb is applied (*verb coverage*)—affect how "easily" children learn verbs. To quantify these values, the data from the adult comprehension task were used.

To represent the degree of *boundary overlap*, the entropy value was calculated for *each action*. If many verbs are accepted for a given action, it means that the video originally created to represent one verb can be described using other verbs. Hence, the boundary of the verb with other neighboring verbs is somewhat continuous, and the verb has a high degree of boundary overlap with other verbs. On the other hand, if only one verb is accepted for the action across different adult participants, there is little boundary overlap with other verbs. To quantitatively represent the second predictor, *verb coverage*, the entropy value was obtained for *each verb*. Here, if a given verb was accepted for many different actions—that is, if the verb covers a wide range of action instances—the entropy for the verb will be high. In addition, the influence of *word frequency* in the model was examined, as it has been considered as an important predictor in accounting for how early the word enters children's vocabulary (e.g., Li, Zhao, & MacWhinney, 2007).

The results of the regression analyses revealed that the three factors (i.e., frequency, breadth of coverage, and boundary overlap) contributed differently in accounting for the "ease of learning" for the two different definitions of "ease" (i.e., frequency of use by children and convergence with the adult pattern; see Saji et al. 2011 for details). There was also an interesting developmental trend in the relative weights of the three factors. First, using frequency of use by children as the dependent variable, (in which the ease of learning for the 13 verbs was indexed by how willingly children used these verbs), the *word frequency* in the adult corpus made the strongest contribution for all three age groups (3 years: $\beta = 0.65$, $t = 3.5$, $p < 0.01$; 5 years: $\beta = 0.60$, $t = 2.8$, $p < 0.05$; 7 years: $\beta = 0.59$, $t = 3.0$, $p < 0.05$), suggesting that the verbs young children tend to produce frequently are also the ones that they hear most frequently. The degree of *verb coverage* did not make a significant unique contribution to the model (3 years: $\beta = 0.42$, $t = 2.0$, n.s.; 5 years: $\beta = 0.41$, $t = 1.6$, n.s.; 7 years: $\beta = 0.45$, $t = 2.0$, n.s.). On the other hand, the degree of *boundary overlap* contributed to the model in 7-year-olds but not in younger children (3 years: $\beta = -0.27$, $t = -1.5$, n.s.; 5 years: $\beta = -0.39$, $t = -1.9$, n.s.; 7 years: $\beta = -0.49$, $t = -2.6$, $p < 0.05$). These results suggest that older children tend to produce words with distinct boundaries with neighboring verbs.

For the models using the degree of convergence with adults' use of the verbs as the dependent variable, the degree of *boundary overlap* contributed most strongly to the model. The β value for the degree of boundary overlap was significant for all ages (3 years: $\beta = -0.73$, $t = -0.2.9$, $p < 0.05$; 5 years: $\beta = -0.81$, $t = -3.1$, $p < 0.05$; 7 years: $\beta = -0.86$, $t = -3.8$, $p < 0.01$). The negative direction of the β values indicates that the higher the degree of boundary overlap, the lower the degree of convergence in children's use of verbs with that of adults. In this case, for none of the three age groups did *verb coverage* or *word frequency* make a significant contribution.

The results of the regression analyses thus suggest that different factors underlie the two different processes of word learning. At early stages of word learning, fast word-world mapping is very important. There, the input frequency plays a more prominent role than do semantic properties of the target word such as boundary overlap and breadth of meaning: Children produce the words they hear most often. However, for the later process of word learning, the degree of boundary overlap with other verbs is more strongly related to the degree of convergence with adults' use: The more the word has neighboring words with overlapping boundaries, the longer it takes for children to attain adultlike meanings.

Summary and Implications: Object Saliency, Similarity, and Contrast as Driving Forces for Structuring the Verb Lexicon

The study presented in this section clearly shows that verb meaning acquisition in a full sense is a long, protracted process. Children cannot acquire adultlike meanings of individual words by fast-mapping as the acquisition of a word's full meaning requires understanding how a word is different from other words surrounding it. To do so however, children first need to know what semantic domain the word belongs to, what words exist in the domain, and how these words divide the domain. In short, to acquire the full meaning of a word, children need to acquire the representation of the semantic domain as a whole.

How is the representation of the semantic domain structured, and how does it become more adultlike? The result of the study by Saji et al. (2011) suggests that children learning verbs for "carry" first organize the domain by salience of body parts with which the object is held, and by similarity of the objects, such that children apply the same verb for actions involving similar objects. Thus, object saliency and object similarity scaffold children's construction of a lexical system as they fast-map verbs to event scenes. However, as they add new verbs in the lexicon of the given semantic domain, children need to place the newly learned verb in relation to the words that already exist in the lexicon. In so doing, they may rely on object similarity at first, but they need to compare the new verb with the existing ones, align them, and find commonalities and differences between them (e.g., Childers et al., 2016; Childers & Paik, 2009; Gentner & Namy, 2006).

Conclusion

In this chapter, we have outlined a bootstrapping process that children could use to form initial verb meanings from sound and object similarity information available to them (most useful if they are deducing meaning from a single event). We then discussed how children could discover and come to rely on specific linguistic features in their language and then how they could create a lexical system of verb meanings by comparing and contrasting verbs with each other. Children could use information they glean across multiple events for all three of these problems they face, though we have focused on how they may use comparison information particularly for the most complex problems in verb learning. One central mechanism they could use when comparing events is structural alignment, or the ability to compare events by aligning elements across events with each other. We have described experimental evidence showing children's verb learning is consistent with structural alignment theory. In comparing structural alignment to statistical learning, structural alignment is a more specific account of the mental processing children may use when comparing events, whereas in statistical learning, children would attend to any statistical regularity in the input (e.g., see Chaps. 2 and 4, for discussions of statistical learning theory). It is possible that infants start with statistical learning mechanisms and apply those to early stages of verb learning and then use the more specific structural processing as they build the lexicon (see Childers, Bottera, & Howard (2017) for a further comparison of these theories). It is also possible that children use syntactic cues to build their understanding of verb meaning (e.g., Gleitman, 1990; Naigles, 1990; Yuan & Fisher, 2009), though we see these cues as interacting with statistical learning or comparison processes in verb learning (see Fisher, 2002; Scott & Fisher, 2009) as opposed to explaining the whole of verb learning on their own.

Verb learning is a difficult but exciting area of research, which is progressing. By considering what mechanisms children have available to solve these important verb problems in development, we encourage other researchers to continue to investigate more than children's initial verb meanings but how they go beyond those meanings to become productive speakers of a language.

References

Aichison, J. (1987). *Words in the mind: An introduction to the mental lexicon*. Oxford: Basil Blackwell.

Ameel, E., Malt, B., & Storms, G. (2008). Object naming and later lexical development: From baby bottle to beer bottle. *Journal of Memory and Language, 58*, 262–285.

Au, T. K., & Laframboise, D. E. (1990). Acquiring color names via linguistic contrast: The influence of contrasting terms. *Child Development, 61*, 1808–1823.

Au, T. K., & Markman, E. M. (1987). Acquiring word meanings via linguistic contrast. *Cognitive Development, 2*, 217–236.

Bolinger, D. (1977). *Meaning and form*. London and New York: Longman.

Booth, A. E., & Waxman, S. R. (2009). A horse of a different color: Specifying with precision infants' mappings of novel nouns and adjectives. *Child Development, 80*, 15–22.

Bowerman, M. (1982). Reorganizational processes in lexical and syntactic development. In L. R. Gleitman & E. Wanner (Eds.), *Language acquisition: The state of the art* (pp. 319–346). Cambridge: Cambridge University Press.

Bowerman, M. (2005). Why can't you "open" a nut or "break a cooked noodle?": Learning covert object categories in action word meanings. In L. Gershkoff Stowe & D. H. Rakinson (Eds.), *Building object categories in developmental time* (pp. 209–243). Mahwah, NJ: Lawrence Erlbaum.

Carey, S. (1978). The child as word learner. In M. Halle, J. Bresnan, & G. A. Miller (Eds.), *Linguistic theory and psychological reality* (pp. 264–293). Cambridge, MA: MIT Press.

Carroll, J. D., & Chang, J. J. (1970). Analysis of individual differences in multidimensional scaling via an N-way generalization of "Eckart-Young" decomposition. *Psychometrika, 35*, 283–319.

Childers, J. B., Bottera, A., & Howard, T. (2017). Verbs: Learning how speakers use words to refer to actions. In G. Westermann & N. Mani (Eds.), *Current issues in developmental psychology: Early word learning* (pp. 84–96). New York: Routledge.

Childers, J. B., Hirshkowitz, A., & Benavides, K. (2014). Attention to explicit and implicit contrast in verb learning. *Journal of Cognition and Development, 15*, 213–237.

Childers, J. B., & Paik, J. H. (2009). Korean-and English-speaking children use cross-situational information to learn novel predicate terms. *Journal of Child Language, 36*(1), 201–224.

Childers, J. B., Parrish, R., Olson, C. V., Burch, C., Fung, G., & McIntyre, K. P. (2016). Early verb learning: How do children learn how to compare events? *Journal of Cognition and Development, 17*(1), 41–66.

Clark, E. V. (1990). On the pragmatics of contrast. *Journal of Child Language, 17*(2), 417–431.

Cruse, D. A. (1986). *Lexical semantics*. Cambridge: Cambridge University Press.

Ellis, R. (1995). Modified oral input and the acquisition of word meanings. *Applied Linguistics, 16*, 409–441.

Fenson, L., Dale, P. S., Reznick, J. S., Bates, E., Thal, D. J., & Pethick, S. J. (1994). Variability in early communicative development. *Monographs of the Society for Research in Child Development, 59*, i–185.

Fillmore, C. J. (1982). Frame semantics. In The Linguistic Society of Korea (Ed.), *Linguistics in the morning calm: Selected papers from SICOL 1981* (pp. 111–138). Seoul: Hanshin Publishing.

Fillmore, C. J., & Atkins, B. T. (1992). Toward a frame-based lexicon: The semantics of RISK and its neighbors. In A. Lehrer & E. Kittay (Eds.), *Frames, fields, and contrasts* (pp. 75–102). Hillsdale, NJ: Lawrence Erlbaum.

Fisher, C. (1996). Structural limits on verb mapping: The role of analogy in children's interpretations of sentences. *Cognitive Psychology, 31*, 41–81.

Fisher, C. (2002). Structural limits on verb mapping: The role of abstract structure in 2.5-year-olds' interpretations of novel verbs. *Developmental Science, 5*, 55–64.

Fisher, C., Gleitman, H., & Gleitman, L. R. (1991). On the semantic content of subcategorization frames. *Cognitive Psychology, 23*(3), 331–392.

Gentner, D. (1982). *Why nouns are learned before verbs: Linguistic relativity versus natural partitioning* (Center for the Study of Reading, Technical Report No. 257).

Gentner, D. (2010). Bootstrapping the mind: Analogical processes and symbol systems. *Cognitive Science, 34*(5), 752–775.

Gentner, D., & Boroditsky, L. (2001). Individuation, relativity, and early word learning. *Language Acquisition and Conceptual Development, 3*, 215–256.

Gentner, D., & Kurtz, K. J. (2005). Relational categories. In W. K. Ahn, R. L. Goldstone, B. C. Love, A. B. Markman, & P. W. Wolff (Eds.), *Categorization inside and outside the lab* (pp. 151–175). Washington, DC: American Psychological Association.

Gentner, D., Loewenstein, J., & Hung, B. (2007). Comparison facilitates children's learning of names for parts. *Journal of Cognition and Development, 8*(3), 285–307.

Gentner, D., & Namy, L. L. (2006). Analogical processes in language learning. *Current Directions in Psychological Science, 15*(6), 297–301.

Gentner, D., & Rattermann, M. J. (1991). *Language and the career of similarity* (Center for the Study of Reading, Technical Report No. 533). University of Illinois at Urbana-Champaign.

Gleitman, L. (1990). The structural sources of verb meanings. *Language Acquisition, 1*(1), 3–55.

Golinkoff, R. M., Chung, H. L., Hirsh-Pasek, K., Liu, J., Bertenthal, B. I., Brand, R., et al. (2002). Young children can extend motion verb labels to point-light displays. *Developmental Psychology, 38*, 604–614.

Hao, M., Shu, H., Xing, A., & Li, P. (2008). Early vocabulary inventory for Mandarin Chinese. *Behavior Research Methods, 40*, 728–733.

Haryu, E., & Imai, M. (2002). Reorganizing the lexicon by learning a new word: Japanese children's interpretation of the meaning of a new word for a familiar artifact. *Child Development, 73*(5), 1378–1391.

Haryu, E., Imai, M., & Okada, H. (2011). Object similarity bootstraps young children to action-based verb extensions. *Child Development, 82*, 674–686.

Hollich, G., Golinkoff, R. M., & Hirsh-Pasek, K. (2007). Young children associate novel words with complex objects rather than salient parts. *Developmental Psychology, 43*(5), 1051–1061.

Imai, M., & Gentner, D. (1997). A crosslinguistic study of early word meaning: Universal ontology and linguistic influence. *Cognition, 62*, 169–200.
Imai, M., Gentner, D., & Uchida, N. (1994). Children's theories of word meaning: The role of shape similarity in early acquisition. *Cognitive Development, 9*, 45–75.
Imai, M., & Haryu, E. (2001). How do Japanese children learn proper nouns and common nouns without clues from syntax? *Child Development, 72*, 787–802.
Imai, M., Haryu, E., & Okada, H. (2005). Mapping novel nouns and verbs onto dynamic action events: Are verb meanings easier to learn than noun meanings for Japanese children? *Child Development, 76*(2), 340–355.
Imai, M., & Kita, S. (2014). The sound symbolism bootstrapping hypothesis for language acquisition and language evolution. *Philosophical Transactions of the Royal Society B: Biological Sciences, 369*(1651), 20130298.
Imai, M., Li, L., Haryu, E., Okada, H., Hirsh-Pasek, K., Golinkoff, R., et al. (2008). Novel noun and verb learning in Chinese-, English-, and Japanese-speaking children. *Child Development, 79*, 979–1000.
Kantartzis, K., Imai, M., & Kita, S. (2011). Japanese sound-symbolism facilitates word learning in English-speaking children. *Cognitive Science, 35*(3), 575–586.
Landau, B., Smith, L. B., & Jones, S. S. (1988). The importance of shape in early lexical learning. *Cognitive Development, 3*(3), 299–321.
Levin, B. (1993). *English verb classes and alternations: A preliminary investigation*. Chicago: University of Chicago press.
Li, P., Zhao, X., & MacWhinney, B. (2007). Dynamic self-organization and early lexical development in children. *Cognitive Science, 31*, 581–612.
Lidz, J., Gleitman, H., & Gleitman, L. (2003). Understanding how input matters: Verb learning and the footprint of universal grammar. *Cognition, 87*, 151–178.
Lyons, J. (1963). *Structural semantics*. Cambridge: Cambridge University Press.
Lyons, J. (1977). *Semantics* (Vol. 1). Cambridge: Cambridge University Press.
MacWhinney, B. (1987). The competition model. In B. MacWhinney (Ed.), *Mechanisms of language acquisition* (pp. 249–308). Hillsdale, NJ: Lawrence Erlbaum.
Maguire, M. J., Hirsh-Pasek, K., Golinkoff, R. M., & Brandone, A. C. (2008). Focusing on the relation: Fewer exemplars facilitate children's initial verb learning and extension. *Developmental Science, 11*, 628–634.
Markman, E. M. (1989). *Categorization in children: Problems of induction*. Cambridge, MA: MIT Press, Bradford Books.
Markman, E. M. (1990). Constraints children place on word meanings. *Cognitive Science, 14*(1), 57–77.
Markman, E. M., & Hutchinson, J. E. (1984). Children's sensitivity to constraints on word meaning: Taxonomic vs. thematic relations. *Cognitive Psychology, 16*, 1–27.
Markman, E. M., & Wachtel, G. F. (1988). Children's use of mutual exclusivity to constrain the meanings of words. *Cognitive Psychology, 20*(2), 121–157.
Miller, G., & Johnson-Laird, P. (1976). *Language and perception*. Cambridge, MA: Harvard University Press.
Mori, T., & Yoshida, T. (1990). *Technical book of data analysis for psychology*. Kyoto: Kitaoji Press.
Muehleisen, V., & Imai, M. (1997). Transitivity and the incorporation of ground information in Japanese path verbs. In K. Lee, E. Sweetwer, & M. Verspoor (Eds.), *Lexical and syntactic constructions and the construction of meaning* (pp. 329–346). Amsterdam: John Benjamins.
Naigles, L. (1990). Children use syntax to learn verb meanings. *Journal of Child Language, 17*, 357–374.
Namy, L. L., & Clepper, L. E. (2010). The differing roles of comparison and contrast in children's categorization. *Journal of Experimental Child Psychology, 107*, 291–305.
Perniss, P., & Vigliocco, G. (2014). The bridge of iconicity: From a world of experience to the experience of language. *Philosophical Transactions of the Royal Society B, 369*, 1–13.

Pustejovsky, J. (1995). *The generative lexicon*. Cambridge, MA: MIT Press.
Roseberry, S., Hirsh-Pasek, K., Parish-Morris, J., & Golinkoff, R. M. (2009). Live action: Can young children learn verbs from video? *Child Development, 80*, 1360–1375.
Saji, N., Imai, M., Saalbach, H., Zhang, Y., Shu, H., & Okada, H. (2011). Word learning does not end at fast-mapping: Evolution of verb meanings through reorganization of an entire semantic domain. *Cognition, 118*(1), 45–61.
Saussure, F. de (1916/1983). *Course in general linguistics* (R. Harris, Trans). London: Duckworth.
Schaefer, R. P. (1979). Child and adult verb categories. *Kansas Working Papers in Linguistics, 4*, 61–76.
Scott, R. M., & Fisher, C. (2009). Two-year-olds use distributional cues to interpret transitivity-alternating verbs. *Language and Cognitive Processes, 24*(6), 777–803.
Talmy, L. (1975). Syntax and semantics of motion. *Syntax and Semantics, 4*, 181–238.
Talmy, L. (1985). Lexicalization patterns: Semantic structure in lexical forms. In T. Shopen (Ed.), *Language typology and syntactic description* (pp. 36–149). Cambridge: Cambridge University Press.
Tomasello, M. (2000). Do young children have adult syntactic competence? *Cognition, 74*, 209–253.
Tomasello, M., & Barton, M. E. (1994). Learning words in nonostensive contexts. *Developmental Psychology, 30*(5), 639–650.
Tomasello, M., & Kruger, A. C. (1992). Joint attention on actions: Acquiring verbs in ostensive and non-ostensive contexts. *Journal of Child Language, 19*(2), 311–333.
Waxman, S. R., & Klibanoff, R. S. (2000). The role of comparison in the extension of novel adjectives. *Developmental Psychology, 36*, 571–581.
Waxman, S. R., Lidz, J. L., Braun, I. E., & Lavin, T. (2009). Twenty four-month-old infants' interpretations of novel verbs and nouns in dynamic scenes. *Cognitive Psychology, 59*, 67–95.
Yoshida, H. (2012). A cross-linguistic study of sound symbolism in children's verb learning. *Journal of Cognition and Development, 13*, 232–265.
Yuan, S., & Fisher, C. (2009). "Really? She blicked the baby?": Two-year-olds learn combinatorial facts about verbs by listening. *Psychological Science, 20*, 619–626.

Chapter 8
Multiple Examples Support Children's Word Learning: The Roles of Aggregation, Decontextualization, and Memory Dynamics

Catherine Sandhofer and Christina Schonberg

Abstract Young children discover the meaning of words from hearing words used across time and across contexts. Children learn to label not only the specific instances they have experienced, but they also learn the meaning of words appropriately to new instances. Moreover, children remember these word-referent pairs across a period of time, such that they are able to recall the appropriate word after delays of days or weeks. In this chapter, we address these aspects of word learning – how do children generalize instances to new situations and remember word-referent pairs across time? In doing so, we discuss statistical learning as a mechanism for word learning with a specific focus on the processes of aggregation and abstraction. Second, we discuss how multiple examples dynamically support the retention of word-referent pairs.

Young children discover the meaning of words from hearing words used across time and across contexts. Children learn to label not only the specific instances they have experienced (e.g., using "ball" to label their own soft red ball), but children also generalize the meaning of words appropriately to new instances (e.g., pairing the label "ball" with unfamiliar balls). Moreover, children remember these word-referent pairs across a period of time, such that they are able to recall the appropriate word after delays of days or weeks. This chapter focuses on these aspects of word learning – how do children generalize instances to new situations and remember word-referent pairs across time?

Although children are capable of one-trial mapping of words to referents (e.g., Heibeck & Markman, 1987), children typically experience many instances of a

C. Sandhofer (✉)
University of California, Los Angeles, Los Angeles, CA, USA
e-mail: sandhofer@psych.ucla.edu

C. Schonberg
University of Wisconsin–Madison, Madison, WI, USA

J. B. Childers (ed.), *Language and Concept Acquisition from Infancy Through Childhood*, https://doi.org/10.1007/978-3-030-35594-4_8

word before it makes it into their early lexicons. For example, Hart and Risley (1995) estimate that children hear between 300 and 400 utterances an hour. And Roy documented that his son heard approximately 8 million words from the ages of 9–24 months, and 17,529 instances of the word "water" before his first production of the word (Roy, Frank, DeCamp, Miller, & Roy, 2015). Thus, young children's language environments are chock-full of multiple repetitions of words.

Hearing words multiple times appears to benefit learning, such that there is a strong relationship between the frequency of words in language input to children and children's vocabulary acquisition (Hoff & Naigles, 2002). Children with the largest vocabularies have parents who provide the most language input (Hart & Risley, 1995; Hoff, 2003; Hurtado, Marchman, & Fernald, 2008; Huttenlocher, Haight, Bryk, Seltzer, & Lyons, 1991), and the words that are most frequent in a parent's speech to a child tend to be the words that children produce earliest (Hart, 1991; Moerk, 1980) and the words that children comprehend most efficiently (Hurtado et al., 2008).

One reason that multiple presentations benefit word acquisition is that, across presentations, the accompanying but nonessential features are likely to vary. Thus, each presentation provides new information about word meaning to some degree and, in doing so, aids in word generalization. Frequent presentations also benefit word acquisition because multiple presentations, particularly when experienced across time, provide strong support for remembering word-referent pairs. In this chapter, we first discuss statistical learning (also see Johnson, this volume, and Theissen, this volume) as a mechanism for word learning with a specific focus on the processes of aggregation and abstraction. Second, we discuss findings from the memory literature (see Theissen, this volume, for a similar section) that point to mechanisms for how children may remember word-referent pairs across time.

Statistical Learning

Statistical learning is a general learning mechanism that involves extracting statistical regularities in the environment – typically low-level perceptual co-occurrences. Through sensitivity to these low-level co-occurrence patterns in the environment, higher-level structure emerges. Statistical learning has been demonstrated across a broad range of learning tasks including learning distributions of shapes presented in temporal streams and spatial arrays (Tummeltshammer, Amso, French, & Kirkham, 2017), using visual feature co-occurrences to form representations of object integrity (Wu, Gopnik, Richardson, & Kirkham, 2011), and linking statistical distributions of sound sequences to meaning (Graf Estes, Evans, Alibali, & Saffran, 2007).

At its core, statistical learning relies on the existence of statistical regularities within the world. Statistical learning is an especially good fit for understanding how children learn words because language input includes cues to the statistical regularity of language structure at multiple levels. For example, in a landmark study by Saffran, Aslin, and Newport (1996), infants were presented with 2 minutes of a continuous stream of nonsense syllables in which the frequency of one syllable following the next (i.e., the transitional probabilities) indicated the word boundaries. Infants responded differentially to syllable combinations that spanned word boundaries versus syllable combinations that did not span the word boundaries. These results demonstrate that children are sensitive to a set of regularities that may aid in parsing speech into word units, and they are able to extract these statistical probabilities through even a minimal amount of input (i.e., 2 minutes of input).

Recent work exemplifying how statistical learning might solve referential ambiguity about how words map onto meaning comes from the groundbreaking work of Linda Smith, Chen Yu, and colleagues on cross-situational word-referent learning (Yu & Smith, 2007, 2011; Yurovsky, Yu, & Smith, 2013). Because words' meanings are reflected in the statistics of their use, sensitivity to co-occurrence information can lead to discovering word meaning because across instances, there are co-occurrence regularities between words and referents. A growing body of work shows that adults, children, and infants can use co-occurrence information to map words to their meanings (see Scott & Fisher (2012); Suanda, Mugwanya, & Namy (2014) for examples). In these studies, participants learn words under situations of referential ambiguity. In a typical experiment, pictures of two unknown objects are presented to subjects and two unfamiliar words are uttered. For example, the learner might hear the words "natu" and "tikka" while viewing object 1 and object 2. With a single presentation, there is uncertainty as to which object "natu" refers to and which object "tikka" refers to. However, each word and its associated object referent reappear in subsequent presentations with another objects and word pair: on another trial, a child may hear the words "natu" and "wug" while viewing object 1 and object 3. Across presentations, the learner can determine the word's meaning by selecting from those meanings that reliably reoccur across situations and pair the word "natu" with object 1. In studies of this kind, infants, children, and adults are able to successfully map words onto referents. Thus, cross-situational learning provides a compelling model for how co-occurrences can create word learning.

Although infants are able to connect words to meanings in simple cross-situational experiments, the ability to map words to referents becomes disrupted with increasing time intervals between naming events and with greater visual complexity (Smith, Jayaraman, Clerkin, & Yu, 2018; Smith, Suanda, & Yu, 2014; Smith & Yu, 2013; Vlach & Johnson, 2013). In one study, Vlach and Johnson (2013) presented 16- and 20-month-olds with a cross-situational task in which the timing between subsequent presentations of the word-referent pairs varied. Half of the pairings were presented in immediate succession, and half were presented nonadjacently with five other trials separating each presentation of the word-referent pair. Both the 16- and 20-month-old children learned the word-referent pairs that were presented in immediate succession. However, only the 20-month-olds were able to learn the nonadjacent word-referent pairings. This suggests that young infants may encounter more difficulty aggregating information when presentations of words and referents reoccur after a long delay. Because real-world naming instances do not always occur in close temporal proximity, aggregating instances experienced over larger time periods may require extra support to promote memory and generalization.

Below, we discuss context and timing – two omnipresent environmental factors that have a history of demonstrated effects on memory performance – and how these factors may aid in aggregating temporally distant instances.

Multiple Examples Provide Support for Aggregation

Many instances of categories in the real world share similar contexts as well as similarity in features, properties, and functions. For example, consider the category of "toothbrush" and all of the different encounters with toothbrushes that a child experiences across a swath of time. Across these different instances, there are a lot of similarities. The different toothbrushes share similar features: they have bristles made of nylon and handles made of plastic. They are roughly the same shape and the same size. Moreover, from day to day, the toothbrushes also share a lot of contextual similarity: the toothbrushes are kept in the bathroom, they are talked about and used right before bedtime, and they are often proximally and temporally linked with toothpaste.

This background context is in no way intrinsic to being a member of the category of toothbrush. Indeed, a toothbrush is still a toothbrush when it appears in completely different contexts – such as on an archaeological dig. What matters from a statistical learning perspective, however, is the pattern of co-occurrence. The instances of toothbrush co-occur with multiple aspects of the background context: the bathroom, the toothpaste, the time of day, etc. Thus, children's initial category of "toothbrush" might include examples from all the toothbrushes they have experience with (e.g., their own purple toothbrush and their sister's pink glittery toothbrush) but may also be inseparably connected to the contexts in which in toothbrushes are experienced (e.g. brushing one's teeth at bedtime by the bathroom sink).

There is reason to expect that shared context might be helpful for learning and remembering categories. A large body of research has shown that the context in which something is learned has strong effects on memory encoding and retrieval (e.g., Butler & Rovee-Collier, 1989; Godden & Baddeley, 1975; Rovee-Collier & Dufault, 1991; Smith, Glenberg, & Bjork, 1978; Tulving, 1972). In a typical context experiment, learning conditions A and B differ on some dimension, and retrieval conditions A′ and B′ are similar on this dimension to encoding conditions A and B, respectively. The robust finding is that performance is better in conditions in which encoding and retrieval conditions match (A–A′ and B–B′) than when encoding and retrieval conditions do not match (A–B′ and B–A′). That is, maximizing the similarity of the encoding and retrieval contexts benefits retention. Significant context effects have been demonstrated with adult and infant learners in this type of encoding/retrieval paradigm (Butler & Rovee-Collier, 1989; Rovee-Collier & Dufault, 1991; Smith et al., 1978).

Learners are sensitive to background context, as demonstrated in studies showing that the same object is interpreted differently depending on the contextual background (Light & Carter-Sobell, 1970; Perry, Samuelson, & Burdinie, 2013;

Samuelson & Smith, 1998). Even young infants are capable of forming categories on the basis of shared context and form categories such as "objects found in the kitchen" (Mandler, Fivush, & Reznick, 1987; Reznick, 2000). Further, children and adults routinely use linguistic (e.g., Fisher, Klingler, & Song, 2006; Landau, Smith, & Jones, 1992) and discourse context (e.g., Akhtar, 2004) to interpret meaning. Thus, it is well established that children are sensitive to and make use of contextual cues in categorization.

Shared context may also support category formation during development. Initially shared context may do some of the work of categorization by aiding in aggregation. Because context is associated with objects in memory, contextual cues can aid in aggregating discrete instances together in memory.

Indeed, this seems to be the case. Vlach and Sandhofer (2011) experimentally tested the role of context in a novel word learning task. To control for the amount of experience children had with learning a particular category, we presented children with an artificial word learning task, known as the novel noun generalization task. In this task, children are presented with a novel object, and it is named with a novel pseudoword. The objects within each category matched each other by shape, but they differed in color and texture. Children were then tested on whether they could extend the novel word to similar objects.

We purposely chose a background context that had no relevant meaning to the objects; essentially, it was a large cloth napkin that was immaterial to any aspects of the objects. Because we did not want to draw special attention to the context, each object or set of objects was wrapped in cloth. The cloth was turned inside out and kept shut using a clip, creating the appearance of a bag. When the bag was opened during each presentation, the patterned cloth was underneath the object and, as a result, was the visual background. Using this procedure, the context changes appeared to be incidental rather than deliberately made by the experimenter.

Figure 8.1 shows the three conditions in the study. In all three conditions, children were presented with three different instances of an object category, one right after the other. During each presentation, the object was labeled two or three times (e.g., "Look at the dax! See the dax!"). Children were then shown an unlabeled distractor object (e.g., "Wow look at this!"). Finally, children were tested by presenting four objects to the child simultaneously and asking the child to select an object from the same category (e.g., "Give me the dax"). What differed between the three conditions was the amount of contextual support between training and testing. In the *match* condition, the training trials and testing context were all the same. In the *mismatch* condition, all three training context trials had the same colored and patterned cloths, but the testing context differed from that of the training trials. Finally, in the *multiple* condition, children saw a new context for every presentation.

Three age groups participated: 2.5–3 years, 3–4 years, and 4–5 years of age. The youngest children were relatively new word learners, and the oldest children had much more experience with word learning and categorization. Despite the fact that all of the children had the same level of prior experience with the novel objects (i.e., no prior experience), we found clear age-related changes in performance across the three age groups.

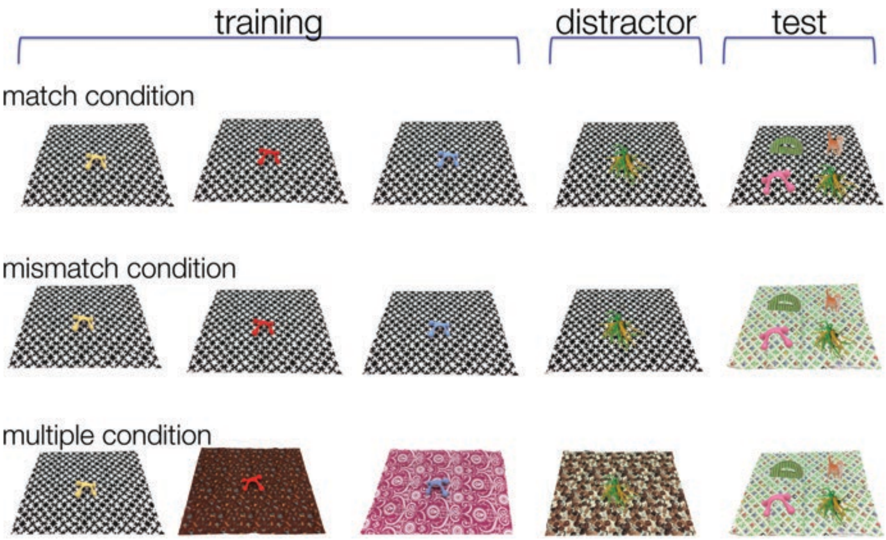


Fig. 8.1 Examples of stimuli in the match, mismatch, and multiple conditions in Vlach and Sandhofer (2011)

Figure 8.2 shows the number of category matches children made in each condition and each age group. When the training context and testing context matched, that is, the context supported aggregation between the training items and the testing items, children in all three age groups selected a large number of correct responses. However, when the training and testing contexts mismatched or varied, the youngest age group's (2.5–3 years) performance was significantly lower than when the training and testing contexts matched. Thus, 2.5- to 3-year-olds seemed to be greatly affected by context manipulation; when they were asked to generalize a category label in a new context, their performance dropped to levels that were no different from chance. By 4 years of age, the context manipulations appeared to have little effect.

Goldenberg and Sandhofer (2013) suggested that difficulty with the varied condition is due to a lack of support for the aggregation of different instances in memory. Successful category learning requires aggregating multiple object-label pairs experienced across time. Shared context may aid the youngest children in forming a category across different instances. Although a shared label may encourage categorization, the more systematically instances occur within a particular context, the more strongly these instances should be associated together. This is in part due to the compounding effects of multiple correlated cues. It is well established that when cues compound or correlate, response to the compound is greater than the response to a single cue (Kehoe, 1986; Rescorla & Coldwell, 1995). Support for aggregation might be particularly important when categories contain variation across different instances. For example, spoons can vary in color, size, and material. By aggregating instances of the experienced objects together, the relevant

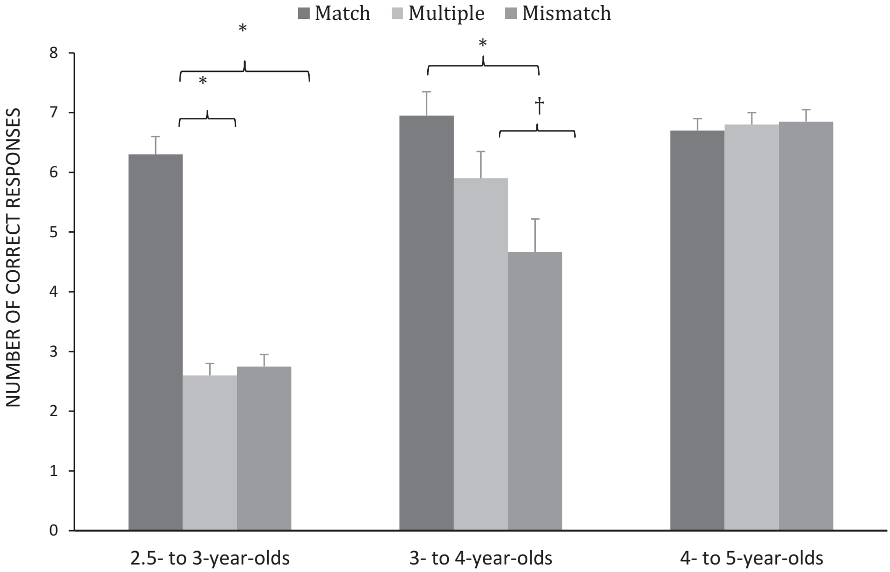


Fig. 8.2 Children's performance by age group and condition in Vlach and Sandhofer (2011). ∗ designates significantly different at $\alpha = 0.05$. † designates marginally different at $\alpha = 0.10$

properties become stronger and the irrelevant properties become weaker. In this way, aggregating across multiple instances highlights the similarities between different objects from the same category (Gentner & Namy, 1999; also see Tversky (1977)). Because the aggregation of memories for specific instances is more likely when there is a large amount of similarity, categories with some degree of variation may be the categories that most benefit from support for aggregation.

Studies with young children indicate that providing multiple redundant correlated cues, cues that point to higher-level structure, leads to greater learning than does providing a single cue. In these studies, children show stronger performance when presented with correlated cues that mutually reinforce each other than when presented with a single cue (e.g., Dueker & Needham, 2005; Thiessen & Saffran, 2003; Yoshida & Smith, 2005). For example, in Yoshida and Smith's (2005) study, when perceptual cues systematically co-occurred with linguistic categories, Japanese-speaking children showed increased performance on generalization tasks in which children hear an unknown exemplar labeled and are asked to select another object that shares that same label. Categorization studies indicate that there are an abundance of correlated cues available for children to take advantage of (Bhatt, Wilk, Hill, & Rovee-Collier, 2004; Madole, Oakes, & Cohen, 1993; Rakison, 2004; Sahni, Seidenberg, & Saffran, 2010; Younger, 2008).

In real-world learning situations, cues indicative of statistical regularity do not just occur in isolation – there is often redundancy in cues that indicate higher-level structure. For example, cues indicating the boundaries between words include stress patterns, transitional probabilities, and phonotactic constraints (Johnson & Jusczyk, 2001; Romberg & Saffran, 2010; Saffran & Kirkham, 2018). Increasing the number of cues increases the likelihood that children will learn. Sloutsky and Robinson (2013) demonstrated this in a shape rule learning study that tested the effects of redundant correlated cues. In this study, 14- to 22-month-old children learned either shape or texture in two distinct contexts that varied in the number of contextual cues (from one to four). Children showed the most learning on the contexts in which they had more correlated cues. As the number of available correlated cues decreased, learning decreased as well.

Thus, as a whole, correlated cues can be a powerful aid to learning and might be especially useful early in the learning process or in more difficult learning situations. Nonadjacent dependencies are one such example of difficult learning. Learning is difficult when information is distributed across nonadjacent presentations. Correlated cues can facilitate this type of nonadjacent learning (Gómez & Lakusta, 2004; Newport & Aslin, 2004; Romberg & Saffran, 2010). In Vlach and Sandhofer's (2011) word learning task, the category instances varied across presentations. To succeed, children needed to aggregate the features that were similar across presentations (i.e., shape). To do so, children were either provided with a single cue to aggregate instances (i.e., the label was the same across instances) or they were provided with two correlated cues (e.g., the label *and* the background context were the same across instances). Novice word learners may require more support than a single aggregative cue. In real-world learning events, this may be especially so because there can be large intervals of time between naming events; correlated contextual cues may be especially beneficial for aggregating across nonadjacent instances.

Multiple Examples Provide Support for Abstraction and Decontextualization

The best support for aggregation of category members might include highly similar objects presented in highly similar contexts. However, these types of presentations might hinder category abstraction and decontextualization. Much research has described the progression from context-bound to more abstract categories. As a whole, children appear to begin learning words with local mappings between context-bound categories and gradually build more abstract categories. This idea that children's words may initially be context-bound has been raised by a number of researchers (e.g., Hoff, 2013). Barrett (1986), for example, reported that his son initially used the word "duck" in a very specific context: only to a rubber duck while knocking the duck off the edge of the bathtub. However, he did not use the word "duck" when the toy duck was played with in other contexts. Similarly, Bloom (1973) reported that her daughter only produced "car" in response to viewing cars from her apartment window, but not for cars viewed up close or for pictures of cars in books.

The fact that children's early words show context dependency is not surprising. It is well documented that memory has context-dependent properties. The contextual effect is so strong that experiencing an object repeatedly in the same context can lead to context dependence, in which a learner fails to retrieve a memory outside the context in which it was learned. This concurs with research suggesting that children's early categories are concrete and context-bound (Huttenlocher, Smiley, & Charney, 1983; Mix, 2002; Mix & Sandhofer, 2007; Quinn, Cummins, Kase, Martin, & Weissman, 1996; Roberts, 1983; Rovee-Collier & Fagen, 1981; Tomasello, 1992, 2000) and studies demonstrating that even though children may seem to categorize objects in one task context, such as visually exploring objects, they may fail to do so in other contexts (e.g., Oakes & Madole, 2000; Younger & Furrer, 2003).

Variability in context across category presentations can protect against context dependency. In Vlach and Sandhofer's (2011) study, 2.5- and 3-year-old children who learned the category names in a single context (supporting aggregation) did significantly worse at extending the label to a new category member when it was presented in a different context (i.e., mismatch condition) than when it was presented in the same context that it had been learned (i.e., match condition). That is, the children exhibited context-dependent learning.

One possibility for this context dependency is that the object name and referent are strongly associated with the context in which it is learned. Such context dependency can be overcome by learning in varied contexts (Jones, Pascalis, Eacott, & Herbert, 2011; Smith et al., 1978). To test this, Goldenberg and Sandhofer (2013) designed a study in which children were either provided with support for aggregation, support for decontextualization, or support for both aggregation and decontextualization. In this study, 24-month-old children were taught labels for novel object categories in one type of contextual condition and were tested for category generalization in a new context.

During the learning phase of the study, children saw five category instances in which the objects within each category matched each other by shape but differed in color and texture. The objects were presented one at a time, and each was presented on its own patterned and colored cloth (the background context). In the aggregation support condition, all five of the colored cloths were identical. Thus, children received extra support for aggregation through the background cloth-correlated cue. In the decontextualization support condition, each of the five colored cloths was different. This signaled to children that the object-label pair was not associated with any specific background context. In the aggregation + decontextualization support condition, three of the colored cloths were identical to each other, and the other two cloths differed from all other cloths.

The results indicated that the children who received support for aggregation + decontextualization performed significantly higher than children who received support for only aggregation or only decontextualization. Moreover, only the children who received support for both aggregation + decontextualization performed at levels higher than chance. This suggests that novice word learners might need two types of contextual support during learning to generalize category membership to a new context. Namely, children need both support to aggregate and support to decontextualize relevant features (see Imai and Childers, this volume, for related findings from studies using progressive alignment).

A second way that variation in the environment might be important for word learning has to do with variation in the features of the objects/categories themselves. For example, if the word "purple" is exclusively paired with a purple stuffed bear, a child might form a very narrow category of purple that does not include purple cars or purple blocks. Moreover, a narrow range of categories might lead to an overly conflated understanding of category membership in which "purple" refers to things that are purple and fuzzy or perhaps simply a wrong understanding in which purple refers to fuzzy animals.

Perry, Samuelson, Malloy, and Schiffer (2010) examined how variability in exemplars during training affected category development (as evidenced by word learning) both in and out of the lab. Specifically, they examined how variability across exemplars affected (1) learning a name for a specific exemplar, (2) generalizing that name to other members of the same category, (3) extending other novel labels to other novel category members, and (4) overall vocabulary development outside the lab setting. In this longitudinal study, 18-month-olds were trained on 12 categories rarely known at their age but commonly known by older children (e.g., toothbrush, bucket).

Over several weeks, infants participated in multiple training sessions for these 12 categories. Half the infants received training on highly similar exemplars for each category (i.e., low within-category variability), and half the infants received training on highly variable exemplars for each category (i.e., high within-category variability). Overall, the 18-month-olds in the high-variability condition were better at generalizing trained category labels to new exemplars and showed faster vocabulary acceleration outside the lab relative to infants in the low-variability condition. Throughout the study, infants in both conditions began attending more to objects' shape, as this is often a reliable marker of category membership in English (see Landau, Smith, & Jones, 1988, 1998); however, infants in the low-variability condition began to overgeneralize based on shape even when it was not a category-defining feature. In contrast, infants in the high-variability condition were more discerning in their generalizations, generalizing based on shape when it was relevant but by other features (e.g., material) when shape was irrelevant. Overall, these results suggest that low within-category variability leads to stricter attentional biases, whereas high variability within a category leads to more flexibility in attention. When there was more variability in exemplars, learners were able to identify both (1) a wider range of features that were acceptable for category inclusion and (2) which features were irrelevant for determining category membership.

In sum, multiple examples are an important source of variability, and different amounts and types of variability within a category affect how the category is formed. Although high within-category variability may not facilitate immediate category learning, it does support category learning in the long term. Experience with multiple examples influences attention (and, thus, memory) and facilitates the identification of category-relevant and category-irrelevant features of subsequent exemplars. However, in generalization tasks, category members might not provide sufficient cues to prior instances because of the variation between category members, so high variability also risks a lower likelihood of aggregation. There appears to be a delicate balance between providing support for decontextualization and providing support for aggregation. However, the balance is expected to shift with development. For example, although 2-year-olds in the Goldenberg and Sandhofer (2013) study benefited when given support for decontextualization and aggregation, by 4 years of age, children can generalize in a new context regardless of the context or level of support (Vlach & Sandhofer, 2011).

Multiple Examples Provide Support for Retention and Memory

By 6 years of age, children typically know somewhere between 6000 and 10,000 words (Bloom & Markson, 1998; Carey, 2010). This is remarkable in that not only do children need to learn how words map onto meanings, but children also need to generalize the meaning to other members of the category. For example, children need to learn that the word "teapot" refers to the white ceramic teapot at home, but they also need to extend the label to new instances they are experiencing for the first time (e.g., the small pink teapot at a tea party). Moreover, children need to retain these word-referent correspondences across time so they can recall or recognize the label at a later date.

How is it that children are able to remember a particular word and recall it weeks or months later? As any student studying for the SAT can attest, knowing the word-referent correspondence at one point in time does not mean that the word will be remembered later.

Recalling word-referent correspondences across time benefits from frequency. Words that are infrequently heard or produced are the ones that are most likely to be forgotten. A long history of research has sought to describe the relationship between memory and forgetting (e.g., Ebbinghaus, 1885/1913). The forgetting curve describes memory retention over time. Although the rate of forgetting depends on multiple factors, a typical curve for adults shows that initially the rate of forgetting is rapid, such that retention for newly acquired information is reduced by 50% after a few days.

Children's word recall seems to show a similar rate of forgetting. Vlach and Sandhofer (2012) tested word retention at three time points: immediately after learning, 1 week after learning, and 1 month after learning. In this study, 3-year-olds participated in a measurement game. At one point during the game, participants heard a single novel word-object correspondence (e.g., "Here is a koba. Let's measure it"). During the test, participants were asked which object was the koba. Figure 8.3 shows the results. When tested immediately after the conclusion of the

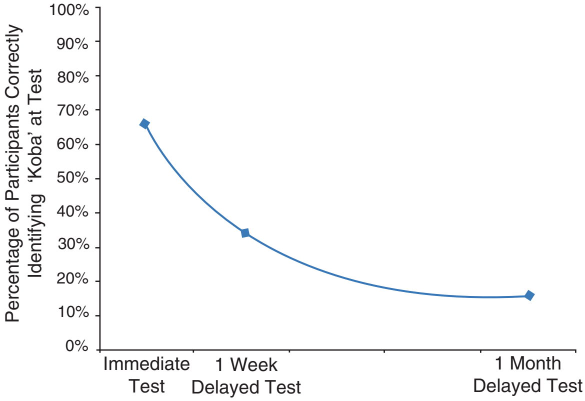


Fig. 8.3 Children's word-object retention in Vlach and Sandhofer (2012)

game, roughly 70% of children were able to identify the correct object. However, after 1 week, approximately 35% of children were able to identify the correct object, and after 1 month, only around 20% of children were able to identify the correct object.

These results indicate that without intervening reminders, children forget words over time (see also, Horst & Samuelson, 2008). However, memory can be increased by providing supports that boost memory. In a second study, Vlach and Sandhofer (2012) provided children with different amounts of memory support by increasing saliency, by providing multiple repetitions of the word at the time of learning, or by having children repeat the word themselves. These types of supports increased the percentage of children who were able to correctly identify the object at 1 week and 1 month, suggesting that word retention may be subject to the same kinds of influences as other aspects of memory.

How can children remember words if they forget them? On the one hand, forgetting words seems problematic if children need to amass a lexicon of approximately 6000 words by age six. It would seem much simpler to get to 6000 if children could hear a word once and remember it forever. However, the content of children's (and adults') lexicons is predictable based on the frequency of occurrence. At age 6, children's lexicons mostly include the 6000 highest-frequency words from their language input. Thus, children know words like "sleep" rather than lower-frequency words like "abdicate." This means that by 6 years of age, children have heard thousands of discrete instances of each of the words in their lexicons.

On the other hand, forgetting across learning events is beneficial for long-term recall (Anderson, Bjork, & Bjork, 2000). The memory of a newly learned word-object correspondence continues to decay across time. Study phase retrieval theories argue that when the word is reencountered, it reactivates the first memory trace and strengthens the memory for the word. If a word is only encountered once, barring some kind of extraordinary event that would make the experience highly salient, the memory trace should decay over time and eventually be forgotten. Study phase retrieval theories argue that forgetting between learning events is critical for the efficiency of memory. Information that reoccurs, and reoccurs frequently, gets

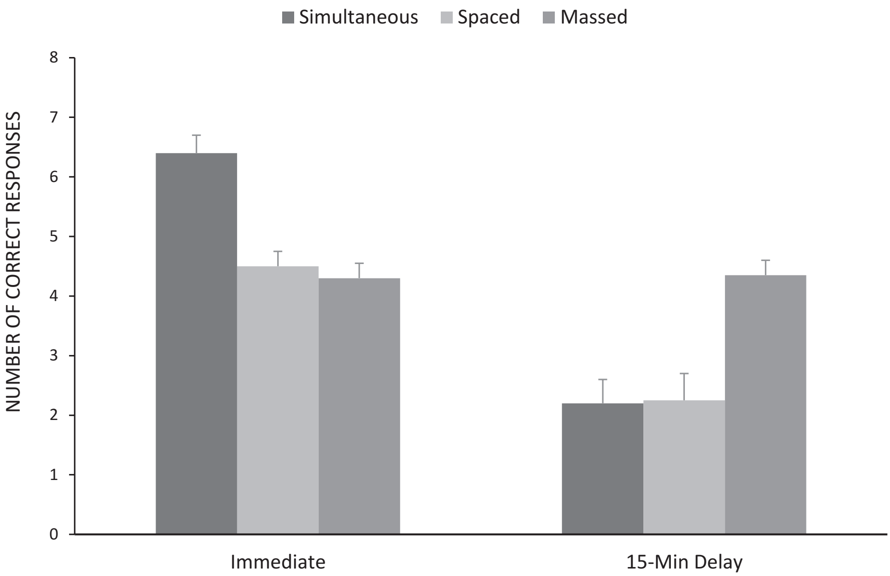


Fig. 8.4 Children's performance in the three conditions of Vlach et al. (2012)

remembered. In real-world word learning in which children experience multiple examples of words at different points in time, the memory trace should be continually reactivated and thus less subject to forgetting. Thus, the ability to retain word-referent correspondences across time so children can recall or recognize the label at a later date is supported by low-level memory processes.

Evidence for this comes from studies in which information is either presented simultaneously or spaced out across time (see Childers & Tomasello, 2002). In one study (Vlach, Ankowski, & Sandhofer, 2012), 2- to 2.5-year-old children learned the names for novel categories in one of three learning schedules: simultaneous, massed, or spaced. In all of the conditions, children saw three different examples of the category, and the amount of labeling and total presentation time was controlled. In the simultaneous schedule, all three instances of the category were presented at the same time. In the massed schedule, the three instances of the category were presented one at a time in immediate succession. In the spaced schedule, there was a 30-second play break between each of the three presentations.

Children were tested using a forced choice test in which children were asked to select another member of the category from an array of objects. When tested immediately following learning, children in all three learning schedules identified the correct object at levels above chance. The standout, though, was the simultaneous schedule. As Fig. 8.4 shows, when tested immediately, children in the simultaneous schedule identified the correct object significantly more often than children in either of the other two schedules.

However, when children were instead tested after a short delay (15 minutes), performance changed dramatically. Children who learned the words in either the simultaneous or massed schedules did not identify the correct object at levels above chance. When tested after a delay, only the children who learned the words under the spaced schedule identified the correct object at levels above chance. Children's performance in the spaced schedule was significantly higher than in the simultaneous schedule or massed schedule.

Why did the spaced schedule result in higher performance after this short delay? One explanation is that the delay between each of the presentations allowed time for forgetting (e.g., Ebbinghaus, 1885; Vlach, 2014). As a result, children in the spaced schedule engaged in deeper retrieval. This type of deeper retrieval should have two effects: (1) it should strengthen the retrievability of both the latent and current memories of the category, and (2) it should slow the rate of forgetting of memories of the category (Cepeda, Vul, Rohrer, Wixted, & Pashler, 2008; Pavlik & Anderson, 2008).

A 30-second interval is considerably smaller than the types of intervals a child encounters in everyday life; however, children also typically need to remember word-referent correspondences for longer than 15 minutes. There is a relationship between the ideal spacing interval – the amount of time that passes between learning events and the retention interval – such that retention over long periods of time benefits from longer intervals of time between each presentation in the learning phase.

There is an important caveat to this result, however. Whether children are able to benefit from spacing depends on the category to be learned and, specifically, children's past learning history. For example, English-learning children are biased to attend to objects' shape over other properties, such as color and texture (e.g., Landau et al., 1988). Slone and Sandhofer (2017) examined the effect of spacing on 2- to 3.5-year-old children's learning of categories organized by shape compared to categories organized by texture or color. Spaced presentations led to significantly better learning of shape categories, but not of texture or color categories, compared with massed presentations. One possibility is that spacing may preferentially benefit children's shape category learning because of children's past learning histories have created attentional biases to the shapes of objects over other features. These attentional biases, in turn, influence memory and retrievability.

Altogether, the results suggest that although children forget words over time, forgetting creates a desirable difficulty such that the amount of effort required to remember words improves long-term performance and, as a result, we are able to recall words over long periods of time (Bjork, 1994; Bjork & Kroll, 2015). Our memory systems tend to advantage information that repeats over time, that we ourselves need to retrieve, and/or information that occurs in highly emotionally salient circumstances. This tends to work out well for a developing lexicon. The words that are most remember are those that are most frequent and that are retrieved most often.

Children's language learning environments provide an incredibly rich backdrop for word learning (see Smith et al. (2018) for a discussion of how children's development in the first few years creates statistical data that increases in richness with the child's expanding sensorimotor abilities). In real-world learning situations, children hear words across multiple time schedules. For example, as a child gets dressed for the day, each of a child's two shoes might be simultaneously labeled. Later when the shoes are taken off and put away one at a time, each shoe might be labeled using a massed learning schedule. And because there is a delay between when shoes are labeled in the morning and later in the evening, children also experience the label using a spaced schedule. Moreover, before a child retrieves and produces their very first word, they have likely heard thousands of instances of that word spaced apart over different periods in time. In this way, a child's entire language history ensures what words children will remember and retrieve.

Conclusion

In this chapter, we have described how experiencing words in variable contexts and at different points in time provides support for (1) aggregating category members together, (2) abstracting the meaning of words, and (3) remembering and retrieving words across time. Multiple examples of words in children's language input are embedded in the process of learning words and critical for successful word learning and retention. Mechanisms such as statistical learning work particularly well with the type of repeated language exposure that children experience, because it suggests one way that low-level co-occurrence patterns in the environment can lead to word learning.

A number of studies examining word learning in young children have demonstrated that children can do quite a lot with a single example. In studies testing in-the-moment induction, an experimenter shows a child an unfamiliar object, labels the object with a novel word, and then observes which other objects the child will extend the word to while the original object is still in view (e.g., "This is a dax. Can you hand me another dax?"). These types of studies have contributed a wealth of information about how children learn words. One strong feature of these studies is that they typically consider how children's past learning history affects testing in-the-moment induction. For example, children who know many nouns for categories organized by shape (e.g., words like spoon in which the spoon shape and not the color or material determines whether an object is a spoon) are much more likely to extend words for solid objects to other objects that share the same shape than are children who know few nouns for categories organized by shape. What is noteworthy here is that even though children may demonstrate one trial learning in the laboratory, they bring an entire history of learning multiple (but different) examples of shape words to bear on the task.

In this chapter, we have discussed how multiple examples aid children's word learning through children's sensitivities to the statistical regularities of their environments. Words are plentiful in children's everyday lives, and children are clearly sensitive to, and benefit from, hearing multiple examples of words. Not only are children's first words among the words that are most frequently used in speech directed to them, but children are most likely to remember and retrieve the words that reoccur over distributed periods of time. As children gain vocabulary and experience, they are better able to learn through other mechanisms; for example, they can learn words solely through linguistic context. But even into adulthood, aggregation, abstraction, and retention remain core components of word learning.

References

Akhtar, N. (2004). Contexts of early word learning. In *Weaving a lexicon* (pp. 485–507). Cambridge, MA: MIT Press.

Anderson, M. C., Bjork, E. L., & Bjork, R. A. (2000). Retrieval-induced forgetting: Evidence for a recall-specific mechanism. *Psychonomic Bulletin & Review, 7*(3), 522–530. https://doi.org/10.3758/BF03214366

Barrett, M. (1986). Early semantic representations and early word-usage. In S. A. Kuczaj & M. Barrett (Eds.), *The development of word meaning: Progress in cognitive development research* (pp. 39–67). New York: Springer-Verlag. Retrieved from www.springer.com/us/book/9781461293262

Bhatt, R. S., Wilk, A., Hill, D., & Rovee-Collier, C. (2004). Correlated attributes and categorization in the first half-year of life. *Developmental Psychobiology, 44*(2), 103–115. https://doi.org/10.1002/dev.10164

Bjork, R. A. (1994). Memory and metamemory considerations in the training of human beings. In J. Metcalfe & A. Shimamura (Eds.), *Metacognition: Knowing about knowing* (pp. 185–205). Cambridge, MA: MIT Press.

Bjork, R. A., & Kroll, J. F. (2015). Desirable difficulties in vocabulary learning. *The American Journal of Psychology, 128*(2), 241–252. https://doi.org/10.5406/amerjpsyc.128.2.0241

Bloom, L. (1973). *One word at a time: The use of single word utterances before syntax*. The Hague, the Netherlands: Mouton.

Bloom, P., & Markson, L. (1998). Capacities underlying word learning. *Trends in Cognitive Sciences, 2*(2), 67–73. https://doi.org/10.1016/S1364-6613(98)01121-8

Butler, J., & Rovee-Collier, C. (1989). Contextual gating of memory retrieval. *Developmental Psychobiology, 22*(6), 533–552. https://doi.org/10.1002/dev.420220602

Carey, S. (2010). Beyond fast mapping. *Language Learning and Development, 6*(3), 184–205. https://doi.org/10.1080/15475441.2010.484379

Cepeda, N. J., Vul, E., Rohrer, D., Wixted, J. T., & Pashler, H. (2008). Spacing effects in learning: A temporal ridgeline of optimal retention. *Psychological Science, 19*(11), 1095–1102. https://doi.org/10.1111/j.1467-9280.2008.02209.x

Childers, J. B., & Tomasello, M. (2002). Two-year-olds learn novel nouns, verbs, and conventional actions from massed or distributed exposures. *Developmental Psychology, 38*(6), 967–978.

Dueker, G., & Needham, A. (2005). Infants' object category formation and use: Real-world context effects on category use in object processing. *Visual Cognition, 12*(6), 1177–1198. https://doi.org/10.1080/13506280444000706

Ebbinghaus, H. E. (1885). *Memory: A contribution to experimental psychology*. New York: Dover. Retrieved from http://search.ebscohost.com/login.aspx?direct=true&db=a9h&AN=95118792&site=ehost-live

Fisher, C., Klingler, S. L., & Song, H. (2006). What does syntax say about space? 2-year-olds use sentence structure to learn new prepositions. *Cognition, 101*(1), B19–B29. https://doi.org/10.1016/j.cognition.2005.10.002

Gentner, D., & Namy, L. L. (1999). Comparison in the development of categories. *Cognitive Development, 14*(4), 487–513. https://doi.org/10.1016/S0885-2014(99)00016-7

Godden, D. R., & Baddeley, A. D. (1975). Context-dependent memory in two natural environments: On land and underwater. *British Journal of Psychology, 66*(3), 325–331. https://doi.org/10.1111/j.2044-8295.1975.tb01468.x

Goldenberg, E. R., & Sandhofer, C. M. (2013). Same, varied, or both? Contextual support aids young children in generalizing category labels. *Journal of Experimental Child Psychology, 115*(1), 150–162. https://doi.org/10.1016/j.jecp.2012.11.011

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. https://doi.org/10.1111/j.1467-7687.2004.00381.x

Graf Estes, K., Evans, J. L., Alibali, M. W., & Saffran, J. R. (2007). Can infants map meaning to newly segmented words?: Statistical segmentation and word learning. *Psychological Science, 18*(3), 254–260. https://doi.org/10.1111/j.1467-9280.2007.01885.x

Hart, B. (1991). Input frequency and children's first words. *First Language, 11*(32), 289–300. https://doi.org/10.1177/014272379101103205

Hart, B., & Risley, T. R. (1995). *Meaningful differences in the everyday experience of young American children*. Baltimore: Paul H. Brookes Publishing Co..

Heibeck, T. H., & Markman, E. M. (1987). Word learning in children: An examination of fast mapping. *Child Development, 58*(4), 1021–1034. https://doi.org/10.2307/1130543

Hoff, E. (2003). The specificity of environmental influence: Socioeconomic status affects early vocabulary development via maternal speech. *Child Development, 74*(5), 1368–1378.

Hoff, E. (2013). *Language development* (5th ed.). Belmont, CA: Cengage Learning. Retrieved from https://smile.amazon.com/Language-Development-published-Cengage-Learning/dp/B00E6T5SV2/ref=sr_1_sc_1?ie=UTF8&qid=1519158927&sr=8-1-spell&keywords=Erika+Hoff%27s+LANGUAGE+DEVELOPMENT%2C+5th+Editio

Hoff, E., & Naigles, L. (2002). How children use input to acquire a lexicon. *Child Development, 73*(2), 418–433. https://doi.org/10.1111/1467-8624.00415

Horst, J. S., & Samuelson, L. K. (2008). Fast mapping but poor retention by 24-month-old infants. *Infancy, 13*(2), 128–157. https://doi.org/10.1080/15250000701795598

Hurtado, N., Marchman, V. A., & Fernald, A. (2008). Does input influence uptake? Links between maternal talk, processing speed and vocabulary size in Spanish-learning children. *Developmental Science, 11*(6), F31–F39. https://doi.org/10.1111/j.1467-7687.2008.00768.x

Huttenlocher, J., Haight, W., Bryk, A., Seltzer, M., & Lyons, T. (1991). Early vocabulary growth: Relation to language input and gender. *Developmental Psychology, 27*(2), 236–248. https://doi.org/10.1037/0012-1649.27.2.236

Huttenlocher, J., Smiley, P., & Charney, R. (1983). Emergence of action categories in the child: Evidence from verb meanings. *Psychological Review, 90*(1), 72–93. https://doi.org/10.1037/0033-295X.90.1.72

Johnson, E. K., & Jusczyk, P. W. (2001). Word segmentation by 8-month-olds: When speech cues count more than statistics. *Journal of Memory and Language, 44*(4), 548–567. https://doi.org/10.1006/jmla.2000.2755

Jones, E. J. H., Pascalis, O., Eacott, M. J., & Herbert, J. S. (2011). Visual recognition memory across contexts. *Developmental Science, 14*(1), 136–147. https://doi.org/10.1111/j.1467-7687.2010.00964.x

Kehoe, J. E. (1986). Summation and configuration in conditioning of the rabbit's nictitating membrane response to compound stimuli. *Journal of Experimental Psychology: Animal Behavior Processes, 12*, 186–195. https://doi.org/10.1037/0097-7403.14.3.320

Landau, B., Smith, L., & Jones, S. (1998). Object shape, object function, and object name. *Journal of Memory and Language, 38*(1), 1–27. https://doi.org/10.1006/jmla.1997.2533

Landau, B., Smith, L. B., & Jones, S. (1992). Syntactic context and the shape bias in children's and adults' lexical learning. *Journal of Memory and Language, 31*(6), 807–825. https://doi.org/10.1016/0749-596X(92)90040-5

Landau, B., Smith, L. B., & Jones, S. S. (1988). The importance of shape in early lexical learning. *Cognitive Development, 3*(3), 299–321. https://doi.org/10.1016/0885-2014(88)90014-7

Light, L. L., & Carter-Sobell, L. (1970). Effects of changed semantic context on recognition memory. *Journal of Verbal Learning and Verbal Behavior, 9*(1), 1–11. https://doi.org/10.1016/S0022-5371(70)80002-0

Madole, K. L., Oakes, L. M., & Cohen, L. B. (1993). Developmental changes in infants' attention to function and form-function correlations. *Cognitive Development, 8*(2), 189–209. https://doi.org/10.1016/0885-2014(93)90014-V

Mandler, J. M., Fivush, R., & Reznick, J. S. (1987). The development of contextual categories. *Cognitive Development, 2*(4), 339–354. https://doi.org/10.1016/S0885-2014(87)80012-6

Mix, K. S. (2002). The construction of number concepts. *Cognitive Development, 17*(3), 1345–1363. https://doi.org/10.1016/S0885-2014(02)00123-5

Mix, K. S., & Sandhofer, C. M. (2007). Do we need a number sense? In M. J. Roberts (Ed.), *Integrating the Mid* (pp. 293–326). Hove, UK: Psychology Press.

Moerk, E. L. (1980). Relationships between parental input frequencies and children's language acquisition: A reanalysis of Brown's data∗. *Journal of Child Language, 7*(1), 105–118. https://doi.org/10.1017/S0305000900007054

Newport, E. L., & Aslin, R. N. (2004). Learning at a distance I. Statistical learning of non-adjacent dependencies. *Cognitive Psychology, 48*(2), 127–162. https://doi.org/10.1016/S0010-0285(03)00128-2

Oakes, L. M., & Madole, K. L. (2000). The future of infant categorization research: A process-oriented approach. *Child Development, 71*(1), 119–126. https://doi.org/10.1111/1467-8624.00125

Pavlik, P. I., & Anderson, J. R. (2008). Using a model to compute the optimal schedule of practice. *Journal of Experimental Psychology: Applied, 14*(2), 101–117. https://doi.org/10.1037/1076-898X.14.2.101

Perry, L. K., Samuelson, L. K., & Burdinie, J. B. (2013). Highchair philosophers: The impact of seating context-dependent exploration on children's naming biases. *Developmental Science, 17*(5), 757–765. https://doi.org/10.1111/desc.12147

Perry, L. K., Samuelson, L. K., Malloy, L. M., & Schiffer, R. N. (2010). Learn locally, think globally: Exemplar variability supports higher-order generalization and word learning. *Psychological Science, 21*(12), 1894–1902.

Quinn, P. C., Cummins, M., Kase, J., Martin, E., & Weissman, S. (1996). Development of categorical representations for above and below spatial relations in 3- to 7-month-old infants. *Developmental Psychology, 32*(5), 942–950. https://doi.org/10.1037/0012-1649.32.5.942

Rakison, D. H. (2004). Infants' sensitivity to correlations between static and dynamic features in a category context. *Journal of Experimental Child Psychology, 89*(1), 1–30. https://doi.org/10.1016/j.jecp.2004.06.001

Rescorla, R. A., & Coldwell, S. E. (1995). Summation in autoshaping. *Animal Learning & Behavior, 23*(3), 314–326. https://doi.org/10.3758/BF03198928

Reznick, J. S. (2000). Interpreting infant conceptual categorization. *Journal of Cognition and Development, 1*(1), 63–66. https://doi.org/10.1207/S15327647JCD0101N_7

Roberts, K. (1983). Comprehension and production of word order in stage I. *Child Development, 54*(2), 443–449. https://doi.org/10.2307/1129705

Romberg, A. R., & Saffran, J. R. (2010). Statistical learning and language acquisition. *Wiley Interdisciplinary Reviews: Cognitive Science, 1*(6), 906–914. https://doi.org/10.1002/wcs.78

Rovee-Collier, C., & Dufault, D. (1991). Multiple contexts and memory retrieval at three months. *Developmental Psychobiology, 24*(1), 39–49. https://doi.org/10.1002/dev.420240104

Rovee-Collier, C. K., & Fagen, J. W. (1981). The retrieval of memory in early infancy. *Advances in Infancy Research, 1*, 225–254.

Roy, B. C., Frank, M. C., DeCamp, P., Miller, M., & Roy, D. (2015). Predicting the birth of a spoken word. *Proceedings of the National Academy of Sciences, 112*(41), 12663–12668. https://doi.org/10.1073/pnas.1419773112

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science, 274*(5294), 1926–1928.

Saffran, J. R., & Kirkham, N. Z. (2018). Infant statistical learning. *Annual Review of Psychology, 69*(1), 181–203. https://doi.org/10.1146/annurev-psych-122216-011805

Sahni, S. D., Seidenberg, M. S., & Saffran, J. R. (2010). Connecting cues: Overlapping regularities support cue discovery in infancy. *Child Development, 81*(3), 727–736. https://doi.org/10.1111/j.1467-8624.2010.01430.x

Samuelson, L. K., & Smith, L. B. (1998). Memory and attention make smart word learning: An alternative account of Akhtar, Carpenter, and Tomasello. *Child Development, 69*(1), 94–104.

Scott, R. M., & Fisher, C. (2012). 2.5-Year-olds use cross-situational consistency to learn verbs under referential uncertainty. *Cognition, 122*(2), 163–180. https://doi.org/10.1016/j.cognition.2011.10.010

Slone, L. K., & Sandhofer, C. M. (2017). Consider the category: The effect of spacing depends on individual learning histories. *Journal of Experimental Child Psychology, 159*, 34–49. https://doi.org/10.1016/j.jecp.2017.01.010

Sloutsky, V. M., & Robinson, C. W. (2013). Redundancy matters: Flexible learning of multiple contingencies in infants. *Cognition, 126*(2), 156–164. https://doi.org/10.1016/j.cognition.2012.09.016

Smith, L. B., Jayaraman, S., Clerkin, E., & Yu, C. (2018). The developing infant creates a curriculum for statistical learning. *Trends in Cognitive Sciences, 22*(4), 325–336. https://doi.org/10.1016/j.tics.2018.02.004

Smith, L. B., Suanda, S. H., & Yu, C. (2014). The unrealized promise of infant statistical word–referent learning. *Trends in Cognitive Sciences, 18*(5), 251–258. https://doi.org/10.1016/j.tics.2014.02.007

Smith, L. B., & Yu, C. (2013). Visual attention is not enough: Individual differences in statistical word-referent learning in infants. *Language Learning and Development, 9*(1), 25–49. https://doi.org/10.1080/15475441.2012.707104

Smith, S. M., Glenberg, A., & Bjork, R. A. (1978). Environmental context and human memory. *Memory & Cognition, 6*(4), 342–353. https://doi.org/10.3758/BF03197465

Suanda, S. H., Mugwanya, N., & Namy, L. L. (2014). Cross-situational statistical word learning in young children. *Journal of Experimental Child Psychology, 126*, 395–411. https://doi.org/10.1016/j.jecp.2014.06.003

Thiessen, E. D., & Saffran, J. R. (2003). When cues collide: Use of stress and statistical cues to word boundaries by 7- to 9-month-old infants. *Developmental Psychology, 39*(4), 706–716. https://doi.org/10.1037/0012-1649.39.4.706

Tomasello, M. (1992). *First verbs: A case study of early grammatical development*. Cambridge: Cambridge University Press. Retrieved from https://www.cambridge.org/core/journals/journal-of-child-language/article/div-classtitletomasellom-first-verbs-a-case-study-of-early-grammatical-development-cambridge-cambridge-university-press-1992-pp-ix-373div/36680A6ED73141AECB77286404D7259C

Tomasello, M. (2000). Do young children have adult syntactic competence? *Cognition, 74*(3), 209–253.

Tulving, E. (1972). Episodic and semantic memory. In *Organization of memory* (p. xiii, 423). Oxford: Academic Press.

Tummeltshammer, K., Amso, D., French, R. M., & Kirkham, N. Z. (2017). Across space and time: Infants learn from backward and forward visual statistics. *Developmental Science, 20*(5), e12474. https://doi.org/10.1111/desc.12474

Tversky, A. (1977). Features of similarity. *Psychological Review, 84*(4), 327–352. https://doi.org/10.1037/0033-295X.84.4.327

Vlach, H., & Sandhofer, C. M. (2012). Fast mapping across time: Memory processes support children's retention of learned words. *Frontiers in Psychology, 3*, 46. https://doi.org/10.3389/fpsyg.2012.00046

Vlach, H. A. (2014). The spacing effect in Children's generalization of knowledge: Allowing children time to forget promotes their ability to learn. *Child Development Perspectives, 8*(3), 163–168. https://doi.org/10.1111/cdep.12079

Vlach, H. A., Ankowski, A. A., & Sandhofer, C. M. (2012). At the same time or apart in time? The role of presentation timing and retrieval dynamics in generalization. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 38*(1), 246–254. https://doi.org/10.1037/a0025260

Vlach, H. A., & Johnson, S. P. (2013). Memory constraints on infants' cross-situational statistical learning. *Cognition, 127*(3), 375–382. https://doi.org/10.1016/j.cognition.2013.02.015

Vlach, H. A., & Sandhofer, C. M. (2011). Developmental differences in children's context-dependent word learning. *Journal of Experimental Child Psychology, 108*(2), 394–401. https://doi.org/10.1016/j.jecp.2010.09.011

Wu, R., Gopnik, A., Richardson, D. C., & Kirkham, N. Z. (2011). Infants learn about objects from statistics and people. *Developmental Psychology, 47*(5), 1220–1229. https://doi.org/10.1037/a0024023

Yoshida, H., & Smith, L. B. (2005). Linguistic cues enhance the learning of perceptual cues. *Psychological Science, 16*(2), 90–95. https://doi.org/10.1111/j.0956-7976.2005.00787.x

Younger, B. A. (2008). Developmental change in infant categorization: The perception of correlations among facial features. *Child Development, 63*(6), 1526–1535. https://doi.org/10.1111/j.1467-8624.1992.tb01712.x

Younger, B. A., & Furrer, S. D. (2003). A comparison of visual familiarization and object-examining measures of categorization in 9-month-old infants. *Infancy, 4*(3), 327–348. https://doi.org/10.1207/S15327078IN0403_02

Yu, C., & Smith, L. B. (2007). Rapid word learning under uncertainty via cross-situational statistics. *Psychological Science, 18*(5), 414–420. https://doi.org/10.1111/j.1467-9280.2007.01915.x

Yu, C., & Smith, L. B. (2011). What you learn is what you see: Using eye movements to study infant cross-situational word learning. *Developmental Science, 14*(2), 165–180. https://doi.org/10.1111/j.1467-7687.2010.00958.x

Yurovsky, D., Yu, C., & Smith, L. B. (2013). Competitive processes in cross-situational word learning. *Cognitive Science, 37*(5), 891–921. https://doi.org/10.1111/cogs.12035

Chapter 9
Mechanisms for Evaluating Others' Reliability When Learning Novel Words

David M. Sobel, Elena Luchkina, and Kristen Tummeltshammer

Abstract Word learning is a social act. Because there is an arbitrary relation between words and their meaning, children must learn words from other people. Other people, however, are not always reliable sources of knowledge. People can be ignorant, hold false beliefs, or simply be deceptive. How do children evaluate the reliability of sources of knowledge for word learning? This chapter investigates the possibility that children possess multiple mechanisms for evaluating such reliability and possess such mechanisms very early in development. We suggest that infants not only track the accuracy of others using statistical learning mechanisms but also incorporate their existing knowledge of the world into judgments of reliability to make judicious inferences about the knowledge of a speaker and the pragmatics of a communicative act. Moreover, we suggest that as children get older, both low-level associative mechanisms and higher-level cognitive processes influence the way in which children track and use others' reliability as sources of knowledge.

Over the past 10 years, there has been growing interest in young children's capacity to learn from other people's verbal testimony. There is now ample evidence that young children can evaluate the reliability of others' verbal information and use that evaluation to make judgments about their knowledge of novel words (see, e.g., Harris, 2012; Mills, 2013; Sobel & Kushnir, 2013 for reviews). This research focuses on children's judgments about the referents of novel labels generated by speakers who, in the past, have differed in their accuracy of labeling familiar objects. Preschoolers appear capable of tracking that history of accuracy and generalizing from others' novel labels judiciously (e.g., Clément, Koenig, & Harris, 2004;

The first author was supported by NSF 1420548 and 1661068 during the writing of this chapter and while this research was conducted. The third author was supported by NIMH F32MH108278-01 during the writing of this chapter. We thank Dima Amso, Kathleen Corriveau, Natasha Kirkham, and James Morgan for helpful discussion about the research discussed in this chapter.

D. M. Sobel (✉) · E. Luchkina · K. Tummeltshammer
Brown University, Providence, RI, USA
e-mail: david_sobel_1@brown.edu

J. B. Childers (ed.), *Language and Concept Acquisition from Infancy Through Childhood*, https://doi.org/10.1007/978-3-030-35594-4_9

Koenig, Clément, & Harris, 2004). They also appear to generalize others' category of knowledge from accuracy information (e.g., Koenig & Harris, 2005a; Kushnir, Vredenburgh, & Schneider, 2013; Sobel & Corriveau, 2010), and by the age of 4, children are sensitive to the probability with which an individual is likely to make an error (see, e.g., Pasquini, Corriveau, Koenig, & Harris, 2007). Further, preschoolers can use their epistemic knowledge to discount accuracy information and thus integrate accuracy information with other kinds of knowledge to infer whether another is a reliable[1] source of information (e.g., Nurmsoo & Robinson, 2009). Assessing the reliability of others' information is important not only for learning the meaning of words but also as a basis for social learning and cultural transmission (e.g., Bergstrom, Moehlmann, & Boyer, 2006; Harris & Koenig, 2006; Mascaro & Sperber, 2009).

Elsewhere, we have documented a parallel between how children learn causal relations from observed data and how children establish others' reliability for social learning (Sobel & Kushnir, 2013). Early in development, children's causal reasoning is based primarily on associative information. Infants possess statistical learning mechanisms for registering event structure (e.g., Haith, 1993; Kirkham, Slemmer, & Johnson, 2002). With experience, children integrate existing conceptual and social knowledge with this associative information to make more sophisticated causal inferences or to discount regularities that denote spurious associations (Denison & Xu, 2010; Madole & Cohen, 1995; Sobel & Kirkham, 2006, 2007, 2012). In this chapter, our goal is to elucidate and support a similar mechanism for assessing reliability in others for selective word learning. Children initially begin by registering statistical regularity among informants and their accuracy (see Chaps. 2, 4, and 8). Such a mechanism forms the basis of assessing reliability but is quickly incorporated with a more rational system that integrates one's existing knowledge with accuracy information to make reliability judgments.

To outline the arguments in this chapter, we begin by considering children's ability to track the accuracy of others' information. Our goal is to show that infants have the capacity to track accuracy information within the first year of life. They then begin to integrate other pieces of existing knowledge into their judgments. Even in infancy, existing knowledge allows children to make judicious inferences about both the knowledge of the speaker and the pragmatics of the communicative act and either update their knowledge with the new information or resist it based on their inferences. We conclude by showing that even though this more sophisticated learning mechanism explains children's selective word learning, children also retain the associative information on which this more rational system is built.

[1] Throughout this chapter, we will use the term *accuracy* to refer to whether informants generate appropriate labels for familiar objects or, more generally, appropriate or expected information. We will use the term *reliability* to refer to whether informants are expected to provide consistent information over time. That is, reliability can be based on past accuracy, but reliability could emerge from other factors as well if accuracy information is unavailable.

Associative Origins of Selective Learning

Infants come into the world equipped with statistical learning capacities that allow them to detect regularity in a complex, noisy natural environment (e.g., Kidd, Piantadosi, & Aslin, 2014; Kirkham et al., 2002; Haith, 1993; Saffran, Aslin, & Newport, 1996; Tummeltshammer & Kirkham, 2013). Such a learning mechanism offers a means by which children might come to recognize others as accurate sources of information. By tracking accuracy over time, children may establish a likelihood estimate of an informant's future accuracy, which may be applied to future information generated by that informant.

Early studies of selective trust were often motivated by this interpretation (Clément et al., 2004). Such an account also explains why children use the accuracy information informants generate to make novel inferences about other information those informants generate, even when such generalization is not warranted (e.g., halo effects, such as making inferences about which informant will be more prosocial based on who labels familiar objects correctly; see Brosseau-Liard & Birch, 2010). More notably, this interpretation is consistent with arguments made by Jaswal et al. (2010; following philosophical claims by Reid, 1764), who suggest that because informants usually provide us with accurate information, children rely on the testimony of others because they have generalized this association. Because people usually provide accurate information, we assume that the information people provide us is accurate. Recently, Heyes (2015) extended this argument to suggest that social learning of nonverbal information (e.g., imitation) can be explained by associative mechanisms. While she is agnostic as to whether associative mechanisms contribute to selective word learning, we suspect that such a mechanism is consistent with the views we articulated above.

Early trust in speakers' accuracy is present in very young children. By 16 months, children expect interlocutors to provide appropriate labels for objects; when children hear an informant generate an inaccurate label for a familiar object, they look longer at that informant (Koenig & Echols, 2003). In another study, when introduced to a single speaker who either labeled familiar objects accurately and inaccurately or feigned ignorance about the labels of the objects (among other conditions), 2-year-olds retained a novel label generated by that speaker for a novel object equally often among these three conditions (Krogh-Jespersen & Echols, 2012). Jaswal (2010; Jaswal et al., 2010) suggested that 3-year-olds struggle to learn that individuals who are consistently inaccurate sources of information about the location of objects are unreliable; 4-year-olds, in contrast, learn to distrust such inaccurate informants relatively quickly. Such findings parallel findings from other studies contrasting the accuracy of two speakers (Clements et al., 2004; Koenig et al., 2004); the 3-year-olds in these studies tended to use accuracy information as the basis for word learning less frequently than did older children. These findings led Koenig and Harris (2005b) to suggest that young children "fail to effectively inhibit their default assumption that other people's beliefs are true" (p. 458) or, more generally, that children have a specific "robust bias to trust" (Jaswal et al., 2010, p. 1541) others' verbal information.

This conclusion, however, is controversial as there is also evidence that very young children track accuracy information and use that information as the basis for selective learning. Koenig and Woodward (2010) demonstrated that there were some conditions under which 2-year-olds would rely on others' past accuracy selectively. This assertion was supported by Brooker and Poulin-Dubois (2013), who found that 18-month-olds could track and learn from others selectively under certain simplified circumstances. Both of these findings, however, rely on slightly different methods for assessing reliability. They presented children with both an accurate and inaccurate speaker, establishing the accuracy of the two informants on the same object. In Krogh-Jespersen and Echols' procedure, children only interact with a single informant who is either accurate or inaccurate and do not see a contrast between an accurate and inaccurate speaker. Vanderbilt, Heyman, and Liu (2014) showed that when two speakers labeled the same object differently, that contrast facilitates children's reliability judgments; when children observe only one speaker label objects accurately or inaccurately, it isn't until much later in development (usually around age 4) that children can discount an inaccurate speaker's labels. The Krogh-Jespersen and Echols' procedure seems much more ecologically valid; how often do children – particularly very young children tasked with learning words for the first time – directly observe two people label the same object differently?

Given this concern, we (Luchkina, Sobel, & Morgan, 2018, Experiment 1) modified these previous investigations in several ways. We introduced 18-month-olds to videos of two informants. The first informant labeled a set of familiar[2] objects accurately while the second labeled a different set of familiar objects inaccurately (using labels that would be known to the child). Thus, while children observed an accurate and an inaccurate speaker, these two speakers never labeled the same object and never presented the contrast that Vanderbilt et al. suggested was critical to improving children's sensitivity to accuracy. Then, between subjects, one of the two informants brought out two novel objects and gives each object a novel label (just like in the Krogh-Jespersen and Echols' procedure).

Using an intermodal preferential-looking paradigm (IPLP), children were then assessed on the mapping of the referents of these labels. On *familiar object* trials, children saw two familiar objects and were asked to identify one of the familiar objects using a familiar label (so if children saw a ball and a shoe, they might hear, "Look, a shoe! Where's the shoe?"). Critically, the speaker on all the test trials was novel – this way, we could test whether infants generalized the label for the object, as opposed to learn that a particular speaker simply uses these labels. Regardless of whether the accurate or inaccurate speaker labeled novel objects with novel names, children should look to the appropriate object on familiar object trials. On *novel object* trials, children saw the two novel objects that had previously been labeled by the speaker, and a novel voice asked them to identify one of the objects using a label that had been generated by the speaker. Here, the expectation was that children

[2] Familiar and known labels were confirmed from pilot testing and parental reports of toddlers' productive and receptive vocabularies filled out before the experiment commenced.

would differentiate between the accurate and the inaccurate speaker conditions. Unlike most other selective word learning experiments that recruit older children and used manual forced-choice procedure, in our study, toddler's word learning outcomes were evaluated based on looking time in an intermodal preferential-looking procedure. Because we tested 18-month-olds, a manual task may have been less sensitive to their word knowledge for various information-processing reasons (e.g., limited understanding of the experimenter's instructions or the inability to suppress the desire to pick up a more attractive object).

Figure 9.1a shows the pattern of looking times to the target objects. On the trials in which children saw two familiar objects and heard a familiar label, they looked at the appropriate object above chance levels, regardless of whether the accurate or inaccurate speaker labeled the novel objects. This suggests that children can participate in the task and recognize familiar labels for familiar objects regardless of condition. What is more critical is how children performed when they were shown the two novel objects and heard one of the novel labels. On these trials, children's looking time differed between the trials, showing preferential looking only in the accurate condition. Only when a previously accurate informant generated novel labels for novel objects did 18-month-olds generalize those labels to other speakers.

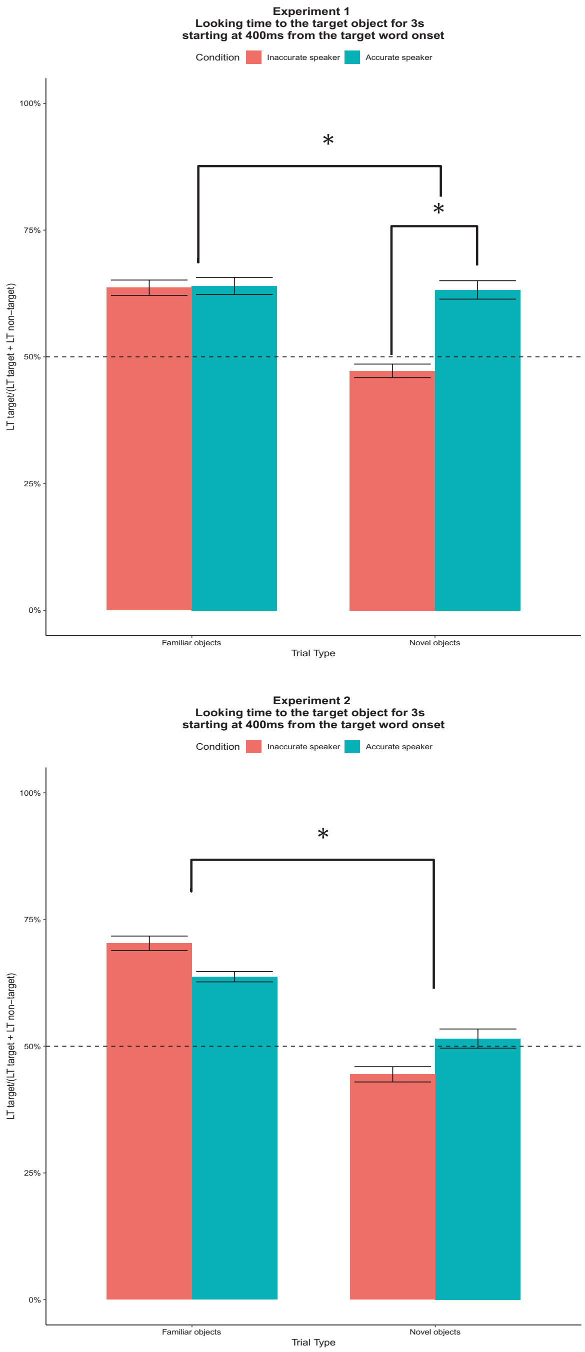
These results suggest that toddlers are able to discount information from inaccurate speakers even in the more ecologically valid situations as those suggested by Krogh-Jespersen and Echols (2012) in which only one speaker provides information about their accuracy. Moreover, there is some evidence that children are making this inference based on more than just associative mechanisms, as children respond to novel speakers, and not the same informant. This suggests that they generalize the label and not just that a particular speaker calls an object with a particular label. What this work does not show, however, is that children are really using more sophisticated mechanisms for selective word learning – a point we will discuss in the next section.

Do More Sophisticated Mechanisms Underlie Children's Social Learning: A Gaze Following Example

Previous findings on children's selective word learning have suggested that toddlers might not be able to discount novel labels for novel objects even if speakers were unreliable when faced with only a single speaker. The Luchkina et al. (2018) findings suggest that 18-month-olds can be judicious in their evaluation of others' information for word learning, even when there is only information presented by a single speaker.

Such findings are consistent with multiple studies outside of the domain of word learning. Sixteen-month-olds distinguish between individuals who reliably gaze where they search for hidden objects and individuals who do not (Poulin-Dubois & Chow, 2009). Fourteen-month-olds imitate a competent but not an incompetent

Fig. 9.1 Proportion of toddlers' looking time at the target objects in Luchkina et al. (2018) when familiarized by statements (Experiment 1, top figure) or questions (Experiment 2, bottom figure)



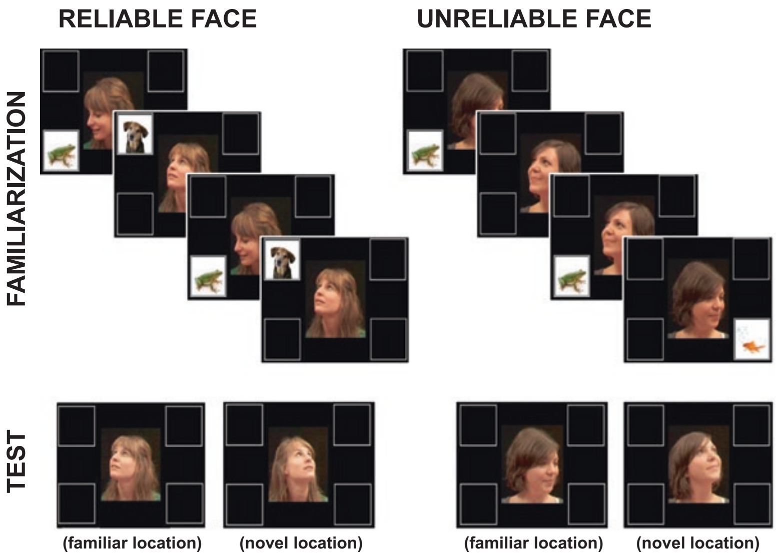


Fig. 9.2 Familiarization and test trials presented in Tummeltshammer et al. (2014)

model (Zmyj, Buttelmann, Carpenter, & Daum, 2010), and even 12-month-olds prefer objects demonstrated by individuals who appear more knowledgeable about the objects (Stenberg, 2013). All of these data suggest that infants use the reliability of others' social cues to draw inferences about the world.

What are the origins of these abilities? Tummeltshammer, Wu, Sobel, and Kirkham (2014) demonstrated that 8-month-olds could track the accuracy of potential informants in a gaze-monitoring paradigm and use that reliability information judiciously to modify future behavior. Their procedure is shown in Fig. 9.2. They introduced infants to two individuals on a screen. During familiarization, each individual turned and gazed at particular locations in space over four different trials: one always predicted the appearance of an interesting video in that location, while the other was only accurate in predicting the location of the video on one of four trials. At test, the informants turned and gazed at locations where they had predicted the onset of the video during familiarization as well as locations they had never predicted before (*test* and *generalization* trials, respectively). On both types of trials, infants followed the gaze of the reliable informant but not the unreliable informant. Thus, as young as 8 months, infants are judiciously responding to the accuracy information they observe.

One interpretation of all of the data presented in the previous section is that infants are using the statistical learning capacities available to them early in infancy to make judgments about others' reliability. In the word learning experiments, children might associate the word they hear, the object they see, and the speaker who utters the word together to form a unit; then they use that unit – that is, those co-occurrences – to make a judgment about the valence of that speaker: accurate speakers co-occur with reliable novel information, inaccurate speakers less so. Preference about whom to trust is based on preferring accuracy or avoiding inaccuracy (e.g., Lucas & Lewis, 2010).

There are difficulties with this argument. Registering statistical regularities among events alone does not provide children with the information necessary to make the same kinds of inferences as older children. Consider work described by Madole and Cohen (1995). They presented 14- and 18-month-olds with stimuli that had specific links between different parts of an object and the functions of those parts. At test, results showed that 14-month-olds registered any kind of statistical regularity among object parts and their functions, including correlating the shape of one part of an object with the function of another part (a mechanistically implausible regularity). Eighteen-month-olds, given the same exposure, did not learn these types of correlations. The development here involves learning to ignore regularities that are not salient because of their mechanistic implausibility. Children's knowledge of causal mechanisms allows them to recognize that this particular regularity is not mechanistically plausible and thus should be ignored.[3] More generally, the existing knowledge that children have affects how they might attend to and register correlational information (or affects the output of their statistical learning process).

Knowing that faces can cue the spatial locations of events might be critical for 8-month-olds to interpret the regularity in Tummeltshammer et al.'s (2014) procedure. Indeed, we suggested that infants were incorporating the accuracy information they observed with their own knowledge of faces as social cues. To investigate this possibility, in a second experiment, Tummeltshammer et al. (2014) replicated their procedure using dynamic, arrow-like stimuli instead of faces. Unlike faces, these stimuli were completely novel, but the regularities presented in the study were sufficient to generalize the accuracy information judiciously. Indeed, 8-month-olds registered a difference between the two arrows for locations that had previously been shown to them but could not generalize this accuracy to novel locations. This finding suggests that they registered the associative information they observed, but unlike faces, that statistical regularity was not salient enough for them to be able to generalize to novel situations. Integrating this finding with the first experiment with faces suggests that 8-month-olds potentially integrate the associative accuracy information they observe with their experience in following gaze from faces or their prior knowledge of faces as informative cues to locations in space. These experiences potentially afforded the face stimuli to be more salient for tracking accuracy than the arrow stimuli.

To test this hypothesis, we replicated the study using faces with two additional age groups: 5-month-olds and 12- to 13-month-olds (Tummeltshammer, Sobel, & Kirkham, 2019). Children's responses across these ages are shown in Fig. 9.3. Both age groups interpreted the face stimuli differently than did the 8-month-olds. The younger infants essentially treated the face stimuli in the same way as the 8-month-

[3] Or is registered but ignored; these alternatives have not been investigated empirically. We favor the explanation presented in the main text but recognize that this is an empirical question.

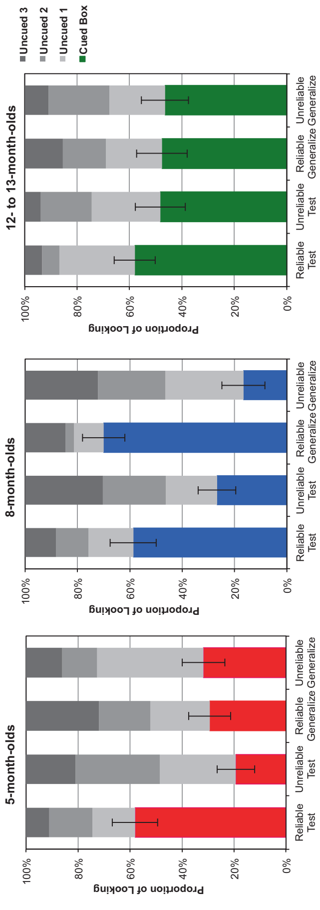


Fig. 9.3 Proportions of looking to cued and uncued locations on test and generalization trials in Tummeltshammer et al. (2019)

olds treated the arrow stimuli. They were able to follow the gaze of the accurate face when it cued a familiar location, but they could not generalize this accuracy to novel locations. In contrast to the 8-month-olds, they did not engage in any kind of systematic response when given the inaccurate face. Five-month-olds (of course) do not have the same experience or capacity to follow gaze as do older infants, and as a result, the accuracy information presented here might not have been sufficient for them to use to generalize. Five-month-olds also may not have the capacity to interpret these stimuli in any way other than through statistical regularity. This would be consistent with the development that Sobel & Kirkham (2006, 2007) found between 5- and 8-month-olds regarding their causal reasoning abilities. They found that 8-month-olds registered conditional independence among events in ways that were inconsistent with their just recognizing statistical regularity among events. Five-month-olds, in contrast, responded to the same stimuli in a manner consistent with their registering the statistical regularity but ignoring the conditional independence among events.

The older (12- to 13-month-old) infants, in contrast, performed quite differently. They followed and generalized the gaze of the accurate face, just like the 8-month-olds, but they also followed and generalized the gaze of the inaccurate informant. Because these children have so much more experience with faces and have been accumulating evidence of gaze as an informative spatial cue for many months (e.g., Reid, Striano, Kaufman, & Johnson, 2004; Scaife & Bruner, 1975; Senju, Csibra, & Johnson, 2008), it is quite possible that their prior knowledge of faces as reliable spatial cues was weighted more heavily than the (relatively small) amount of information indicating that faces can be unreliable social cues that were available in the experiment. This idea is consistent with the "robust bias to trust" interpretation (Jaswal et al., 2010) suggested above but reinterprets these findings not as indicating a bias toward informants' credulity but as stronger evidence of a sensitivity to communicative intent (similar to the argument made by Csibra & Gergely, 2009, regarding natural pedagogy). Indeed, many of Jaswal's studies suggesting that 3-year-olds have this robust bias involve children coming to recognize that their interlocutors can be intentionally deceitful, which may have a prolonged developmental trajectory (e.g., Polak & Harris, 1999; Talwar & Lee, 2002; see Lee, 2013, for a review).

The above example suggests that infants have the capacity to integrate their experiences of gaze following with their tracking of informants' spatial cueing accuracy. We propose that a similar mechanism underlies the way in which children engage in word learning. Moreover, we wish to suggest that the statistical learning mechanism is not simply replaced by more rational one that is more model-based. Rather, we suggest that the more model-free system still influences certain aspects of children's judgments about others' epistemic trust. In this way, arguments like the one made by Lucas and Lewis (2010) and Heyes (2015) – that is, using associative mechanisms to make these social inferences – are not rejected but rather just represent an incomplete story regarding the mechanisms by which children learn from others.

Questions Can Answer Questions About Mechanisms of Selective Word Learning

We have suggested that children's early statistical learning capacities give rise to their ability to track the accuracy of information generated by others. With experience, children integrate those statistical regularities with other pieces of prior knowledge to make judgments about the reliability of others, and they base inferences about others' information on those reliability judgments. The development observed in our face studies above reflects this integration. Five-month-olds only appreciate the regularity among events, but by 8 months, infants are integrating their representations of that regularity with more general prior knowledge that they possess about the informativeness of faces. This allows them to discriminate between accurate and inaccurate faces. By 13 months, because the prior probability of informants' gaze cueing accuracy is so high, the familiarization events do not provide sufficient data to treat the inaccurate informant as an unreliable source of future knowledge.

Critically, we wanted to present children with a selective word learning environment that manipulated the statistical regularity among the speaker, label, and object but did not present any information about whether the informant was accurate or inaccurate. That way, we could determine whether children were only using this statistical regularity or were integrating that information with other pieces of prior knowledge to make reliability judgments.

Luchkina et al. (2018), Experiment 2) replicated their procedure but changed a critical aspect of the familiarization trials. Instead of speakers bringing out objects and making a statement about the label of the object (i.e., bringing out a star and saying "This is a star."), speakers brought out the objects and asked a question about the identity of the object (i.e., bringing out a star and saying, "Is this a star?"). By using questions, we could present children with speakers who label information that preserved or did not preserve the speaker-label-referent association while not indicating specific differences in their epistemic competence. Critically, 15-month-olds understand subject questions (Seidl, Hollich, & Jusczyk, 2003), and infants as young as 7 months are sensitive to prosodic cues that indicate the difference between questions and statements (e.g., Soderstrom, Ko, & Nevzorova, 2011). Liszkowski, Carpenter, and Tomasello (2008) showed that 12-month-olds understand the pragmatics of questions based on the speaker's access to information and can respond to questions appropriately. These findings point to the possibility that at 18 months (the age investigated by Luchkina et al.), toddlers will differentiate between statements containing accurate label information and questions containing the same information. Both contain the same statistical regularity between the accurate or inaccurate label, referent, and speaker, which would be sufficient for a mechanism that only calculates this information. Children's understanding of questions at this age suggests that they will not treat the questions as indicating differences in the reliability of the speakers if they are using a more sophisticated mechanism than one that only registers this statistical regularity.

Figure 9.1b shows 18-month-olds' performance on the IPLP measure when asked to identify the label uttered by either the accurate[4] or inaccurate speaker. Just like the previous experiment with statements, 18-month-olds showed no difference in looking time between the accurate and inaccurate speakers when a novel speaker asked them to identify familiar labels for familiar objects. This serves as an important control because there is no reason for children to differentiate between the conditions on these trials (i.e., they already know the meaning of these words). Critically, unlike the previous finding, 18-month-olds treated both the accurate and inaccurate question-asker as unreliable sources. They showed no difference in their looking time when asked to identify novel labels spoken by a novel speaker, and they responded at chance levels to both speakers, indicating that they did not learn to associate the novel label with these novel objects based on this exposure. These data are quite similar to performance of the 8-month-olds in Tummeltshammer et al. (2014) arrow condition; the familiarization data was either not strong enough or salient enough to produce a difference among the conditions.

These data suggest that infants are not just using the associative information between a speaker's labels and the referents of those labels as the basis for generalizing novel verbal information. There are two open questions from these findings. The first is why 18-month-olds treat both the accurate and inaccurate speaker as unreliable sources of new knowledge. Investigating how older children respond to a similar familiarization phase allows us to consider the extent to which children can integrate existing knowledge into these judgments. The second question is about the specific mechanism children are using, and again, investigations with older children might shed light on this question.

To answer these questions, we extended the Luchkina et al. procedure to consider how 3- and 4-year-olds treat informants who ask questions about familiar objects using accurate or inaccurate labels (Luchkina, Morgan, & Sobel, in press, Experiment 1). Children were exposed to the same familiarization phase in which the speakers asked questions about the labels for familiar objects, one using the appropriate label and the other using an inappropriate one. Then, between subjects, one of those speakers generated novel labels for novel objects (i.e., one of the two speakers labeled a novel object a "lif" and another novel object a "neem"). At test, the experimenter presented children with two objects: one was a novel object they had just seen labeled in the video, and the second was a novel object that had not been shown in the video. Children were asked to select an object corresponding with the novel label from the video (i.e., the choice would be between the object labeled a "lif" and a novel object not seen previously, and children would be asked, "Which one is a lif?").

If children's judgments about speaker reliability are based only on the speaker-label-referent associations, then children should show evidence for a difference

[4] Note that accuracy here refers to whether the speaker asked a question with the appropriate label for the object or an inappropriate label. That is, the accurate speaker is not more reliable than the inaccurate speaker on a rational account. But if one is only registering statistical regularities, then the accurate speaker should be treated as more reliable than the inaccurate one.

between these conditions. When speakers uses the accurate label in their questions, children should treat the informant as reliable; when speakers use inaccurate labels in their questions, children should treat the informant as unreliable. Of course, this was not what Luchkina et al. (2018) found in an 18-month-old sample – 18-month-olds did not learn from either speaker's questions, as opposed to when those speakers made statements. That pattern of response might have indicated that 18-month-olds did not understand the procedure. Alternatively, it could have indicated that children thought both speakers were epistemically incompetent, after all, why should one ask questions about familiar objects if one knows the meaning of their labels?

We found that both 3- and 4-year-olds remembered the labels generated by the accurate and inaccurate speaker equally across conditions and at levels that were well above chance (Fig. 9.4a). Unlike the 18-month-olds in Luchkina et al. (2018), 3- and 4-year-olds treated both question-askers as reliable sources of knowledge. Where preschoolers differed was in their response time (shown in Fig. 9.4b). They were slower to give correct responses when the label was spoken by the inaccurate speaker than by the accurate speaker. In a follow-up experiment using a similar method (Luchkina et al. in press, Experiment 2), similar findings emerged when children were asked to make a disambiguation inference as opposed to simply recognize the labels they heard the speaker generate. In this case, during test trials (again, in which children observed a novel object labeled by the speaker and an unfamiliar unlabeled novel object), children were asked to select an object corresponding to an unfamiliar novel label. In this disambiguation paradigm, children showed no difference in their inferences – they chose the completely novel object when that label was generated by both the accurate and inaccurate speakers. Their response times did not differ between conditions and were significantly longer than in the first experiment (see Fig. 9.4b). This difference makes sense given the increased demands of the disambiguation task. Rather than responding to the association between the label and the speaker, children must recognize that the speaker is referring to the novel object not previously shown in the video because she generated a different novel label. While the associative mechanism might be involved in making links between the familiar object, the familiar label, and the accuracy of the speaker, in this experiment, one must also infer that the speaker intends to refer to the other object, which is a more sophisticated inference.

Together, these experiments suggest that reliability judgments in preschoolers are driven by a rational process that originates with recognizing associations among speakers, labels, and objects. Children generalize these associations to new information introduced by the same speaker (e.g., Brosseau-Liard & Birch, 2010; Heyes, 2015) and then evaluate that associative information based on their own knowledge. Much like children's causal reasoning abilities potentially begin with their ability to associate information together (Sobel & Kirkham, 2012), children's selective word learning might begin with such associative mechanisms but begin to integrate existing knowledge into interpreting those associations (Sobel & Kushnir, 2013).

In these studies, it is likely that children begin to integrate their experiences with questions, and their pragmatic expectations for social interactions, as well as their

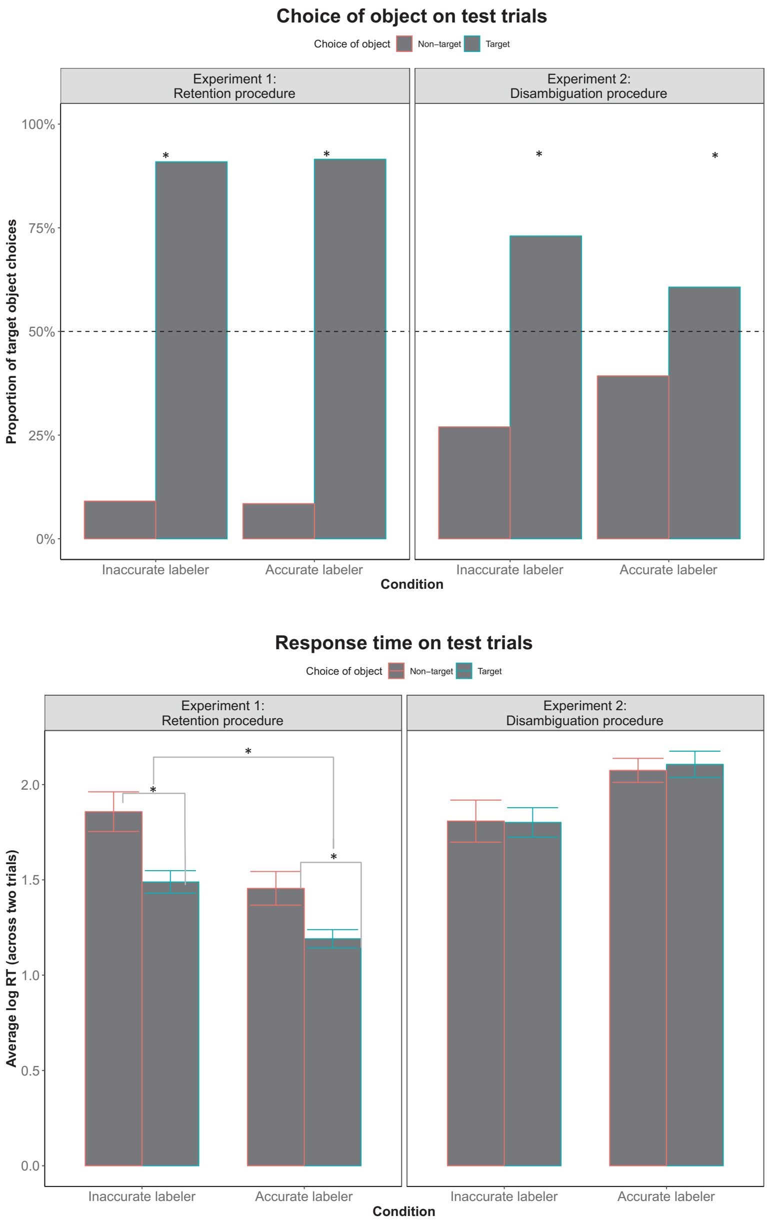


Fig. 9.4 (**a**, **b**) Mean number of preschoolers' correct responses to the experimenter's request in Luchkina et al. (2018), top figure; asterisks indicate above-chance proportions of target object choices) and preschoolers' mean response time to the experimenter's request (bottom figure; asterisks indicate significant differences in response times)

specific appreciation of how yes/no questions communicate truth value information (Choi, 1991). Such a mechanism also explains various findings during the preschool years in which children integrate social or metacognitive knowledge into their judgments of others' reliability. Among other pieces of knowledge, preschoolers evaluate others' age (Jaswal & Neely, 2006), intentions (Heyman, Sritanyaratana, & Vanderbilt, 2013), social group membership (e.g., Kinzler, Corriveau, & Harris, 2011), and familiarity (e.g., Corriveau & Harris, 2009) when evaluating others' reliability for word learning.

The main point that we wish to make here is that despite the prevailing effect of children's knowledge on their judgments of reliability, there might be lingering effects of associative mechanisms on children's selective word learning. Many situational factors that influence selective learning are rational in that they contribute to inferences about others' epistemic knowledge. For example, from the child's perspective, adults have more knowledge than children and indeed might be more generally reliable as sources of novel information (Jaswal & Neely, 2006). Critically, children also recognize that peers might have more knowledge in certain domains (not the meaning of novel words but the names of novel Pokémon characters) than adults (VanderBorght & Jaswal, 2009). Similarly, membership in a linguistic community potentially indicates linguistic as well as conceptual or conventional knowledge related to that community. This would license inferences based on speakers' accent (Kinzler et al., 2011) or appropriate use of syntactic structures (Corriveau, Kinzler, & Harris, 2013; Sobel & Macris, 2013).

Concluding Thoughts: Beyond Selective Word Learning

We have suggested that children possess a rational mechanism for selective learning, which integrates a statistical learning mechanism that tracks the accuracy of an individual with other pieces of prior knowledge about informants to make inferences about their reliability. We have also suggested that the mechanism through which children track others' accuracy is still present and potentially used by young children. Others (e.g., Hermes, Behne, Bich, Thielert, & Rakoczy, 2018; Hermes, Behne, & Rakoczy, 2018) have suggested that children's social learning is best explained by a dual-process model; they showed that children make inferences via a rational mechanism similar to the one we have proposed but are more likely to be influenced by simpler, more associative mechanisms when under cognitive load. We are agnostic as to whether we agree with the idea that the two systems operate independently, but we do agree with the possibility that both associative and rational mechanisms underlie selective word learning.

But on this view, there is nothing special about word learning as a domain to investigate selective learning. Indeed, we have tried to use examples outside of word learning to motivate our model. We posit that integrating statistical learning with real-world knowledge not only allows children to engage in selective word learning but selective social learning more generally. It may be that selective learning underlies not only certain aspects of linguistic competency but cultural and conventional competency more generally (see, e.g., Kline, 2015). Whether the rational mechanism we have suggested here underlies all such learning is an open empirical question but is one that we believe offers a great deal of promise.

References

Bergstrom, B., Moehlmann, B., & Boyer, P. (2006). Extending the testimony problem: Evaluating the truth, scope, and source of cultural information. *Child Development, 77*, 531–538.

Brooker, I., & Poulin-Dubois, D. (2013). Is a bird an apple? The effect of speaker labeling accuracy on infants' word learning, imitation, and helping behaviors. *Infancy, 18*(s1), E46–E68.

Brosseau-Liard, P. E., & Birch, S. A. J. (2010). "I bet you know more and are nicer too!": What children infer from others' accuracy. *Developmental Science, 13*(5), 772–778.

Choi, S. (1991). Children's answers to yes-no questions: A developmental study in English, French, and Korean. *Developmental Psychology, 27*(3), 407.

Clément, F., Koenig, M., & Harris, P. (2004). The ontogenesis of trust. *Mind and Language, 19*(4), 360–379.

Corriveau, K., & Harris, P. L. (2009). Preschoolers continue to trust a more accurate informant 1 week after exposure to accuracy information. *Developmental Science, 12*(1), 188–193.

Corriveau, K. H., Kinzler, K. D., & Harris, P. L. (2013). Accuracy trumps accent in children's endorsement of object labels. *Developmental Psychology, 49*(3), 470–479.

Csibra, G., & Gergely, G. (2009). Natural pedagogy. *Trends in Cognitive Sciences, 13*(4), 148–153.

Denison, S., & Xu, F. (2010). Twelve-to 14-month-old infants can predict single-event probability with large set sizes. *Developmental Science, 13*(5), 798–803.

Haith, M. M. (1993). Future-oriented processes in infancy: The case of visual expectations. In C. E. Granrud (Ed.), *Visual perception and cognition in infancy* (pp. 235–264). Hillsdale, NJ: Erlbaum.

Harris, P. L. (2012). *Trusting what you're told: How children learn from others*. Cambridge, MA: Harvard University Press.

Harris, P. L., & Koenig, M. A. (2006). Trust in testimony: How children learn about science and religion. *Child Development, 77*(3), 505–524.

Hermes, J., Behne, T., Bich, A. E., Thielert, C., & Rakoczy, H. (2018). Children's selective trust decisions: Rational competence and limiting performance factors. *Developmental Science, 21*, e12527.

Hermes, J., Behne, T., & Rakoczy, H. (2018). The development of selective trust: Prospects for a dual-process account. *Child Development Perspectives, 12*, 134–138.

Heyes, C. (2015). Homo imitans? Seven reasons why imitation couldn't possibly be associative. *Philosophical Transactions of the Royal Society B: Biological Sciences, 371*(1686), 20150069.

Heyman, G. D., Sritanyaratana, L., & Vanderbilt, K. E. (2013). Young Children's Trust in Overtly Misleading Advice. *Cognitive Science, 37*(4), 646–667.

Jaswal, V. K. (2010). Believing what you're told: Young children's trust in unexpected testimony about the physical world. *Cognitive Psychology, 61*(3), 248–272.

Jaswal, V. K., Croft, A. C., Setia, A. R., & Cole, C. A. (2010). Young children have a specific, highly robust bias to trust testimony. *Psychological Science, 21*(10), 1541–1547.

Jaswal, V. K., & Neely, L. A. (2006). Adults don't always know best preschoolers use past reliability over age when learning new words. *Psychological Science, 17*(9), 757–758.

Kidd, C., Piantadosi, S. T., & Aslin, R. N. (2014). The goldilocks effect in infant auditory attention. *Child Development, 85*(5), 1795–1804.

Kinzler, K. D., Corriveau, K. H., & Harris, P. L. (2011). Children's selective trust in native-accented speakers. *Developmental Science, 14*(1), 106–111.

Kirkham, N. Z., Slemmer, J. A., & Johnson, S. P. (2002). Visual statistical learning in infancy: Evidence for a domain general learning mechanism. *Cognition, 83*(2), B35–B42.

Kline, M. A. (2015). How to learn about teaching: An evolutionary framework for the study of teaching behavior in humans and other animals. *Behavioral and Brain Sciences, 38*, e31.

Koenig, M. A., Clément, F., & Harris, P. L. (2004). Trust in testimony: Children's use of true and false statements. *Psychological Science, 15*(10), 694–698.

Koenig, M. A., & Echols, C. H. (2003). Neuropsychological developmental change in a case with Noonan syndrome: Longitudinal assessment. *Brain & Development, 25*(4), 291–293.

Koenig, M. A., & Harris, P. L. (2005a). Preschoolers mistrust ignorant and inaccurate speakers. *Child Development, 76*(6), 1261–1277.

Koenig, M. A., & Harris, P. L. (2005b). The role of social cognition in early trust. *Trends in Cognitive Sciences, 9*, 457–459.

Koenig, M. A., & Woodward, A. L. (2010). Sensitivity of 24-month-olds to the prior inaccuracy of the source: Possible mechanisms. *Developmental Psychology, 46*(4), 815.

Krogh-Jespersen, S., & Echols, C. H. (2012). The influence of speaker reliability on first versus second label learning. *Child Development, 83*(2), 581–590.

Kushnir, T., Vredenburgh, C., & Schneider, L. A. (2013). "Who can help me fix this toy?" The distinction between causal knowledge and word knowledge guides preschoolers' selective requests for information. *Developmental Psychology, 49*(3), 446–453.

Lee, K. (2013). Little liars: Development of verbal deception in children. *Child Development Perspectives, 7*(2), 91–96.

Liszkowski, U., Carpenter, M., & Tomasello, M. (2008). Twelve-month-olds communicate helpfully and appropriately for knowledgeable and ignorant partners. *Cognition, 108*(3), 732–739.

Lucas, A. J., & Lewis, C. (2010). Should we trust experiments on trust? *Human Development, 53*(4), 167–172.

Luchkina, E., Morgan, J, & Sobel, D. M. (in press). Questions can answer questions about the mechanisms of preschoolers' selective word learning. *Child Development*.

Luchkina, E., Sobel, D. M., & Morgan, J. (2018). Eighteen-month-olds selective generalize words from accurate speakers to novel contexts. *Developmental Science, 21,* e12663.

Madole, K. L., & Cohen, L. B. (1995). The role of object parts in infants' attention to form-function correlations. *Developmental Psychology, 31*(4), 637–648.

Mascaro, O., & Sperber, D. (2009). The moral, epistemic, and mindreading components of children's vigilance towards deception. *Cognition, 112*(3), 367–380.

Mills, C. M. (2013). Knowing when to doubt: Developing a critical stance when learning from others. *Developmental Psychology, 49*(3), 404.

Nurmsoo, E., & Robinson, E. J. (2009). Identifying unreliable informants: Do children excuse past inaccuracy? *Developmental Science, 12*(1), 41–47.

Pasquini, E. S., Corriveau, K. H., Koenig, M., & Harris, P. L. (2007). Preschoolers monitor the relative accuracy of informants. *Developmental Psychology, 43*(5), 1216–1226.

Polak, A., & Harris, P. L. (1999). Deception by young children following noncompliance. *Developmental Psychology, 35*(2), 561–568.

Poulin-Dubois, D., & Chow, V. (2009). The effect of a Looker's past reliability on infants' reasoning about beliefs. *Developmental Psychology, 45*(6), 1576–1582.

Reid, T. (1785). An inquiry into the human mind (1764). *Philosophical Works, 1*.

Reid, V. M., Striano, T., Kaufman, J., & Johnson, M. H. (2004). Eye gaze cueing facilitates neural processing of objects in 4-month-old infants. *Neuroreport, 15*(16), 2553–2555.

Saffran, J., Aslin, R., & Newport, E. (1996). Statistical learning by eight-month-old infants. *Science, 274*(5294), 926–928.

Scaife, M., & Bruner, J. S. (1975). The capacity for joint visual attention in the infant. *Nature, 253*(5489), 265–266.

Seidl, A., Hollich, G., & Jusczyk, P. W. (2003). Early understanding of subject and object Wh-questions. *Infancy, 4*(3), 423–436.

Senju, A., Csibra, G., & Johnson, M. H. (2008). Understanding the referential nature of looking: Infants' preference for object-directed gaze. *Cognition, 108*(2), 303–319.

Sobel, D. M., & Corriveau, K. H. (2010). Children monitor individuals' expertise for word learning. *Child Development, 81*(2), 669–679.

Sobel, D. M., & Kirkham, N. Z. (2006). Blickets and babies: The development of causal reasoning in toddlers and infants. *Developmental Psychology, 42*(6), 1103–1115.

Sobel, D. M., & Kirkham, N. Z. (2007). Bayes nets and babies: Infants' developing statistical reasoning abilities and their representation of causal knowledge. *Developmental Science, 10*(3), 298–306.

Sobel, D. M., & Kirkham, N. Z. (2012). The influence of social information on children's statistical and causal inferences. *Advances in Child Development and Behavior, 43*, 321–350.

Sobel, D. M., & Kushnir, T. (2013). Knowledge matters: How children evaluate the reliability of testimony as a process of rational inference. *Psychological Review, 120*(4), 779–797.

Sobel, D. M., & Macris, D. M. (2013). Children's understanding of speaker reliability between lexical and syntactic knowledge. *Developmental Psychology, 49*(3), 523–532.

Soderstrom, M., Ko, E. S., & Nevzorova, U. (2011). It's a question? Infants attend differently to yes/no questions and declaratives. *Infant Behavior and Development, 34*(1), 107–110.

Stenberg, G. (2013). Do 12-month-old infants trust a competent adult?. *Infancy, 18,* 873–904.

Talwar, V., & Lee, K. (2002). Development of lying to conceal a transgression: Children's control of expressive behaviour during verbal deception. *International Journal of Behavioral Development, 26*(5), 436–444.

Tummeltshammer, K. S., & Kirkham, N. Z. (2013). Learning to look: Probabilistic variation and noise guide infants' eye movements. *Developmental Science, 16*(5), 760–771.

Tummeltshammer, K. S., Sobel, D. M., & Kirkham, N. Z. (2019). *Infants' developing capacity for tracking accuracy in gaze following.* Manuscript in preparation, Brown University.

Tummeltshammer, K. S., Wu, R., Sobel, D. M., & Kirkham, N. Z. (2014). Infants track the reliability of potential informants. *Psychological Science, 25*(9), 1730–1738.

Vanderbilt, K. E., Heyman, G. D., & Liu, D. (2014). In the absence of conflicting testimony young children trust inaccurate informants. *Developmental Science, 17*(3), 443–451.

VanderBorght, M., & Jaswal, V. K. (2009). Who knows best? Preschoolers sometimes prefer child informants over adult informants. *Infant and Child Development, 18*(1), 61–71.

Zmyj, N., Buttelmann, D., Carpenter, M., & Daum, M. M. (2010). The reliability of a model influences 14-month-olds' imitation. *Journal of Experimental Child Psychology, 106*(4), 208–220.

Chapter 10
The Search for Invariance: Repeated Positive Testing Serves the Goals of Causal Learning

Elizabeth Lapidow and Caren M. Walker

This research was supported by funding from the National Defense Science and Engineering Graduate Fellowship awarded to EL. We would like to thank Gail Heyman and Craig McKenzie for their insights, discussion, and feedback on these ideas. *What is invariant does not emerge unequivocally except with a flux. The essentials become evident in the context of changing nonessentials.*

James Gibson, 1979

Abstract Positive testing is characteristic of exploratory behavior, yet it seems to be at odds with the aim of information seeking. After all, repeated demonstrations of one's current hypothesis often produce the same evidence and fail to distinguish it from potential alternatives. Research on the development of scientific reasoning and adult rule learning have both documented and attempted to explain this behavior. The current chapter reviews this prior work and introduces a novel theoretical account—the Search for Invariance (SI) hypothesis—which suggests that producing multiple positive examples serves the goals of causal learning. This hypothesis draws on the interventionist framework of causal reasoning, which suggests that causal learners are concerned with the *invariance* of candidate hypotheses. In a probabilistic and interdependent causal world, our primary goal is to determine whether, and in what contexts, our causal hypotheses provide accurate foundations for inference and intervention—not to disconfirm their alternatives. By recognizing the central role of invariance in causal learning, the phenomenon of positive testing may be reinterpreted as a rational information-seeking strategy.

Human learners are intuitively exploratory: We acquire new knowledge from the outcomes of our actions. However, in order for exploration to support learning, at

E. Lapidow (✉) · C. M. Walker
Department of Psychology, University of California – San Diego, La Jolla, CA, USA
email: elapidow@ucsd.edu

J. B. Childers (ed.), *Language and Concept Acquisition from Infancy Through Childhood*, https://doi.org/10.1007/978-3-030-35594-4_10

least some of these actions must serve to evaluate our *existing* knowledge. Despite this need for informative "hypothesis testing" in everyday learning, decades of research examining self-directed experimentation suggests that learners rarely choose informative tests. That is, instead of selecting actions to test *whether* their current hypothesis is correct, both children and adults tend to prefer "positive tests": actions that will produce an effect assuming their current hypothesis *is* correct (see Klayman, 1995; Zimmerman, 2007).

To illustrate, suppose you drop an ice cube on the floor and it shatters. As a learner, you might form an initial hypothesis that "impact with an unyielding surface causes ice to shatter." This hypothesis is also a causal explanation for your observation: indicating how one variable (*X*) *makes a difference* to the state of another variable (*Y*). According to traditional interpretations of Popper's (1959) falsificationist approach, testing this hypothesis would require disconfirming its alternatives. That is, assessing whether "*X* is the cause of *Y*" requires "negative tests" or actions to determine whether *Y* occurs in the absence of *X*. Here, since *Y* is "shattering" and *X* is "impacting an unyielding surface," you should drop an ice cube on a *yielding* surface (*not X*), like rubber or cotton, to determine whether it will shatter.

However, learners rarely choose this kind of disconfirming action during their exploration. Instead, they are much more likely to *repeat* the initial observation: for example, to pick up another ice cube and drop it on the same surface or a similar one. This tendency to generate multiple positive examples is a puzzling characteristic of self-directed learning since it does not initially appear to be informative. After all, these repeated demonstrations often produce the same evidence and do not distinguish between the current hypothesis (i.e., impacting an unyielding surface) and potential alternatives (e.g., impacting *any* surface at a particular speed) since they are consistent with both. Why then do self-directed learners consistently and repeatedly conduct positive tests?

In this chapter, we propose a novel answer to this question: the Search for Invariance (SI) hypothesis, which suggests that observing multiple positive examples may facilitate learning by allowing us to assess the *invariance* of our causal theories. That is, by repeatedly activating a hypothesized cause and checking if its anticipated effect occurs, positive tests generate information about the degree to which this relationship holds across time and contexts. Given that the majority of the causal relationships we encounter are probabilistic and interdependent, determining the degree of invariance is critical for utilizing causal knowledge as a basis for action and inference. In order to test whether and when *X* (e.g., impacting an unyielding surface) reliably brings about *Y* (e.g., shattering in ice), it is necessary to *repeat X* (e.g., dropping more ice on similar surfaces) and observe whether *Y* occurs *again*.

The current chapter will unpack this claim that the tendency to engage in positive testing is motivated by (and serves) our goals as causal learners. First, we will outline the empirical evidence for the use of a "positive testing strategy," during exploratory learning, and review existing theoretical accounts that have been proposed to explain it. We will then introduce the Search for Invariance (SI) hypothesis and explain its foundations within theories of causality. After establishing this background, we will aim to apply our novel account to reinterpret some of the primary examples of positive testing in exploratory learning and address potential objections and misinterpretations (e.g., sufficiency).

Positive Testing Strategy

A variety of learning and reasoning behaviors have been linked to (and confounded with) the positive testing strategy (PTS), so it is important to establish a working definition of this term. For the purposes of the current discussion, PTS will be treated as a phenomenon of hypothesis *testing* and defined as the tendency (or preference) to select actions with the highest probability of producing the expected effect, if the current hypothesis were correct.[1] That is, we will focus on cases in which the learner assesses a hypothesis by examining its positive instances—either checking whether the expected effect occurs when the hypothesized conditions are met or checking whether the conditions of the hypothesis are met when the event occurs (Klayman & Ha, 1987).

We will, therefore, *not* address accounts that focus primarily on whether young learners are able to generate hypotheses and evaluate their fit to evidence more broadly (e.g., Carey, Evans, Honda, Jay, & Unger, 1989; Kuhn, 1989; Kuhn et al., 1988). This discussion is also not intended to address "*confirmation bias*," the failure to seek and consider (or even to avoid and distort) conflicting evidence, which is often presented alongside PTS in adult research[2] (for reviews, see Klayman, 1995; Nickerson, 1998). While the ability (and willingness) to reconcile an existing theory with new evidence is critical for exploration to support learning, it falls outside our specific focus on the generation of evidence through self-directed action.

PTS in Scientific Reasoning

Inhelder and Piaget (1958) were the first to experimentally examine the understanding and use of the principles of experimentation in children. Later researchers adapted their methodologies to assess and improve scientific reasoning (e.g., Kuhn & Angelev, 1976; Kuhn & Brannock, 1977; Siegler & Liebert, 1975), and a tremendous body of research has grown out of this initial work (see Zimmerman, 2000, 2007; Zimmerman

[1] As discussed below, there are many accounts of this behavior, not all of which use the term "PTS." Additionally, while the term has also been used to describe learners' motivation for conducting positive tests, we will restrict our use of "PTS" to refer to observable behavior.

[2] The exact nature of the relationship between PTS and confirmation bias differs between accounts. PTS is variously suggested to be (a) an instance of (Nickerson, 1998; Wason, 1962), (b) a source of (see Nickerson, 1998 for review), and (c) a departure from confirmation bias (Klayman, 1995; Klayman & Ha, 1987).

& Klahr, 2018 for reviews). Studies typically present children with multivariate contexts and assess their ability to systematically combine and isolate these variables to reveal causal relationships. In some cases, participants are instructed to determine the variable(s) causally related to an outcome (e.g., which chemicals cause a color reaction when mixed; Kuhn & Phelps, 1982). In others, children are asked to determine whether and how variable(s) make a difference to a certain outcome (e.g., which features of a race car determine its speed; Schauble, 1990). Another common approach is to indicate a variable of interest and ask participants to test hypotheses about its effect (e.g., the operation performed by a computer input command; Dunbar & Klahr, 1989; Klahr & Dunbar, 1988; Klahr, Fay, & Dunbar, 1993).

The bulk of this research finds the development of experimentation skills to be slow and error-filled (e.g., Dunbar & Klahr, 1989; Inhelder & Piaget, 1958; Klahr et al., 1993; Klahr & Chen, 2003; Kuhn, 1989; Tschirgi, 1980; Valanides, Papageorgiou, & Angeli, 2014). Importantly, many of these errors resemble PTS: Children tend to repeatedly choose actions and experiments that are expected to generate an effect (e.g., the color reaction, the fastest car) if their causal hypothesis were correct. This behavior is often interpreted as driven by children's desire to "demonstrate the correctness" of their hypotheses (e.g., Dunbar & Klahr, 1989; Inhelder & Piaget, 1958; Klahr et al., 1993; Kuhn & Phelps, 1982). Other researchers have viewed these explorations as evidence of a misplaced focus on an action's *tangible* outcomes rather than its informative potential (e.g., Kuhn & Phelps, 1982; Schauble, 1990; Schauble, Glaser, Duschl, Schulze, & John, 1995; Siler & Klahr, 2012; Siler, Klahr, & Price, 2013; Tschirgi, 1980; Zimmerman & Glaser, 2001). That is, rather than trying to learn the relations between cause and effect, children seem to select experiments in order to reproduce positive outcomes and avoid negative ones.

Tschirgi (1980) is perhaps the most cited example of PTS in scientific reasoning (see Croker & Buchanan, 2011; Klayman & Ha, 1987 for discussions). In this study, 2nd-, 4th-, and 6th-grade children and adults were asked to choose an experiment to prove that a variable was causally responsible for either a positive or negative outcome. In one scenario, a character bakes a cake using one of two types of each of three ingredients (a flour, a sweetener, and a fat), and the cake comes out well. The character believes that the type of *sweetener* causes this outcome while the types of *flour* and *fat* do not matter. Participants were then given a choice between three potential experiments and asked to select one to prove the character's hypothesis. They could (1) change the suspected cause and keep the other two variables constant ("VARY"), (2) change the other two variables and keep the suspected cause constant ("HOLD"), or (3) change all three variables ("CHANGE ALL"). According to Tschirgi, the VARY option, which isolates the suspected causal variable, is the only informative test of the character's hypothesis. However, participants of all ages only preferred this option when the outcome of the initial scenario was *negative* (i.e., when the cake came out badly). Otherwise, learners preferred the HOLD option—keeping the variable of interest constant and changing the others. Tschirgi (1980) explains this finding as evidence that children and adults tend to select experiments based on practical rather than "logical" concerns.

PTS in Rule Learning

There is also extensive evidence of PTS within adult's hypothesis testing (see McKenzie, 2004 for review). The classic example of this behavior is the 2-4-6 task (Wason, 1960),[3] which asks participants to determine the rule used to generate sequences of three numbers. At the start of the task, the experimenter provides an example of a sequence generated by the rule (e.g., "2, 4, 6"). Participants are then able to experiment by generating novel sequences and requesting feedback from the experimenter about whether these follow the rule. Notably, the predominant strategy is to test cases that *fit* the rule one has in mind: to test positive instances of one's current hypothesis. For example, most adults, given the "2, 4, 6" example, form the hypothesis that the rule is "ascending even numbers." In gathering evidence, the majority of participants will test sequences that follow this rule (e.g., "10, 12, 14," "−2, 0, 2," "104, 106, 108," etc.) and treat affirmative feedback from the experimenter as evidence serving to increase the strength of their belief in the accuracy of their hypothesis.

The problem with using a positive testing strategy in this task is that it misses the correct (and more general) rule, "ascending numbers," and provides no evidence to suggest that an error has been made. Wason (1960) interpreted this as evidence of participants arriving at their hypotheses through a process of "enumeration" rather than "elimination." In other words, learners assume that confirming evidence alone is enough to justify their conclusions. While some participants do eventually discover the correct rule, most of them only do so after *many* positive tests of incorrect rules. Wason's task and his conclusion that learners have a general tendency to only consider verification (Wason, 1960; Wason & Johnson-Laird, 1972) have both been used extensively as evidence of adults' biased hypothesis testing and reasoning (e.g., Gorman & Gorman, 1984; Mahoney & DeMonbreun, 1977; Tukey, 1986; Tweney et al., 1980; Wetherick, 1962).

Theories of PTS

Given that PTS is a widespread and well-documented phenomenon, there have been numerous theoretical attempts to account for it. These accounts may be roughly separated into two categories, depending on whether PTS is explained as a means of generating an outcome or as a means of generating evidence.

[3] Wason also created another classic task of hypothesis testing, the Selection Task (1968), which falls outside the scope of the current discussion.

PTS as a Means of Generating Outcomes

As mentioned above, some have suggested that learners prefer PTS because their selection of actions is motivated by tangible outcomes, rather than information value (e.g., Kuhn & Phelps, 1982; Tschirgi, 1980). Studies of early scientific reasoning, for example, often suggest that young learners begin with an incorrect intuition about the *function* of experimentation. Children are described as preoccupied with "making events happen" (e.g., making a good cake, producing a color reaction) rather than identifying the causal factors responsible for these outcomes.

Based on this evidence, Schauble, Klopfer, and Raghavan (1991) proposed the "science versus engineering models" of self-directed behavior: Individuals can adopt either a "science goal" (to determine causal relationships) or an "engineering goal" (to generate or reproduce a particular effect) in interactions with their environment. According to the authors, children incorrectly approach scientific reasoning tasks using an "engineering" model, which is concerned with generating outcomes and stops whenever its target (or an acceptable approximation) is achieved. Schauble (1990) distinguishes this account of PTS from the one offered by Klayman and Ha (1987). In particular, the engineering model does *not* claim that learners are seeking falsification or information of any kind. Instead of trying to determine the causal relationships between variables and outcomes, young explorers manipulate variables in an attempt to bring about desirable outcomes (Schauble, 1990). In other words, according to this theory, learners' interventions are often uninformative because *information is not their goal*. The strongest version of this account implies that children do not differentiate between understanding an event and making it occur. Although later empirical work provides evidence against this claim (Sodian, Zaitchik, & Carey, 1991; see also Lapidow & Walker, 2019), the "science versus engineering" explanation of PTS continues to be cited and used as a framework for understanding choice behavior within scientific reasoning (e.g., Siler et al., 2013; Siler & Klahr, 2012).

Other researchers have made suggestions similar to the "science versus engineering" hypothesis but concerning adult choice behavior (e.g., Friedrich, 1993; Schwartz, 1982; Vogel & Annau, 1973). These interpretations suggest we employ PTS because we value maximizing current over long-term gains. One early articulation of this idea comes from Einhorn and Hogarth (1978), who describe a conflict between acting in accordance with the current hypothesis (i.e., maximizing current success) and acquiring information to improve it (i.e., increasing long-term success). As a result, in contexts where the potential cost of false positives is more damaging than that of false negatives, learners may be entirely justified in gathering positive evidence.

These "outcome-focused" theories of PTS all point to the tension between "exploitation," taking an action that is known to have a high likelihood of success, and "exploration," taking a less certain (or less rewarding) action in order to improve one's epistemic status. Schauble and colleagues (1990, 1991) suggest that this tension results from a lack of understanding on the part of the learner, while Einhorn and Hogarth (1978) and others point to situational factors that lead the learner to prioritize exploitation. Regardless of the source, this is a real tension in behavior and has been shown to influence decision-making on a variety of tasks. For example, Heyman and Dweck (1992) report that an individual's response to challenges and failures may be explained by whether they hold a "performance goal" (to prove one's competence) or a "learning goal" (to improve one's skills).[4] A recent study has also shown that sensitivity to the distinction between production and investigation influences the actions of causal learners as a rational adaption to context demands (Yoon, MacDonald, Asaba, Gweon, & Frank, 2018). However, regardless of whether it is due to a lack of understanding or situational pressures, the tension between productive and informative actions is likely not the *only* factor responsible for PTS.

PTS as a Means of Generating Evidence

The accounts reviewed above argue that positive tests reflect learners' desire to bring about tangible effects. However, other accounts have suggested that PTS is an *intentional* (though not always valid) form of hypothesis testing. For example, Wason's (1960, 1972) interpretation of adults' behavior on the 2-4-6 task suggested that learners believe that their positive tests provided valid conclusive evidence. Wason based this analysis on the prescriptions for hypothesis testing laid out by Popper (1959), who argued that instances that *verify* a hypothesis are always ambiguous since they can occur even if the hypothesis is ultimately incorrect. Learners should therefore aim to collect counterevidence when testing hypotheses since falsifying evidence is always *conclusive* (i.e., observing a single black swan can overturn an entire lifetime of positive evidence that "'all swans are white"). Further, this type of observation cannot be countered by positive evidence later on (e.g., no number of white swans observed *after* the single black one will make the statement that "all swans are white" true). Popper's prescription for scientific hypothesis testing, therefore, is to make the falsification of alternatives—not the generation of positive evidence—the primary goal of experimentation. By comparing participants' behavior to this standard, Wason concluded that PTS is a logical or cognitive failing of our intuitive hypothesis testing, and many others have echoed this interpretation (e.g., Baron, Beattie, & Hershey, 1988; Devine, Hirt, & Gehrke, 1990; Skov & Sherman, 1986; Wason & Johnson-Laird, 1972).

Indeed, hints of this perspective appear in the account of PTS suggesting that children aim to "demonstrate the correctness" of their hypotheses during scientific reasoning tasks (e.g., Dunbar & Klahr, 1989; Inhelder & Piaget, 1958; Klahr et al., 1993; Kuhn & Phelps, 1982). For example, Klahr and colleagues (1993) gave 3rd- and 6th-grade children and adults either a plausible or implausible hypothesis for the operation of one input in a simple programming system. If the starting hypothesis was plausible, participants tended to go about generating positive evidence of its

[4] While this distinction resembles Schauble et al.'s (1991) notion of "science versus engineering models," Heyman and Dweck's (1992) account is agnostic about the immediate goal. You could, therefore, conceivably have either a "learning" or a "performance" goal while following *either* a "science" or an "engineering" model.

validity. When the starting hypothesis was implausible, however, participants were more likely to set up experiments to discriminate between this initial claim and self-proposed alternatives. Tellingly, these participants (and young children in particular) were often "sidetracked" by generating positive evidence for their rival hypothesis (Klahr et al., 1993).

In other accounts, PTS is not assumed to be driven by an error or illusion of validity but is treated instead as one of several potentially useful inquiry strategies available to the learner. The classic form of this argument comes from Klayman and Ha's (1987) analysis of the 2-4-6 task. These authors distinguish between the uses of falsification as a *goal* versus as a *method* of hypothesis testing. The goal is what Popper (1959) prescribes, but the method is not necessary to achieve it. The two are confounded in Wason's task because most participants begin with a hypothesis that is more specific than the correct one. As a result, testing negative instances of one's hypothesis (falsification as method) is the *only* means of generating disconfirming evidence (falsification as goal). Klayman and Ha argue that this circumstance is neither necessary nor typical in real-world learning. More often, correct hypotheses are a "minority phenomenon," so most tests of positive instances will result in negative responses (Klayman & Ha, 1987). Given this "rarity," PTS is argued to be a more efficient means of seeking informative disconfirmation than negative tests and therefore a reasonable hypothesis testing heuristic.

Klayman and Ha's account has since been broadly adopted to explain PTS in adult learning (see Coenen, Nelson, & Gureckis, 2018), and similar arguments have been successfully applied to other instances of PTS beyond the 2-4-6 task (e.g., Oaksford & Chater, 1994). Navarro and Perfors (2011) have also extended this argument by demonstrating that PTS is a "near-optimal" learning strategy in contexts where correct hypotheses are "sparse" and provide further justification for learners' intuitive assumption of sparsity. Importantly, these accounts present PTS as a default heuristic approach to hypothesis testing. That is, in the absence of more specific information, learners employ PTS because it is cognitively inexpensive and often effective (Klayman & Ha, 1987).

In contrast, other accounts have stressed that learners are sensitive to their learning context and *selectively* employ PTS. For example, McKenzie and Mikkelsen (2000) show that manipulating participants' beliefs about the rarity of events changes the degree to which they appeal to PTS. Furthermore, recent computational work finds that intervention choice is best captured by a model employing a *mix* of PTS and expected information gain in a way that is sensitive to task demands. Coenen, Rehder, and Gureckis (2015) modeled PTS in causal learning as a preference to intervene on variables with the greatest proportion of downstream effects, treating each consistent outcome as a point of positive evidence for the hypothesis. Their analysis compared this model, and one designed to favor interventions most likely to distinguish between competing hypotheses (i.e., with the highest expected information gain) to participants' chosen interventions. They found that a model employing a mixture of *both* strategies best captured the intervention choices of adult causal learners. Results also revealed that strategy selection was sensitive to the learning context: When the task feedback indicated that one strategy was insufficient for distinguishing the true causal structure, or when placed under time pressure, participants shifted flexibly between the PTS and information gain models. Thus, the most recent research suggests that PTS is *not* a default heuristic for generating evidence but an adaptive and efficient hypothesis-testing strategy employed by context-sensitive learners.

Other Accounts and Overlapping Evidence

While these two broad categories of accounts—positive tests as a means of generating outcomes and positive tests as a means of generating evidence—help to organize much of the existing literature, some prior accounts do not neatly conform to *either* category. For example, a recent "self-teaching" model of active learning (Yang, Vong, Yu, & Shafto, 2019) proposes that PTS may be a by-product of the way that interventions are chosen. That is, although the ideal learner usually selects actions according to expected information gain, selections that deviate from this accord with PTS. In this context, PTS is not presented as a mistaken focus (e.g., Schauble et al., 1991) or a logical error (e.g., Wason, 1960), but it is also not presented as an adaptive learning tool (e.g., Coenen et al., 2015).

It is also not always possible to distinguish between generating evidence and generating effects within a learner's behavior. In many cases of causal learning, the action expected to maximize the probability of the most likely hypothesis is *also* expected to have the most positive tangible outcomes.[5] For example, McCormack, Bramley, Frosch, Patrick, and Lagnado (2016) presented children with a three-component causal system and three competing hypotheses: a common cause (activating component A causes components B and C to activate) and two causal chains (activating A causes B to activate, which causes C to activate, or activating A causes C to activate, which causes B to activate). Five- to six-year-olds preferred to *repeatedly* activate component A (the root node in all hypotheses). Although activating A is expected to activate all the other components in the system, it is not clear what motivates this action. It accords with "generating effects" theories of PTS, since turning on the root node of a system "makes the most things happen," but also with "generating evidence" theories, since this action also tests the greatest proportion of downstream connections.

A similar study by Meng, Bramley, and Xu (2018) tested 5- to 7-year-olds on a modified version of Coenen et al.'s (2015) task and also found a preference for this type of intervention. Although children's intervention choice was best captured by a model incorporating both PTS and information gain (as with adults), this mix was heavily skewed toward PTS (i.e., intervening on the node with the greatest proportion of dependent causal links), with the vast majority of children using PTS as their primary intervention strategy. Again, we observe evidence that young causal learners preferentially select positive tests, and again, the source of this preference remains unclear.

[5] See Bramley, Lagnado, and Speekenbrink (2015) for an in-depth treatment of this overlap between expected probability gain and expected utility gain models of intervention.

The Current Theory: Positive Testing and Causal Learning

Rather than explaining PTS as an error, bias, or by-product of our inquiry strategies, the current proposal (the Search for Invariance [SI] hypothesis) presents an alternative account that draws on evidence describing our strengths as self-directed *causal* learners. In contrast to the difficulties documented in the scientific and rule learning domains, we excel at exploratory causal learning from an early age.

Considerable evidence shows that children spontaneously and preferentially explore what is most likely to be informative, given their current causal beliefs (see Schulz, 2012 for a review). Research in causal learning typically presents children (3- to 6-year-olds) with evidence about a novel physical device and then allows them to interact with it during a period of free play. When the evidence is ambiguous or violates their current theories, children engage in significantly more exploration (Bonawitz, van Schijndel, Friel, & Schulz, 2012; Gweon & Schulz, 2008; Schulz & Bonawitz, 2007; Schulz, Standing, & Bonawitz, 2008). They are also more likely to take potentially informative actions during this exploration (Lapidow & Walker, 2019; Schulz & Bonawitz, 2007; van Schijndel, Visser, van Bers, & Raijmakers, 2015).

For example, Cook, Goodman, and Schulz (2011) found that 4- to 5.5-year-olds select and even spontaneously invent informative interventions in their exploration of an ambiguous causal system. Children were introduced to a toy that played music when beads were placed on top of it. They were taught either that *all* beads caused the toy to activate or that only *some* did while the rest were inert. The beads could be snapped together to form two-bead pairs, and a snapped pair of novel beads were given to children during their free play. This pair caused the toy to activate, but it was impossible to tell from this observation alone whether *both* beads in the pair had the power to activate the toy or only *one*. Faced with this ambiguity, children in the "some-beads" condition took informative actions: pulling the pair apart and trying the beads on the toy in isolation in order to disambiguate their causal status. In contrast, children in the "all-beads" condition, for whom the pair was *not* perceived as ambiguous, did not produce these actions. Furthermore, when given an ambiguous pair that was permanently glued together, several children in the "some-beads" condition spontaneously turned the pair on its end, conducting informative hypothesis tests to isolate the beads in a way they had never seen demonstrated.

The fact that young children are such voracious and effective self-directed learners in these contexts raises the possibility that causal learners may have different information-seeking goals than those typically assumed in studies of scientific experimentation. This view of self-directed causal learners as "intuitive scientists" (Brewer & Samarapungavan, 1991; Carey, 1985; Gopnik & Meltzoff, 1997; Karmiloff-Smith, 1988) is the impetus for the SI hypothesis: that PTS is an informative means of assessing causal *invariance* across the set of examples tested, and thus, normatively motivated by the concerns of causal learning. In a causal world of predominately probabilistic and interdependent relationships, repeated generation and investigation of positive instances provide critical information about the reliability and consistency of our causal hypotheses. In other words, PTS is useful and informative because it allows learners to examine the degree to which a dependency between variables continues to hold over time and across contexts. In this way, the SI hypothesis suggests that our goals as causal learners might explain our behavior as intuitive scientists.

Causal Invariance and Interventionism

In order to more fully describe the SI hypothesis and establish its relevance for interpreting PTS, we will first situate the concept of invariance within theories of causality.

Numerous theories of causality and causal explanation include a central idea that the function of causal knowledge is to highlight patterns of dependence that will generalize to future contexts (see Hitchcock, 2012 for details). A few notable examples of this include notions of "sensitivity" (Lewis, 1974), "robustness" (Redhead, 1987), "non-contingency" (Kendler, 2005), "exportability" (Lombrozo & Carey, 2006), "insensitivity" (Ylikoski & Kuorikoski, 2010), "portability" (Weslake, 2010), and "transportability" (Pearl & Bareinboim, 2011). Indeed, this sensitivity to regularities in variable input is critical for knowledge acquisition in other domains as well (e.g., see Wu, Gopnik, Richardson, & Kirkham, 2011).

Although these accounts vary, Sloman's (2005) description of "invariance" provides a sense of the common ground among them. He argues that, in every domain, aspects that are *invariant* across instances hold the most valuable information for learners. These aspects represent the "stable, consistent, and reliable properties that hold across time and across different instantiations of a system" (Sloman, 2005, p. 15). Knowledge of invariance therefore allows learners to predict, explain, and manipulate events in the world. The concept of invariance is not defined exclusively in terms of causality; for example, recognizing the invariant statistical regularities among syllables within streams of continuous speech supports early language learning (e.g., Saffran, Aslin, & Newport, 1996; Saffran, Johnson, Aslin, & Newport, 1999). However, the causal relations that govern observable events are usually highly systematic and generalizable. Causal knowledge is, therefore, a key source of invariance, and stable causal principles are often the most reliable basis for inference and interaction available to us (Sloman, 2005).

Much of Sloman's account draws on the interventionist perspective on causal explanation (e.g., Woodward, 1997, 2003, 2006, 2010), which defines a causal relationship in terms of the invariance between variables following some change. That is: If X causes Y, then intervening to change the value of X would result in a change to the value of Y. A causal explanation is thus a true claim describing how some factors *make a difference* to others. To illustrate this, suppose that two variables, X and Y, are observed to co-occur. If an intervention changing the value of X maintains this correlation (i.e., it leads to a corresponding change in the value of Y), then the relationship between X and Y is *invariant* (it continues to hold), under at least some interventions. This relationship is regarded as causal and can be used as a basis for inference and manipulation. To make this idea more concrete, imagine that *X* is fertilizer and *Y* is plant growth. If the statement "Fertilizer causes plant growth" is an accurate causal explanation, we would expect changes to the amount of fertilizer (*X*) to lead to *systematic* changes in the growth of the plant (*Y*) (Woodward, 1997).

Note that the interventionist concept of invariance outlined so far defines a causal explanation as representing a kind of *counterfactual dependence*: It indicates how changing the factors included in the explanation would lead to a difference in the phenomenon being explained (Woodward, 1997, 2003). For example, the hypothesis that "impacting an unyielding surface causes ice to shatter" provides a causal explanation for shattering and implies the counterfactual that, in the absence of such an impact, shattering would not occur. Causal knowledge allows us not only to reason about how one factor makes a difference to another but also to exploit those dependencies in our actions. Interventionism views causal learning and reasoning as rooted in our "highly practical interest" in manipulation and control of our environment (Woodward, 2003, p. 10). As a result, the value of identifying causal relationships is inherently tied to the knowledge of the actions this information supports.

In addition to its role in defining causal relationships, invariance is also an important quality expressed by causal relationships. While the former captures the continuity of causal relationships when related variables change, the latter emphasizes the "stability" of those relationships across contexts and conditions. That is, *X* is a cause of *Y* if and only if an intervention on *X* changes the value of *Y* in at least some background circumstances *b*. These "background circumstances" are aspects of a situation that are not explicitly represented by *X* or *Y* but that are critical to the meaning and commitments of causal explanations (Blanchard, Vasilyeva, & Lombrozo, 2018). For example, "Fertilizer causes plant growth" is only true if the plant is also getting water, sunlight, and oxygen. These are some of the *background circumstances* of the causal relationship between fertilizer and plant growth; these are both critically important to our understanding of the causal claim and not explicitly stated as part of the relation.

Of course, for any causal relationship, there are changes to both the related variables and the background conditions under which the relationship will *not* hold. The interventionist definition of causality only requires a relationship to hold under at least *some* conditions (e.g., fertilizer *is* a cause of plant growth, because of the dependence between them, even though this relationship will not hold if the plant is kept in the dark or the fertilizer is infested with harmful bugs). Our notion of invariance includes our understanding of these conditions. Vasilyeva and colleagues (2018) explain this "stability" component of invariance as a combination of "breadth," or generality (the range of background circumstances in which the generalization holds) and "guidance," or accuracy (the support a causal explanation provides for generalization to new circumstances). This dual consideration reveals the importance that causal learners place both on identifying causal relationships that are most reliable *and* knowing the contexts in which they may be relied on. Empirical evidence supporting interventionism's role for invariance in our causal thinking comes primarily from studies looking at highly similar concepts, such as "explanation generality" (e.g., Friedman, 1974; Gelman, Star, & Flukes, 2002; Johnston, Sheskin, Johnson, & Keil, 2018; Kitcher, 1981; Strevens, 2009; Walker, Lombrozo, Legare, & Gopnik, 2014). Very recently, computational (Morris et al., 2018) and behavioral (Vasilyeva et al., 2018) studies have begun to look explicitly at interventionist invariance and find that it both reflects and influences our causal judgments.

According to interventionism then, the purpose of causal learning lies in acquiring representations of counterfactual dependence between variables that generalize to guide future action and inference. However, learners are almost never in an epistemic position to form exceptionless generalizations, which would require specifying *every* relevant contributing and enabling factor involved in causing an event. Furthermore, such a specific explanation would also be severely limited in usefulness as it would be applicable to fewer new situations than a less complete account. Instead, causal learners acquire causal explanations that are incomplete generalizations and augment them with evaluations of their invariance over time and across contexts. Given this, it makes sense for causal learners to value information about invariance and to seek to uncover it through the exploration of multiple examples.[6] This is the foundation of the Search for Invariance (SI) hypothesis.

The Search for Invariance (SI) Hypothesis

The SI hypothesis proposes that the positive testing strategy (PTS) may provide a means of assessing this critical characteristic of causal invariance during exploration and hypothesis testing. Since no causal relationship is without exceptions or holds in all circumstances, knowing the extent and contexts in which relationships are invariant is critical for causal knowledge to guide action and reasoning. It is therefore incumbent upon causal learners to determine the invariance of putative causal relationships by asking, for example, the following: Is this dependency reliable? Is it generalizable? And, if so, with what probability and in what contexts? Activating the hypothesized causal variable (or examining cases in which it has already been activated) allows the learner to assess whether the effect behaves as hypothesized in the current context. Repeating these interventions provides evidence for how consistently and extensively this relationship holds. Negative tests—while assessing whether an alternative cause might also bring about the effect in the current context—can not, by definition, provide this information. Thus, positive tests are potentially informative for self-directed learners, *regardless* of their expectations about the sparsity or rarity of causal relationships, by providing evidence about their relative invariance.

[6] See also Chaps 3 and 7, for other instances of how the observation of multiple examples influences how learners generalize their knowledge.

To further illustrate this claim, the following sections will return to three key examples of PTS to demonstrate how reinterpreting these results in terms of the learner's Search for Invariance provides a novel and more complete account of these behaviors.

As an Alternative to "Engineering Desirable Outcomes"

Tschirgi (1980) provides an excellent example of how the SI hypothesis reframes the tendency to conduct positive tests to bring about desirable outcomes. Recall that when asked to select one of three actions to prove that the type of sweetener was the reason that the cake was good and that the types of fat and flour did not matter, the majority of participants chose to HOLD the suspected causal variable and change the others *rather* than VARY only the suspected cause (Tschirgi, 1980).

From the perspective of the SI hypothesis, the HOLD option presents a valuable test of the circumstances under which the suspected causal dependency holds. Baking another cake in which the type of sweetener is kept constant, and all the other ingredients are changed, allows the learner to assess whether this relationship is invariant under different background circumstances.[7] Another example of this kind of hypothesis testing in scientific reasoning comes from Zimmerman and Glaser (2001), who asked 6th graders to test the claim that coffee grounds are good for plants. The authors found that the majority of students designed a series of positive tests; that is, they checked the outcome of adding coffee grounds to a variety of different plants, thereby testing the hypothesized effect of this intervention across a variety of background conditions.

On the other hand, it might be argued that assessing invariance is illogical in a situation in which the causal status of the hypothesized relationship has not yet been conclusively established. Zimmerman (2007) makes precisely this objection, explaining that the participants in Tschirgi (1980) failed to first confirm the claim that the sweetener produces good cake in a controlled manner. However, it is not clear that this confirmation is necessary. The hypothesis presented in the task makes two distinct claims: (1) Good cake is causally dependent on the sweetener, and (2) good cake is causally independent of the type of flour and fat. Tschirgi sees (1) as the *only* claim participants are being asked to test. However, there is nothing to stop participants from choosing to evaluate (2), which would make HOLD, and not VARY, the disconfirming test. This duality means that HOLD is just as valid a test of the hypothesis stated in the problem text as VARY. Further, HOLD has the additional attraction of also assessing the invariance of the causal relationship between the sweetener and good cake that was singled out by the prompt.[8]

[7] In fact, since the scenarios only contained two values for each variable, the HOLD option tests invariance for *all* possible kinds (though not combinations) of other factors.

[8] This is not the only study in the scientific reasoning literature with such ambiguities. Assumptions about parameters—the number of causal variables and whether their effects are independent or interdependent, probabilistic, or deterministic—are regularly made by experimenters but not conveyed to participants or considered when evaluating their behavior. Ongoing work in our lab aims to remove these ambiguities to better assess children's intuitive experimentation.

As an Alternative to "Seeking Confirmatory Evidence"

The SI hypothesis also provides an explanation for participants' repeated testing of instances that are consistent with their current hypothesis, as seen in Wason's 2-4-6 task. Again, learners are testing the invariance of their hypothesized rule, but it is not the same quality of invariance as the one tested in the previous example. In that case, the learner's goal was to assess the range of background circumstances in which a causal claim holds, what Vasilyeva and colleagues (2018) call "breadth." In contrast, explorations in the Wason task are more concerned with "guidance," the *accuracy* of a causal hypothesis for developing expectations about novel circumstances. By checking multiple sets that are all consistent with their current hypothesis (e.g., "increasing even numbers") across distinct instances, learners can determine whether the rule invariantly identifies sets with the target property. Further evidence for this interpretation is what Klayman and Ha (1987) call *limit testing*: Within their repeated positive tests, participants in the 2-4-6 task commonly select extreme or unusual instances of their hypothesized rule. For example, a participant considering the hypothesis "increasing even numbers" might choose to test the set: −2, 0, 2 (Klayman, 1995; Klayman & Ha, 1987).

Again, it might be objected that the current hypothesis has not yet been verified. It is true that evidence of the invariance of "increasing even numbers" does not rule out the possibility that the rule is actually "increasing numbers." However, if any instance of the "increasing even numbers" rule *ever* fails one of these tests of invariance, the learner will know that *neither* hypothesis is correct. If learners were simply hoping to generate confirmatory evidence or produce affirmative responses from the experimenter, then we would *not* expect them to preferentially test their hypotheses at the regions of highest uncertainty (i.e., at the *limits*). Instead, these investigations serve as a stress test of the invariance of their hypothesis, even at its boundaries.

Beyond Evidence of Sufficiency

Another likely objection to the SI hypothesis is to claim that invariance is not meaningfully distinguishable from *sufficiency*. A sufficient cause is adequate, but not required, for bringing about an effect (Klayman & Ha, 1987), and researchers have previously proposed the use of sufficient hypotheses as an explanation of PTS. These accounts tend to emphasize pragmatic motivations (e.g., a desire to achieve or avoid specific outcomes) over epistemic ones. For example, Friedrich (1993) suggests that human cognition, which was shaped by a drive to ensure survival, is better suited to identifying sufficient mechanisms than determining truth. Similarly, Schwartz (1982) explains PTS as a kind of error avoidance, motivated in part by the conditions of reward and reinforcement. In other words, once the learner identifies a sufficient cause, they may feel no compulsion to determine whether that condition is also necessary (Nickerson, 1998).

It is difficult to distinguish the SI hypothesis from this alternative account since the occurrence of PTS does not, in itself, indicate the motivation behind it. For example, in the 2-4-6 task, selecting sets that are sufficient (i.e., have a high probability of producing an affirmative response based on what is currently known) and selecting sets that test the invariance of the hypothesized rule can lead to the same actions. Despite this challenge, we maintain that invariance is *not* reducable to sufficiency.

To explain this position, it is essential first to recognize that causal inference and rule-based inference involve different notions of sufficiency. Rule learning typically assumes that there is only one correct rule, which must be both sufficient and necessary (Klayman & Ha, 1987). Further, these conditions are defined in terms of propositional logic (see Johnson-Laird & Byrne, 2002). For example, in the 2-4-6 task, any set that follows the correct rule will have the target property (the rule is sufficient), and all sets that have the target property will follow the rule (the rule is necessary). Causation, on the other hand, does *not* follow the rules of standard inference, and causal logic involves assumptions that cannot be captured by the principles of propositional logic (see Sloman & Lagnado, 2005, 2015).

To illustrate this difference, consider these two pairs of hypotheses: (a) *sets of numbers increasing by two have the target property* and (b) *sets of increasing numbers have the target property* as opposed to (c) *sex causes pregnancy* and (d) *embryo fertilization causes pregnancy*. Both (a) and (c) are cases in which the antecedents ("*increasing by two"* and "*sex"*) are sufficient but not necessary for their consequents ("*having the target property*" and "*pregnancy"*). Sets of numbers increasing by two *will* be in the target set, and sex *will* (under certain background conditions) be a cause of pregnancy. However, the truth of hypotheses (b) and (d) accounts for *why* the antecedents of (a) and (c) bring about their consequents, making them unnecessary. Sex is not necessary for pregnancy (which can also occur through in vitro fertilization), and increasing by two is not necessary for a set to have the target property (as all sets of numbers increasing by two are also sets of increasing numbers).

However, there is also a critical difference here: (a) and (b) are hypotheses about rules, while (c) and (d) are hypotheses about causes. The domain of rule discovery requires that there be only one correct rule that is both necessary and sufficient for the target property. In contrast, the domain of causal reasoning allows for multiple possible causes to exist simultaneously. As a result, (b) being true means (a) cannot be the correct rule, but (d) being true does not mean that (c) cannot be a correct causal explanation. Put another way, the hypothesis that "*the rule for the target property is numbers increasing by two*" is incorrect, but the hypothesis that "*sex is a cause of pregnancy*" is not. The difference between propositional and causal logic means that the impact of one statement's truth value on another's differs between the domains of rule and causal learning. Lack of necessity makes a rule false, but it does *not* make a cause false. That said, it *does* make it less invariant.

Judgments of necessity and sufficiency are critical to our reasoning in the causal domain, but in a way that is unique to the domain. Necessity captures our intuition that causal variables ought to *make a difference* to outcomes: We judge dropping an ice cube on the ground as the cause of it shattering since, in the absence of the first event, the second would not have occurred. Indeed, a wealth of evidence shows such evaluations are central in our causal judgments (e.g., Gerstenberg, Goodman, Lagnado, & Tenenbaum, 2014; Gerstenberg, Peterson, Goodman, Lagnado, & Tenenbaum, 2017; Icard, Kominsky, & Knobe, 2017; Morris et al., 2018; Wells & Gavanski, 1989).

Causal sufficiency is also intertwined with the notion of necessity. For example, when we credit one event (e.g., dropping an ice cube) as the difference-making cause of another event (e.g., shattering), we understand that the necessity of the first event for the occurrence of the second is *dependent on the context* in which both take place. That is, dropping the ice cube will only cause it to shatter under certain background conditions (e.g., gravity). We also understand that shattering might occur in other cases, even in the absence of dropping (e.g., if someone hits the ice cube with a hammer).

Thus, unlike a rule, a single causal variable is never sufficient *or* necessary for bringing about an effect in and of itself (Mackie, 1974). In causal reasoning, these qualities exist only given certain background conditions in which the occurrence of the variable makes a difference to the outcome. Our knowledge of invariance, of what these conditions are and the degree to which they hold, captures this. It thereby requires *consideration of necessity* and cannot be reduced to assessment or employment of sufficiency.

As an Account of Previously Ambiguous Evidence

Finally, the SI hypothesis also provides a consistent interpretation of learners' tendency to intervene on the root node in McCormack et al. (2016) and Meng et al. (2018).[9] Recall that according to interventionism, the commitments a causal hypothesis makes about the potential for action and manipulation are central to our reasoning about it. In fact, according to Woodward, the difference between competing causal models is *understood* in terms of the difference in the predicted outcomes of interventions (2006, p. 61). This suggests that a causal learner might not feel the need to distinguish between causal hypotheses that make the same predictions about the majority or more salient interventions on a system.

The problems used in McCormack et al. (2016) and Meng et al. (2018) place participants in precisely this situation: The competing possible causal structures all indicate the same variable as the root node of the system. Activating this node, therefore, cannot distinguish between the competing hypotheses as the expected outcome is the same in each. However, considered in light of the SI hypothesis, children's preference for this action is not necessarily uninformative. Precisely because all the hypotheses predict that intervention on this variable will activate all the other variables in the system, the interventionist difference-making perspective may not meaningfully distinguish between these possibilities. Given this, the primary concern for a causal learner would be to assess the degree to which manipulation of this putative root cause reliably leads to its predicted effects, rather than to disambiguate exactly how it does so, which fits the behavior seen in the task.

[9] As a reminder, the "root node" is the starting point for the causal model of the system. Here, the structure is either a common cause (activating component A causes components B and C to activate) or a causal chain (activating A causes B to activate, which causes C to activate, etc.), meaning component A is the root node in both cases.

Conclusion: Relationship to Truth

The claim that we are "intuitive scientists" in our exploratory learning is well established as part of the account of human inquiry (Coenen et al., 2018). However, the fact that self-directed learners often choose to conduct repeated positive tests of their hypotheses rather than (apparently) more informative interventions has historically complicated this claim. Repeated positive testing is characteristic of exploratory behavior across development, yet it would seem to be at odds with the aims of self-directed information seeking. In this chapter, we introduced a novel account of this behavior—the Search for Invariance (SI) hypothesis—which suggests that seeking multiple positive examples may in fact serve the information-seeking goals of causal learning.

To summarize, the SI hypothesis draws on the interventionist framework of causal reasoning, which suggests that causal learners are concerned with the *invariance* of candidate hypotheses. In a probabilistic and interdependent causal world, our primary goal is to determine whether (and in what contexts) our current hypothesis provides an accurate basis for inference and intervention—not to disconfirm alternatives. By recognizing the central role of invariance in causal learning, positive testing may be reinterpreted as a rational and necessary information-seeking strategy. The SI hypothesis therefore provides an explanation of PTS that accords with theories that portray self-directed learning as intuitive science. Of course, empirical work is needed to establish the importance of invariance to learning and to specify *how* learners form estimates of invariance from the multiple examples they generate, what type of examples are needed, and how many. By providing a novel approach to PTS, we hope that the SI hypothesis will serve as a promising theoretical foundation to guide future work.

That said, by aligning PTS with theories of the intuitive experimentation, it is also important to acknowledge that the "learner-as-scientist" approach typically emphasizes increasing the *accuracy* of current knowledge as the primary goal of self-directed exploration. Indeed, Gopnik and colleagues (e.g., Gopnik, 1998, 2000; Gopnik & Walker, 2013) have variously asserted that the raison d'etre of our self-directed causal learning is to form veridical causal models of the world. According to this view, learning as an "intuitive scientist" ought to be characterized by movement toward more accurate knowledge—and the proposal that causal learners are more concerned with assessing the invariance of a causal explanation than whether it is more accurate than alternatives may seem initially incompatible.

However, recall that the interventionist account of causality is inherently tied to action. As causal learners, we are concerned with refining the accuracy of our causal models but *only* insofar as this meaningfully improves our ability to predict, explain, and manipulate the world. These goals do not require absolute accuracy; they require *reliability.* The SI hypothesis does not aim to imply that learners are disinterested in evaluating the truth of competing hypotheses but that the concerns and priorities of causal learning will determine which aspects of these hypotheses are most informative to evaluate.

References

Baron, J., Beattie, J., & Hershey, J. C. (1988). Heuristics and biases in diagnostic reasoning: II. Congruence, information, and certainty. *Organizational Behavior and Human Decision Processes, 42*(1), 88–110.

Blanchard, T., Vasilyeva, N., & Lombrozo, T. (2018). Stability, breadth and guidance. *Philosophical Studies, 175*(9), 2263–2283.

Bonawitz, E. B., van Schijndel, T. J. P., Friel, D., & Schulz, L. (2012). Children balance theories and evidence in exploration, explanation, and learning. *Cognitive Psychology, 64*(4), 215–234. https://doi.org/10.1016/j.cogpsych.2011.12.002

Bramley, N. R., Lagnado, D. A., & Speekenbrink, M. (2015). Conservative forgetful scholars: How people learn causal structure through sequences of interventions. *Journal of Experimental Psychology: Learning Memory and Cognition*. https://doi.org/10.1037/xlm0000061

Brewer, W. F., & Samarapungavan, A. (1991). Children's theories vs. scientific theories: Differences in reasoning or differences in knowledge? In R. R. Hoffman & D. S. Palermo (Eds.), *Cognition and the symbolic processes: Applied and ecological perspectives* (pp. 209–232). Hillsdale, NJ: Erlbaum.

Carey, S. (1985). *Conceptual change in childhood. The MIT series in learning development and conceptual change*. Cambridge, MA: MIT Press. https://doi.org/10.1016/S0016-6995(85)80176-5

Carey, S., Evans, R., Honda, M., Jay, E., & Unger, C. (1989). 'An experiment is when you try it and see if it works': A study of grade 7 students' understanding of the construction of scientific knowledge. *International Journal of Science Education, 11*(5), 514–529. https://doi.org/10.1080/0950069890110504

Coenen, A., Nelson, J. D., & Gureckis, T. M. (2018). Asking the right questions about the psychology of human inquiry: Nine open challenges. *Psychonomic Bulletin & Review*, 1–41. https://doi.org/10.3758/s13423-018-1470-5

Coenen, A., Rehder, B., & Gureckis, T. M. (2015). Strategies to intervene on causal systems are adaptively selected. *Cognitive Psychology, 79*, 102–133.

Cook, C., Goodman, N. D., & Schulz, L. E. (2011). Where science starts: Spontaneous experiments in preschoolers' exploratory play. *Cognition, 120*(3), 341–349. https://doi.org/10.1016/j.cognition.2011.03.003

Croker, S., & Buchanan, H. (2011). Scientific reasoning in a real-world context: The effect of prior belief and outcome on children's hypothesis-testing strategies. *British Journal of Developmental Psychology, 29*(3), 409–424.

Devine, P. G., Hirt, E. R., & Gehrke, E. M. (1990). Diagnostic and confirmation strategies in trait hypothesis testing. *Journal of Personality and Social Psychology, 58*(6), 952.

Dunbar, K., & Klahr, D. (1989). Developmental differences in scientific discovery processes. In D. Klahr & K. Kotovsky (Eds.), *Complex information: the impact of Herbert A. Simon* (pp. 109–143). Hillsdale, NJ: L. Erlbaum Associates. https://doi.org/10.1109/cipe.2004.1428147

Einhorn, H. J., & Hogarth, R. M. (1978). Confidence in judgment: Persistence of the illusion of validity. *Psychological Review, 85*(5), 395.

Friedman, M. (1974). Explanation and scientific understanding. *The Journal of Philosophy, 71*(1), 5–19.

Friedrich, J. (1993). Primary error detection and minimization (PEDMIN) strategies in social cognition: A reinterpretation of confirmation bias phenomena. *Psychological Review, 100*(2), 298.

Gelman, S. A., Star, J. R., & Flukes, J. (2002). Children's use of generics in inductive inferences. *Journal of Cognition and Development, 3*(2), 179–199.

Gerstenberg, T., Goodman, N., Lagnado, D. A., & Tenenbaum, J. (2014). *From counterfactual simulation to causal judgment*. Proceedings of the annual meeting of the Cognitive Science Society (Vol. 36).

Gerstenberg, T., Peterson, M. F., Goodman, N. D., Lagnado, D. A., & Tenenbaum, J. B. (2017). Eye-tracking causality. *Psychological Science, 28*(12), 1731–1744.

Gibson, J. J. (1979). *The ecological approach to visual perception*. Boston, MA: Houghton Mifflin.

Gopnik, A. (1998). Explanation as orgasm. *Minds and Machines, 8*(1), 101–118. https://doi.org/10.1023/A:1008290415597

Gopnik, A. (2000). Explanation as orgasm and the drive for causal knowledge: The function, evolution, and phenomenology of the theory formation system. In *Explanation and cognition* (pp. 299–323). Cambridge, MA: The MIT Press.

Gopnik, A., & Meltzoff, A. N. (1997). *Words, thoughts, and theories. Learning, development, and conceptual change*. Cambridge, MA: MIT Press. https://doi.org/10.1057/9781137322746

Gopnik, A., & Walker, C. M. (2013). Considering counterfactuals: The relationship between causal learning and pretend play. *American Journal of Play, 6*, 15–28.

Gorman, M. E., & Gorman, M. E. (1984). A comparison of disconfirmatory, confirmatory and control strategies on Wason's 2–4–6 task. *The Quarterly Journal of Experimental Psychology, 36*(4), 629–648.

Gweon, H., & Schulz, L. E. (2008). *Stretching to learn: Ambiguous evidence and variability in preschoolers exploratory play*. Proceedings of the 30th annual meeting of the Cognitive Science Society.

Heyman, G. D., & Dweck, C. S. (1992). Achievement goals and intrinsic motivation: Their relation and their role in adaptive motivation. *Motivation and Emotion, 16*(3), 231–247.

Hitchcock, C. (2012). Portable causal dependence: A tale of consilience. *Philosophy of Science, 79*(5), 942–951. https://doi.org/10.1086/667899

Icard, T. F., Kominsky, J. F., & Knobe, J. (2017). Normality and actual causal strength. *Cognition, 161*, 80–93.

Inhelder, B., & Piaget, J. (1958). *The growth of logical thinking from childhood to adolescence: An essay on the construction of formal operational structures*. New York: Basic Books.

Johnson-Laird, P. N., & Byrne, R. M. J. (2002). Conditionals: A theory of meaning, pragmatics, and inference. *Psychological Review, 109*(4), 646.

Johnston, A. M., Sheskin, M., Johnson, S. G. B., & Keil, F. C. (2018). Preferences for explanation generality develop early in biology but not physics. *Child Development, 89*(4), 1110–1119.

Karmiloff-Smith, A. (1988). The child is a theoretician, not inductivist. *Mind and Language, 3*, 183–195.

Kendler, K. S. (2005). "A gene for…": The nature of gene action in psychiatric disorders. *American Journal of Psychiatry, 162*(7), 1243–1252.

Kitcher, P. (1981). Explanatory unification. *Philosophy of Science, 48*(4), 507–531.

Klahr, D., & Chen, Z. (2003). Overcoming the positive-capture strategy in young children: Learning about indeterminacy. *Child Development, 74*(5), 1275–1296.

Klahr, D., & Dunbar, K. (1988). Dual space search during scientific reasoning. *Cognitive Science, 12*(1), 1–48.

Klahr, D., Fay, A. L., & Dunbar, K. (1993). Heuristics for scientific experimentation: A developmental study. *Cognitive Psychology, 25*(1), 111–146. https://doi.org/10.1006/cogp.1993.1003

Klayman, J. (1995). Varieties of confirmation bias. In D. L. Medin, J. R. Busemeyer, & R. Hastie (Eds.), *The psychology of learning and motivation: Decision making from the perspective of cognitive psychology.* New York: Academic.

Klayman, J., & Ha, Y.-W. (1987). Confirmation, disconfirmation, and information in hypothesis testing. *Psychological Review, 94*(2), 211.

Kuhn, D. (1989). Children and adults as intuitive scientists. *Psychological Review*. https://doi.org/10.1037/0033-295X.96.4.674

Kuhn, D., Amsel, E., O'Loughlin, M., Schauble, L., Leadbeater, B., & Yotive, W. (1988). *The development of scientific thinking skills*. San Diego, CA: Academic Press.

Kuhn, D., & Angelev, J. (1976). An experimental study of the development of formal operational thought. *Child Development*, 697–706.

Kuhn, D., & Brannock, J. (1977). Development of the isolation of variables scheme in experimental and "natural experiment" contexts. *Developmental Psychology, 13*(1), 9.

Kuhn, D., & Phelps, E. (1982). The development of problem-solving strategies. *Advances in Child Development and Behavior, 17*(C), 1–44. https://doi.org/10.1016/S0065-2407(08)60356-0

Lapidow, E., & Walker, C. M. (2019). *Does the intuitive scientist conduct informative experiments?: Children's early ability to select and learn from their own interventions*. 41st annual meeting of the Cognitive Science Society.

Lewis, D. (1974). Causation. *The Journal of Philosophy, 70*(17), 556–567.

Lombrozo, T., & Carey, S. (2006). Functional explanation and the function of explanation. *Cognition, 99*(2), 167–204. https://doi.org/10.1016/j.cognition.2004.12.009

Mackie, J. L. (1974). *The cement of the universe: A study of causation*. Oxford: Clarendon Press.

Mahoney, M. J., & DeMonbreun, B. G. (1977). Psychology of the scientist: An analysis of problem-solving bias. *Cognitive Therapy and Research, 1*(3), 229–238.

McCormack, T., Bramley, N. R., Frosch, C., Patrick, F., & Lagnado, D. A. (2016). Children's use of interventions to learn causal structure. *Journal of Experimental Child Psychology*. https://doi.org/10.1016/j.jecp.2015.06.017

McKenzie, C. R. M. (2004). Hypothesis testing and evaluation. In D. Koehler & N. Harvey (Eds.), *Blackwell handbook of judgment and decision making* (pp. 200–219). Oxford: Blackwell.

McKenzie, C. R. M., & Mikkelsen, L. A. (2000). The psychological side of Hempel's paradox of confirmation. *Psychonomic Bulletin & Review, 7*(2), 360–366.

Meng, Y., Bramley, N. R., & Xu, F. (2018). *Children's causal interventions combine discrimination and confirmation*. Proceedings of the 40th annual conference of the Cognitive Science Society.

Morris, A., Phillips, J. S., Icard, T. F., Knobe, J., Gerstenberg, T., & Cushman, F. (2018). Causal judgments approximate the effectiveness of future interventions. https://doi.org/10.31234/osf.io/nq53z

Navarro, D. J., & Perfors, A. F. (2011). Hypothesis generation, sparse categories, and the positive test strategy. *Psychological Review, 118*(1), 120.

Nickerson, R. S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology, 2*(2), 175–220.

Oaksford, M., & Chater, N. (1994). A rational analysis of the selection task as optimal data selection. *Psychological Review, 101*(4), 608.

Pearl, J., & Bareinboim, E. (2011). *Transportability of causal and statistical relations: A formal approach*. Twenty-fifth AAAI conference on Artificial Intelligence.

Popper, K. R. (1959). The logic of scientific discovery. *Physics Today*. https://doi.org/10.1063/1.3060577

Redhead, M. (1987). *Incompleteness, nonlocality, and realism: A prolegomenon to the philosophy of quantum mechanics*. Oxford: Clarendon Press.

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science, 274*(5294), 1926–1928. https://doi.org/10.1126/science.274.5294.1926

Saffran, J. R., Johnson, E. K., Aslin, R. N., & Newport, E. L. (1999). Statistical learning of tone sequences by human infants and adults. *Cognition.* https://doi.org/10.1016/S0010-0277(98)00075-4

Schauble, L. (1990). Belief revision in children: The role of prior knowledge and strategies for generating evidence. *Journal of Experimental Child Psychology, 49*(1), 31–57. https://doi.org/10.1016/0022-0965(90)90048-D

Schauble, L., Glaser, R., Duschl, R. A., Schulze, S., & John, J. (1995). Students' understanding of the objectives and procedures of experimentation in the science classroom. *The Journal of the Learning Sciences, 4*(2), 131–166.

Schauble, L., Klopfer, L. E., & Raghavan, K. (1991). Students' transition from an engineering model to a science model of experimentation. *Journal of Research in Science Teaching, 28*(9), 859–882. https://doi.org/10.1002/tea.3660280910

Schulz, L. E. (2012). The origins of inquiry: Inductive inference and exploration in early childhood. *Trends in Cognitive Sciences*. https://doi.org/10.1016/j.tics.2012.06.004

Schulz, L. E., & Bonawitz, E. B. (2007). Serious fun: Preschoolers engage in more exploratory play when evidence is confounded. *Developmental Psychology*. https://doi.org/10.1037/0012-1649.43.4.1045

Schulz, L. E., Standing, H. R., & Bonawitz, E. B. (2008). Word, thought, and deed: The role of object categories in children's inductive inferences and exploratory play. *Developmental Psychology*. https://doi.org/10.1037/0012-1649.44.5.1266

Schwartz, B. (1982). Reinforcement-induced behavioral stereotypy: How not to teach people to discover rules. *Journal of Experimental Psychology: General, 111*(1), 23.

Siegler, R. S., & Liebert, R. M. (1975). Acquisition of formal scientific reasoning by 10-and 13-year-olds: Designing a factorial experiment. *Developmental Psychology, 11*(3), 401.

Siler, S. A., & Klahr, D. (2012). Detecting, classifying, and remediating: Children's explicit and implicit misconceptions about experimental design. *Psychology of Science: Implicit and Explicit Processes*. https://doi.org/10.1093/acprof:oso/9780199753628.003.0007

Siler, S. A., Klahr, D., & Price, N. (2013). Investigating the mechanisms of learning from a constrained preparation for future learning activity. *Instructional Science, 41*(1), 191–216.

Skov, R. B., & Sherman, S. J. (1986). Information-gathering processes: Diagnosticity, hypothesis-confirmatory strategies, and perceived hypothesis confirmation. *Journal of Experimental Social Psychology, 22*(2), 93–121.

Sloman, S. A. (2005). *Causal models : How people think about the world and its alternatives*. Oxford, NY: Oxford University Press.

Sloman, S. A., & Lagnado, D. A. (2005). Do We "do"? *Cognitive Science, 29*(1), 5–39. https://doi.org/10.1207/s15516709cog2901_2

Sloman, S. A., & Lagnado, D. A. (2015). Causality in thought. *Annual Review of Psychology, 66*, 223–247.

Sodian, B., Zaitchik, D., & Carey, S. (1991). Young children's differentiation of hypothetical beliefs from evidence. *Child Development, 62*(4), 753–766.

Strevens, M. (2009). *Depth: An account of scientific explanation*. Cambridge, MA: Harvard University Press.

Tschirgi, J. E. (1980). Sensible reasoning: A hypothesis about hypotheses. *Child Development, 51*(1), 1–10. https://doi.org/10.2307/1129583

Tukey, D. D. (1986). A philosophical and empirical analysis of subjects' modes of inquiry in Wason's 2–4–6 task. *The Quarterly Journal of Experimental Psychology Section A, 38*(1), 5–33.

Tweney, R. D., Doherty, M. E., Worner, W. J., Pliske, D. B., Mynatt, C. R., Gross, K. A., et al. (1980). Strategies of rule discovery in an inference task. *Quarterly Journal of Experimental Psychology, 32*(1), 109–123.

Valanides, N., Papageorgiou, M., & Angeli, C. (2014). Scientific investigations of elementary school children. *Journal of Science Education and Technology, 23*(1), 26–36. https://doi.org/10.1007/s10956-013-9448-6

van Schijndel, T. J. P., Visser, I., van Bers, B. M. C. W., & Raijmakers, M. E. J. (2015). Preschoolers perform more informative experiments after observing theory-violating evidence. *Journal of Experimental Child Psychology, 131*, 104–119. https://doi.org/10.1016/j.jecp.2014.11.008

Vasilyeva, N., Blanchard, T., & Lombrozo, T. (2018). Stable causal relationships are better causal relationships. *Cognitive Science, 42*(4), 1265–1296.

Vogel, R., & Annau, Z. (1973). An operant discrimination task allowing variability of reinforced response patterning 1. *Journal of the Experimental Analysis of Behavior, 20*(1), 1–6.

Walker, C. M., Lombrozo, T., Legare, C. H., & Gopnik, A. (2014). Explaining prompts children to privilege inductively rich properties. *Cognition, 133*(2), 343–357. https://doi.org/10.1016/j.cognition.2014.07.008

Wason, P. C. (1960). On the failure to eliminate hypotheses in a conceptual task. *Quarterly Journal of Experimental Psychology, 12*(3), 129–140.

Wason, P. C. (1962). Reply to wetherick. *Quarterly Journal of Experimental Psychology, 14*(4), 250.

Wason, P. C. (1968). Reasoning about a rule. *Quarterly Journal of Experimental Psychology, 20*(3), 273–281.

Wason, P. C., & Johnson-Laird, P. N. (1972). *Psychology of reasoning: Structure and content* (Vol. 86). Cambridge, MA: Harvard University Press.

Wells, G. L., & Gavanski, I. (1989). Mental simulation of causality. *Journal of Personality and Social Psychology, 56*(2), 161.

Weslake, B. (2010). Explanatory depth. *Philosophy of Science, 77*(2), 273–294.

Wetherick, N. E. (1962). Eliminative and enumerative behaviour in a conceptual task. *Quarterly Journal of Experimental Psychology, 14*(4), 246–249.

Woodward, J. (1997). Explanation, invariance, and intervention. *Philosophy of Science, 64*, S26–S41.

Woodward, J. (2003). *Making things happen: A theory of causal explanation*. New York: Oxford University Press.

Woodward, J. (2006). Sensitive and insensitive causation. *The Philosophical Review, 115*(1), 1–50.

Woodward, J. (2010). Causation in biology: Stability, specificity, and the choice of levels of explanation. *Biology and Philosophy, 25*(3), 287–318.

Wu, R., Gopnik, A., Richardson, D. C., & Kirkham, N. Z. (2011). Infants learn about objects from statistics and people. *Developmental Psychology, 47*(5), 1220–1229. https://doi.org/10.1037/a0024023

Yang, S. C., Vong, W. K., Yu, Y., & Shafto, P. (2019). A unifying computational framework for teaching and active learning. *Topics in Cognitive Science, 11*, 316–337.

Ylikoski, P., & Kuorikoski, J. (2010). Dissecting explanatory power. *Philosophical Studies, 148*(2), 201–219.

Yoon, E. J., MacDonald, K., Asaba, M., Gweon, H., & Frank, M. C. (2018). *Balancing informational and social goals in active learning*. Proceedings of the 40th annual meeting of the Cognitive Science Society.

Zimmerman, C. (2000). The development of scientific reasoning skills. *Developmental Review, 20*(1), 99–149. https://doi.org/10.1006/drev.1999.0497

Zimmerman, C. (2007). The development of scientific thinking skills in elementary and middle school. *Developmental Review, 27*(2), 172–223. https://doi.org/10.1016/j.dr.2006.12.001

Zimmerman, C., & Glaser, R. (2001). Testing positive versus negative claims: A preliminary investigation of the role of cover story on the assessment of experimental design skills. *CSE Technical Report*.

Zimmerman, C., & Klahr, D. (2018). Development of scientific thinking. In J. Wixted (Ed.), *Stevens' handbook of experimental psychology and cognitive neuroscience* (4th ed., pp. 1–25). New York: Wiley. https://doi.org/10.1002/9781119170174.epcn407

Chapter 11
Multiple Exemplars of Relations

Stella Christie

Abstract *My mom, mother rabbit, mother country*—these are all multiple exemplars of the relational concept "mother." How do we come to understand that these are exemplars of the same concept? This chapter explains the mechanisms for learning *about* multiple exemplars, particularly multiple exemplars of *relations*. I discuss why perceiving relational exemplars is difficult, and how *structure mapping theory* (Gentner, Cogn Sci Multidiscip J 7(2):155–170, 1983) provides precise learning mechanisms that learners use to understand relational exemplars. Given the scope of the problems and solutions, a new area of research emerges: *social relational learning*. The social world is fundamentally characterized by relations such as kinships, friendships, alliances, and social hierarchies; these relations govern behavior and have far-reaching consequences (friends help; foes do not). Understanding social learning as a relational learning problem gives insight to how learners acquire complex knowledge about their social world—such as differentiating various exemplars of friends versus foes.

You see a picture of four women. Whom does this picture represent? Are they a quartet of *musicians*, are they *mothers*, or are they (merely) a group of unrelated women? In other words, what sort of *exemplar* does this picture represent? To decide this, you look for certain cues: Are they carrying musical instruments? Are they pushing strollers? You look for such cues as markers of potential relations, knowing full well that such cues can mislead you: many mothers will not be pushing a stroller and many non-mothers may well be pushing one. Indeed, recognizing a mother on a picture cannot be reduced to checking one surface feature because being a mother is a relational concept. As you hold the picture in your hand, you focus on some cues and discard others. Many cues invariably escape your attention: Did you notice that each woman holds a piece of paper or a cell phone? Do they display tickets? Are the women *passengers*? Consciously or not, you focus on some potential relations and ignore others. Why and why the mind deals with *exemplars of relations* is the focus of this chapter.

S. Christie (✉)
Tsinghua University, Beijing, China
e-mail: schrist3@swarthmore.edu; christie@mail.tsinghua.edu.cn

J. B. Childers (ed.), *Language and Concept Acquisition from Infancy Through Childhood*, https://doi.org/10.1007/978-3-030-35594-4_11

Exemplars of Relations—Unique Challenges

Learning about and from multiple exemplars of relations has a unique complexity and set of challenges, which are unlike those encountered in non-relational contexts. Potential relations constantly compete for our attention with surface features (e.g., perhaps the women on the picture are wearing blue jeans). We may focus our attention on one candidate relation and ignore another, but even before we do that, we must ask what relations have been overlooked. Were they overlooked because they were not salient enough or, perhaps, because the learner did not even know the relation? Should we assume that a relation must be known before it can be recognized? If not, what allows a previously unknown relation to spring out of a set of (two or twenty-four) exemplars? Indeed, how many exemplars need one see before recognizing a given relation?

These and other questions, which define the landscape of learning from relational exemplars, are the subject of this chapter. I adopt the theory of structural alignment, which posits that recognizing relations is aided by the presence of alignable surface features. After reviewing the problems of recognizing multiple relational exemplars and explaining how structural alignment solves them, I also discuss an example domain where much of this material finds an interesting manifestation: the social world, which is woven from an intricate network of relations.

Why Is It Difficult to Perceive Multiple Exemplars of Relations?

Not Knowing the Relations

Learners may not be able to perceive the sameness of two relations because they do not have the requisite relational concept in the first place. This is true both among novice undergraduates who do not yet possess relational concepts such as *idempotent* or *chirality* and for 2-year-olds who do not yet know the relational concept of *left-right, passenger*, or even *uncle* (Keil, 1998; Waxman & Hall, 1993). These 2-year-olds will not be able to recognize the bearded man on a bus and the tall woman on a train as multiple exemplars of the same concept—a passenger.

One dominant solution to this type of problem is to caption multiple exemplars of relations with relational labels. Relational labels name relational concepts—such as *idempotent* or *passenger*—making them easy to demarcate, reify, and remember (Gentner & Christie, 2010). Relational labels function like a gift wrap: It bestows on a nondescript (or even unknown) item a special status, making it more portable and transferable. Likewise, a labeled concept can be more easily recalled and used in diverse contexts, for example, for a passenger of a bus or a train alike. A large body of evidence shows that relational labels are instrumental for learning relational concepts. For example, Casasola (2005) has shown that 18-month-olds who heard the spatial relational term *on* were better able to recognize multiple exemplars of the support relation than were those who heard general labels. This study illustrates that learning requires hearing relational labels specifically, not just hearing language per se (see also Chap. 3).

The Relations Are Known, but the **Relational Similarity** *Is Not Salient*

At other times, learners may know the relational concept but still have difficulty recognizing new exemplars of the relation. To appreciate this point, let us turn to the simplest relation: the identity relation—recognizing that in an O-O sample, the one O is the same as the other O. Several studies have shown that the concept of identity, or at least the ability to recognize identical repetition, is one of the most fundamental cognitive concepts. For example, 7-month-olds can learn algebraic rules more easily from patterns that contain identical repetition (ABB or AAB) than from other patterns (ABA) (Marcus, Vijayan, Rao, & Vishton, 1999). Newborns show a specific and unique brain activation in the superior temporal and left inferior frontal regions when hearing patterns that involve the identity relation, which does not occur when equally complex patterns without identity relations are played (Gervain, Berent, & Werker, 2012).

Despite an early ability to sense the identity relation (or at least sensing the match repetition), infants do not spontaneously recognize that AA is another exemplar of OO. Seven- and nine-month-olds who were habituated to an exemplar of the identity relation [e.g., fish-fish] did not spontaneously recognize a new identity relation [mouse-mouse] test event as familiar; their looking time did not differ between the new identity relation [mouse-mouse] and a nonidentity relation [mouse-dog] test event (Ferry, Hespos, & Gentner, 2015; see also Chap. 5). In a triad version of this task, 2- and 3-year-olds could not match OO (e.g., square-square) to AA (e.g., circle-circle, choosing randomly between AA (circle-circle) and BC (triangle-pentagon) (Christie & Gentner, 2014). For whatever reason, the relational similarity across these examples of the identity relation is not salient to young children, even though they likely possess the relational concept *identity*.

This problem is compounded when another kind of similarity—an object or a feature similarity—is in competition with the relational one. When the triad task is changed to match OO to either AA (relational match) or OC (object match) (see Fig. 11.1), even 4-year-olds overwhelmingly choose the OC object match (Christie & Gentner, 2007). This is not a wrong choice per se, but it shows that under this circumstance, 4-year-olds are even less likely to recognize that AA and OO are multiple exemplars of the same relation. This tension between relational and object matches is robust in development. Children tend to undergo a *relational shift* sometime during the preschool years (Gentner, 1988; Gentner & Rattermann, 1991; Halford, 1987, 1992) in that they start out focusing on object matches and only later in development focus more on relational matches.

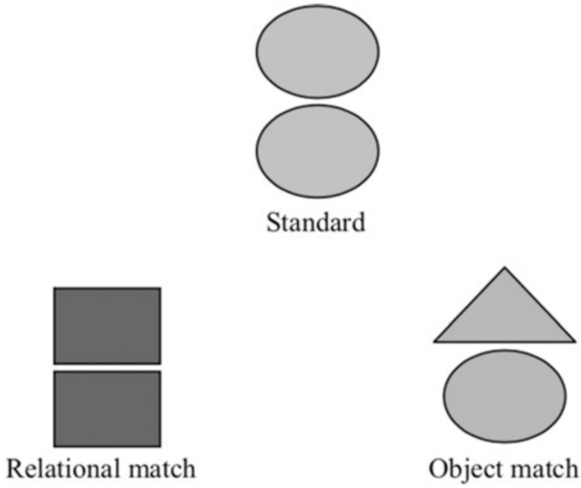


Fig. 11.1 Object versus relational matches

The Solution: Structural Alignment

The dominant solution to the problem of noticing relational commonality is the process of structure mapping. Structure mapping theory (SMT, Gentner, 1983) posits that when learners can align two events, they are more likely to notice the common structure or relations between them, above and beyond noticing feature similarities. For example, 4-year-olds who saw an exemplar of a bicycle were more likely to liken the bicycle with a pair of glasses (because they share round shapes) than to a skateboard (relational commonality; both are vehicles). However, when the 4-year-olds had a chance to compare a bicycle and a tricycle, they were more likely to liken these to the skateboard (Gentner & Namy, 1999; Namy & Gentner, 2002). That is, the alignment process highlights the relational commonality—something that may not be very salient to the learners prior to the comparison.

Learning from multiple exemplars is the essence of SMT. Structure mapping takes one kind of commonality among multiple exemplars (their common surface features) to enable the learning of another type of commonality among exemplars—the relational commonality. In the example above, the initial surface commonality was the round shape of the bicycle and tricycle wheels, and the learning outcome was their common structure (vehicles). Consequently, we should expect that after having abstracted the common relation vehicle, learners will be able to notice other exemplars of this relation. Indeed, children in the Namy and Gentner (2002) now recognized that a skateboard, and not glasses, was the one that shares commonality with the bicycle and tricycle.

Recall that the problem we want to solve—how to recognize multiple exemplars of the same relation—involves two main difficulties. One is that learners may not know the relation in the first place; the other is that learners may know the relation but still fail to perceive the relational similarity across the examples. This may happen because another kind of similarity—feature similarity or object matches—can be more salient than the relational one. Structure mapping, or the alignment of exemplars, addresses both difficulties.

Christie and Gentner (2010) asked if aligning multiple exemplars can result in learning completely new relations. This is an instance of the first difficulty outlined above—not knowing the relation. In this study, we presented arbitrary new relations that are not lexicalized in English in a simple way, for example, black on top, white on bottom (see Fig. 11.2). Unlike the vehicle example (Namy & Gentner, 2002) in which 4-year-olds may already know about vehicles as a category, it is safe to assume that the relations in this study were unfamiliar to the 4-year-old subjects. The question was whether aligning multiple exemplars would result in learning about the novel relation. If so, children should recognize other exemplars of the relation.

To test this, 4-year-olds in the solo condition saw one exemplar of black on top, white on bottom (e.g., a black dog above a white dog), while those in the comparison condition saw two exemplars (e.g., a black dog above a white dog and a black cat above a white cat; Fig. 11.2). The exemplars were labeled with a novel label such as *pepi*. At test, children had to extend this novel label to one of two choices: a picture card depicting an object match (a black dog *next to* a black cat) and a card

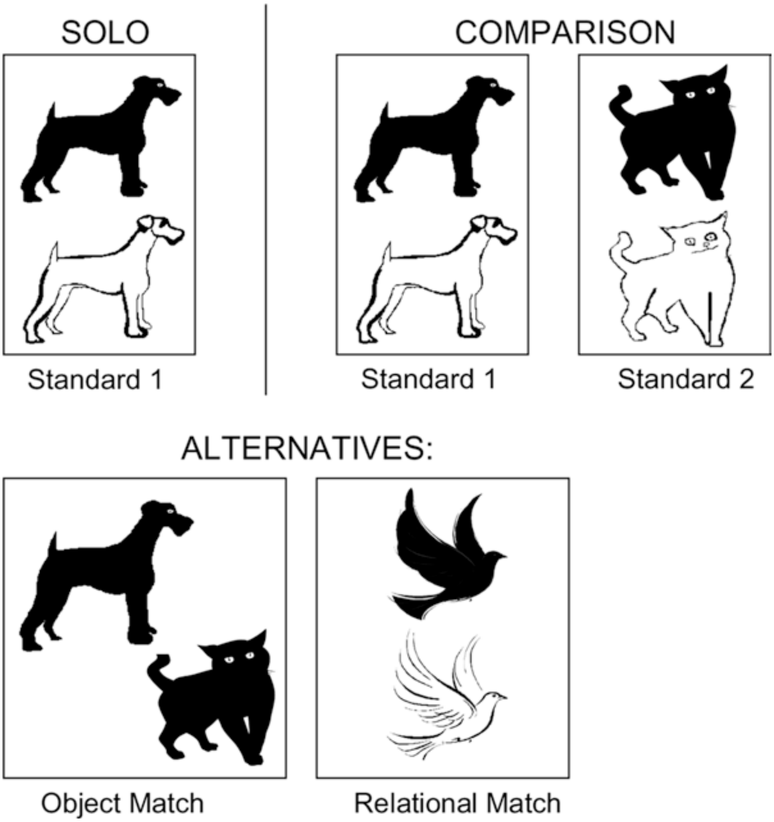


Fig. 11.2 A sample of novel, arbitrary relations for 4-year-olds (black on top, white at the bottom). Children were tested either in the solo (one exemplar) or comparison (two exemplars side by side); both receive the same test, in which they can choose to extend the given novel label either to the object match exemplar or the relational match exemplar

depicting the relational match (a black bird on top and a white bird at bottom). Notice that either choice relies on recognizing some similarity across the multiple exemplars: either a similarity of the shape/features across the examples (the object match) or a multiple exemplar of the relation (relational match). There are no "right" or "wrong" answers in these test trials; rather, we wanted to see if children have a baseline preference for either the object or relational match and whether such a preference could be shifted by the act of comparing and aligning two exemplars.

As expected from prior results, children in the solo condition overwhelmingly chose the object match over the relational match when they extended the novel label. They spontaneously recognized one kind of multiple exemplar—the object match—but not the relational exemplar. Amazingly, this perception changed when they were given an opportunity to compare two exemplars (such as [black dog, white dog] and [black cat, white cat]). After such a comparison, 4-year-olds were more likely to extend the novel label to the relational match. This is notable particularly because in the comparison condition, the feature similarity choice (a black dog *next to* a black cat) contained object matches of *both* standards. Had children persisted in only noticing feature commonalities, they should have been even *more likely* to pick the object match in the comparison group.

Instead, as predicted by structure mapping, comparison—more precisely *aligning* two exemplars—resulted in learners' noticing common relations. We argue that the alignment is crucial; it is not just a matter of more exemplars because in the same study, another group of 4-year-olds saw two exemplars sequentially (thus would have had difficulty aligning them) and chose the object match. The sequential group's behavior did not differ from the solo condition, even though they were shown the same two exemplars as the comparison group. I should note that the exemplars in the sequential condition were presented one after another without any delay. However, because these exemplars contained novel, relatively difficult relations, children likely were not able to hold them in memory and align them. In contrast, children in the comparison condition saw the two exemplars in a spatial juxtaposition which, as seen in Fig. 11.2, we arranged to be optimally aligned. Thus, without the convenient spatial arrangement, children did not appear to have reaped the benefit of comparison and chose the object match instead. These results support our argument that alignment across examples is critical.

We should not confuse the particulars of the methodology—concurrent, sequential, or any other—with the essential ingredient that ushers comparison: alignment. More sophisticated learners who are able to hold exemplars in memory and align them on their own should be able to notice relational commonalities. For example, most adults will have no difficulty in recognizing our new relations [e.g., black on top, white on bottom] after seeing exemplars sequentially. But when the exemplars are sufficiently difficult, even adults should be unable to align them and, consequently, miss relational commonalities. To wit, Loewenstein, Thompson, and Gentner (1999) showed motivated Kellogg MBA students example scenarios of the contingency principle. Those who compared two scenarios simultaneously (comparison group) were 50% better than those who read the scenarios sequentially in using the contingency principle in a later negotiation test.

Comparison resulting in relational learning also overcomes the second difficulty we noted: that children may already know the relations, but the *relational similarity* is not salient. For example, in a Relational Match to Sample Task (RMTS), 2- and 3-year-olds who were tested with the basic identity relation (standard AA, choose between BB and CD) chose at chance. But when they were given an opportunity to align—compare standards AA and EE and then choose either BB or CD—they recognized the other relational exemplar and chose BB as the match to the standards (Christie, 2010).

Comparing What? Similarity of Exemplars

Thus far, I have discussed how comparing multiple exemplars results in noticing relational commonalities. A natural follow-up question is *what kind* of exemplars must be compared to effect relational learning. Suppose that instead of comparing [black dog, white dog] and [black cat, white cat] as in the original study, 4-year-olds compared [black dog, white dog] and another identical card [black dog, white dog]. Would this comparison also produce relational learning? No. In pilot data, we found that when 4-year-olds compared two identical picture cards from the Christie and Gentner (2010) study, they also chose the object match and not the relational match. Namy, Gentner, and Clepper (2007) found that while comparing similar (but discriminable) exemplars of objects helped children to classify the objects based on relational similarity, those who compared nearly identical exemplars classified objects based on perceptual commonality.

The implication is that one cannot just compare any exemplars and expect alignment and structure mapping to take place. The *similarity* of the exemplars matters: Comparing too similar a set of exemplars does not result in learning, while a set of exemplars that is too diverse is more challenging to align. Because there are multiple senses of similarity (see Markman & Gentner, 1993) for a classic view on similarity), here I refer to surface similarity—similarity based on surface features such as shapes or colors (for object sets). With this notion of similarity, the question of how similar exemplars must be to produce relational learning has a predictably ambiguous answer: it depends. There is some optimal level of similarity, or at least a range of similarity, that is optimal for relational learning. What this optimal similarity is depends on the domain to be learned, as well as the learners' prior knowledge.

For example, Jee et al. (2014) asked what similarity comparison allows adult learners to learn about the categorization of complex geographical faults. They found that overall, participants were better at correctly identifying faults if they contrasted *similar* faults than if they contrasted dissimilar faults. Interestingly, while contrasting similar faults worked better overall, the learning outcome also depended on learners' background knowledge. Participants with some geoscience background knowledge (have taken at least one geoscience course in high school or college) benefitted more from the high-similarity comparison than did participants with no background knowledge. That is, for some of the novice learners, the high-similarity comparison as given in this experiment did not yield a learning benefit. It is possible that for these learners, the high similarity was not high enough.

How Many Exemplars to Compare?

In much of the evidence discussed above, the comparison process specifically involves two exemplars. For example, when 4-year-olds learned novel relations (Christie & Gentner, 2010) or when they learned to notice common identity relations in the RMTS task (Christie, 2010), both involved aligning precisely two examples. Two is of course the minimum number of exemplars needed for comparison, but it may also be optimal in setups where learners are explicitly invited to compare. Two exemplars allow a better possibility of going back and forth between the exemplars, encouraging comparison. There is evidence from an infants' eye tracking study (Oakes & Ribar, 2005) that 4-month-old infants who look back and forth between two events during habituation were better able to categorize basic entities such as cats and dogs. The study did not test three or more habituation events side by side, but my prediction is that with three items, infants would not engage in explicit alignment.

Two is not a magic number, however. A very different number of distinct exemplars—essentially, the more the better—was found to be optimal in a study of learning nonadjacent dependency in grammar. The ability to recognize a dependency between syntactic elements separated by other constituents—for example, *is …-ing* in English phrases like *is running* or *is sitting*—plays a crucial role in language learning; it also represents an interesting case of relational learning. An extensive body of work by Gomez and colleagues (see Sandoval & Gomez, 2013 for review) documents nonadjacent dependency learning. The common paradigm in this context is to use artificial languages and ask under what conditions learners can master grammar rules that govern nonadjacent dependency.

In Gomez (2002) (see Chaps. 2 and 4, for a discussion of statistical learning), adults and 18-month-olds were exposed to three-element strings (e.g., *pel-wadim-rud* (aXd)) drawn from one of two artificial languages (L1 or L2). L1 contained one kind of nonadjacent dependencies (aXd, bXe), and L2 contained another kind (aXe, bXf). After hearing multiple instances of these strings, learners had to distinguish which nonadjacent dependency was correct in the language of their training. Learners' training samples included X's (the interposed elements) drawn from different—larger or smaller—sets: Each learner's sample had 2, 6, 12, or 24 distinct X's. The authors found that both adults and 18-month-olds learned the nonadjacent dependencies only from the most variable learning condition (with 24 distinct X's). Since they did not test richer sets of X's, we may speculate that in this case, even greater variability in the sample might further improve relational learning. Mirroring the remarks concerning similarity and comparison, here too learners' prior knowledge may impact the optimal sample variability. As an example, Gomez and Gerken (2000) replicated the variability success with 15-month-olds but found that 12-month-olds failed to learn nonadjacent dependencies under the same learning conditions.

How to explain the apparent need for a large number of exemplars in Gomez's study? It is possible that the learning here is implicit rather than explicit. Learners were not told to compare exemplars, so rather than actively juxtaposing and aligning the exemplars—which works best with two exemplars—they needed several examples before noticing their commonalities. Two types of studies can be done to verify this hypothesis. First, we can repeat Gomez's setup with explicit instructions to compare: Give participants aX1b and aX2b and ask them to identify similarities across these 2 (or 6, 12, 24) examples. If explicit alignment is optimal with 2, we should now expect the variability = 2 condition to produce better learning. Second, we can test whether there are differences between learning from 24 X's sequentially (as in the original study) versus 12 pairs of X's. While there are no explicit instructions to compare here, if alignment can happen better with 2 exemplars, then we expect learning from 12 pairs of exemplars to be equally good as learning from 24 individual items.

One notable difference between explicit and implicit comparison is the speed of learning. Explicit alignment with two exemplars usually gives immediate results of relational abstraction. For example, in the learning novel relations study (Christie & Gentner, 2010), 4-year-olds who compared two examples [black dog on top, white dog at bottom] and [black cat on top, white cat at bottom] immediately noticed the relational commonalities, choosing the relational match option. There were eight trials in this study, but each trial consisted of a novel relation (e.g., trial 2 was the relation symmetry, which was novel for 4-year-olds; trial 3 was big-medium-small, etc.). In effect, the 4-year-olds had to recognize the relational commonality from only one comparison of two exemplars. In the RMTS study (Christie, 2010), across all the eight trials, 2- and 3-year-olds only had to notice one relation—the identity. But here, too, when we examined performance across the eight trials of comparison, we saw no difference between the beginning and end trials: Children correctly chose the identity relational match right after the first comparison trial. In contrast, it took much longer to learn about nonadjacent dependencies in Gomez (2002): Adults were given hundreds of sentences (aXnb) before being tested. Incidentally, the high volume of the learning sample may explain the need for a greater variability: It is possible that 24 distinct X particles helped to sustain attention across the study, while learners who were trained on fewer X's were exhausted by the samples' monotony. Concerning the speed of learning, in many contexts (e.g., infants), it is impossible to explicitly ask learners to compare. Even for the 4-year-olds in our study who learned new relations right away, when we asked "what's the same about these two [standard exemplars]," they mostly gave no reply. The only viable comparison is often implicit.

Sample variability should not be confused with the similarity of exemplars. As an illustration of the difference, consider redoing Gomez's study with 24 *distinct but similar* X elements, for example, substituting X = *wadim, wabim, wapim, wagim…* for the initially used X = *wadim, kicey, puser, fengle*…. What is the matrix of learning outcomes parameterized by variability (the number of distinct X's) and similarity of exemplars? We may speculate that if exemplars are more similar to one another overall, we will need a greater number of distinct exemplars for the same learning outcome. This question has not been addressed in the literature thus far; it still awaits a first exploration.

How Does Comparison Happen?

When exemplars are available for comparison, does comparison automatically happen? Clearly not, as seen in the lower variability condition above. Participants did not learn the nonadjacent dependency when the intervening element X spanned fewer than 24 types. In some ways, this is surprising: Adult subjects had only one thing to pay attention to in the experiment—multiple exemplars of the nonadjacent relation—and still failed to compare them and learn the rule. On the other hand, we often see in everyday learning that young children spontaneously learn relational concepts. They usually do so via implicit comparisons—adults do not tell children to compare all the time. So what invites learners to compare in the first place? I discuss two major invitations to compare: object similarity and language.

Object Similarity Invites Comparison

Recall that children are highly attracted to surface similarity. For example, in Christie (2010), children presented with the standard AA matched it to the object match AC rather than to the relational match BB. As I discussed in the beginning, noticing multiple exemplars of a surface feature does not require learning: Children immediately and spontaneously match red to red and circle to circle. Such unsolicited attention to object matches can impede relational learning, especially when an object match is pitted against a relational match. Under favorable circumstances, however, attention to object matches can also foster relational learning. This happens when noticing object matches catalyzes a closer scrutiny of the exemplars. Once children attend to object similarities among multiple exemplars, they are more likely to compare them.

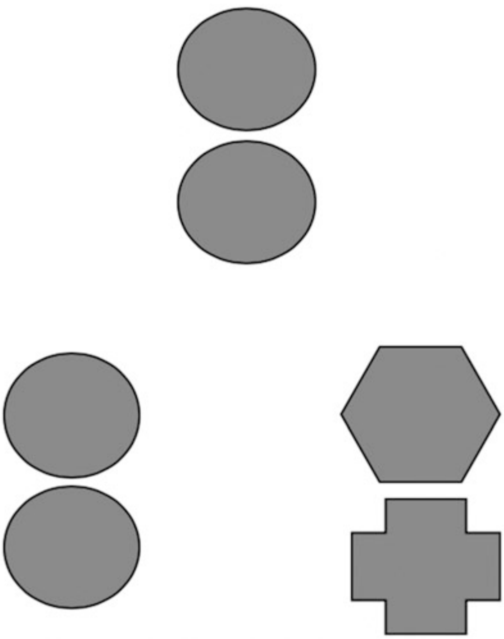
For example, a 3-year-old is unlikely to solve the analogy chicken:chick::dog:? spontaneously, despite knowing the word *puppy* and the parent-offspring relation. On the analogy task chicken:chick::duck:?, however, she is more likely to be successful. What differs between the two is that chicken and duck have a greater degree of surface similarity than do chicken and dog. Having noticed this surface similarity, children may *align* the chicken and the duck's features—their beaks, feathers, wings, and legs. Structure mapping posits that this alignment enables children to then notice the relational commonality so that all elements of the analogy can be organized in terms of their relational roles: chicken and duck as parents and chick and the missing element (duckling) as offspring. The learner can use the offspring relation to draw an inference because the surface similarity allowed her to perceive the relation in the first place.

Let us return to the classic RMTS task to illustrate the effectiveness of object matches in inviting comparison. In the original study, 2- and 3-year-olds failed to match standard AA to relational match BB; they chose randomly between choices BB and CD. (Note that this is the easier version of the task, with no competing object match choices such as AC.) To test if object matches can invite comparison and strengthen children's relational perception, we ran a study where 3-year-olds first saw a version of RMTS with a heightened surface similarity of the relational match: standard AA, choose between AA and CD (Fig. 11.3). Children saw four such trials and were then immediately tested on the original RMTS task (AA, choose BB or CD). That 3-year-olds performed close to 100% correct in the surface-similar RMTS will not be surprising; what is remarkable is their subsequent performance in the original, harder RMTS—they showed a preference for the relational match (Christie, 2010). Four trials of the surface-similar task were enough to shift the preschoolers' choices in the task without surface similarities.

The most likely explanation is that the surface-similar trials prompted the subjects to carry out a comparison, which they then applied in the subsequent trials without surface similarities, choosing the relational match as a consequence. Of course, since the children received four easier trials in the beginning as compared to the baseline RMTS study, one might attribute their success to a simple practice effect. To rule this out, we ran a control group of 3-year-olds who received four original RMTS trials in the beginning of the study. Consistent with our explanation, this group's choices were random.

This type of learning—building up from easy surface similarities to the more challenging relational commonalities—is called *progressive alignment* (Gentner & Medina, 1998; Kotovsky & Gentner, 1996; Loewenstein, Gentner, & Hung, 2007).

Fig. 11.3 Easy surface similarity (match in both object features and relations) helps children to focus on Relational Match to Sample Task (RMTS)



This powerful learning process is robust in children's day-to-day life because it capitalizes on an *existing* bias: that object matches are salient for even the youngest learners. The invitation to compare comes internally from the learner rather than being provided by external factors. Below, I review the evidence for progressive alignment across different ages, tasks, and domains.

Progressive Alignment

Many tests of infants' concepts implicitly rely on progressive alignment. In a typical habituation test, infants are presented with a number of events that represent some relation using identical or very similar objects. For example, in Quinn (1994) which tested infants' knowledge of the spatial relation *above,* all habituation events depicted a diamond-shaped object above a line. A priori, the relation *above* could be showcased using different objects above a line, say a diamond in habituation 1, a square in habituation 2, etc. But this is not the usual practice; habituation trials typically afford surface similarity—the events match in features as well as in relations. Carried out this way, habituation also serves as a learning opportunity: Infants notice a relational commonality among similar-looking events at habituation and are then able to observe it in trial runs.

Studies like Quinn (1994) are not by themselves direct evidence of progressive alignment. To my knowledge, no infant studies thus far have directly tested the importance of object similarity during habituation, contrasting similar-looking habituation events with more varied ones. But there is some indirect evidence of progressive alignment. For example, while 3-month-old infants in Quinn (1994) successfully discriminated above and below spatial categories when familiarized with dots, same age infants failed to form the categorization when they were familiarized to different shapes (i.e., a dot, an arrow, a triangle) (Quinn, Cummins, Kase, Martin, & Weisman, 1996). This suggests that the younger infants may in fact be *learning* the relation when presented with alignable, similar-looking habituation/familiarization events. In this view, habituation events that involve varied-looking stimuli result in a failure to align and to learn the relational commonality.

Direct evidence for learning by multiple exemplars via progressive alignment is abundant among preschoolers (e.g., Childers et al., 2015; Loewenstein et al., 2007). For example, in Loewenstein et al. (2007), 3- to 5-year-olds who learned using progressive alignment could successively learn a novel label referring to a part rather than to a whole object shape, a language learning task that is normally difficult at this age. Children were presented a series of novel characters accompanied by a novel label, for example, "this one has a *blick*." Their task was to map the novel word to the correct body part (navel) as opposed to other visible options such as eyes, nose, etc. Half of the children were presented with the characters in a progressive alignment way—from similar-looking to dissimilar—while the other half saw the same characters in random sequences. The progressive alignment group learned the name for the body part well: They correctly applied *blick* to the navel of a very different-looking character. The control group did not learn the word meaning.

Language Invites Comparison

There is good reason to suspect that language plays a prominent role in comparison learning. But how exactly? I discuss here several specific ways by which language can function as a cognitive tool for inviting alignment (see also Gentner & Christie, 2010; Gentner & Namy, 2006). Distinguishing them is important for making predictions on whether language will invite comparison in a given context. Indeed, although language is ubiquitous, comparison is not always easy to elicit.

Comparative words First and most obvious, many word meanings are comparative: *more, taller, best*. Hearing the word *taller* invites a learner to make a comparison among possible references; hearing *more* compels her to try various alignments—with another carrot, a longer playtime, a stronger emotion.

The resulting comparison is usually category-specific: "That tree is taller" will typically refer to another tree, not to a person. Of course, one could mean "that tree is taller than this man," but such non-alignable comparisons are disfavored in everyday language use. There is evidence that 4-year-olds already use comparative implicature within restricted categories. Barner and Snedeker (2008) showed children novel objects of varying heights and asked children to find the tall and short objects. The objects were labeled with a novel category name—"these are pimwits"—and children had to find the tall pimwit. Results suggested that 4-year-olds were sensitive to the height information within the comparison set, and they concurred on the tallest object among this set of objects. When a second kind of object was introduced (e.g., children now saw pimwits and tulvers), 4-year-olds restricted their tall interpretation within the set: The tall pimwit was the tallest among the pimwit objects, even though it was not the tallest object across both pimwit and tulver objects. Interestingly, category label can act as a determiner of the set: When both object kinds were called under the same name (e.g., these are all pimwits), children now shifted their tall interpretation to the tallest one across this comparison set.

We may ask if the observed preference for alignable comparisons over non-alignable ones is strengthened by the use of comparative words or whether this preference is independent of language use. Currently available evidence is not sufficient to decide this. Young children's strong bias for perceptual similarity suggests that some bias favoring alignable comparisons is in place from the start. Even if this is true, language may work to enhance such a bias.

Systematic language invites systematic structure Another mechanism by which language invites comparison arises from the systematicity afforded by language: A systematic description of a problem invites a correspondingly systematic conceptual representation. For example, Loewenstein and Gentner (2005) showed that young children could make use of language systematicity to infer structures in space. Three- and four-year-olds confronted with a spatial search task heard either a systematic set of terms *top-middle-bottom* or non-systematic terms *on-in-under* applied to boxes with three shelves (see Fig. 11.4). After seeing the experimenter hide a star in the Hide Box, the children were asked to find the star in the

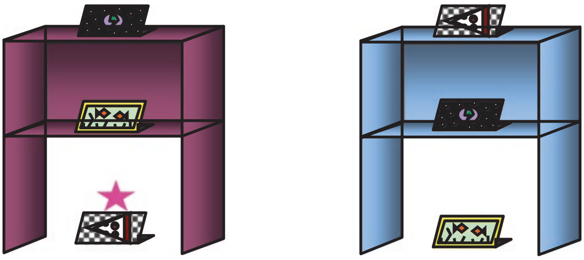


Fig. 11.4 Spatial mapping task used in Loewenstein and Gentner (2005). Children saw a star hidden behind one of the locations (e.g., bottom) at the Hiding Box and had to search in the corresponding location at the Finding Box. Because object match (the pizza picture at the bottom location of the Hiding Box matches with the pizza picture at the top location of the Finding Box) is pitted with relational mapping (bottom to bottom), 4-year-olds had difficulties solving this mapping task

corresponding spatial location in the Find Box. This hiding-finding task was difficult for children because the spatial alignment was pitted against object matches: The top location in the Hide Box *looked like* the middle location in the Find Box, etc. The prediction was that hearing the systematic terms *top-middle-bottom* would allow children to decode the spatial structure more easily than would hearing the non-systematic set *on-in-under*, which does not invite a hierarchical spatial representation. This prediction was borne out: The top-middle-bottom group outperformed the on-in-under group.

What is more, this mechanism can also work across domains. Gentner and Christie (2006) used the same setup as in Loewenstein and Gentner (2005), but this time the systematic set of terms was *one-two-three* while the non-systematic terms were *dog-pig-sheep*. The one-two-three group figured out the spatial structure and successfully solved the search task, while the dog-pig-sheep group persisted in making object match errors. The results held even among those children who had perfect memory for the names of the location: A child who remembered that *dog* referred to the top location still typically performed worse than her peers in the top-middle-bottom group. This is a telling example of how language systematicity fosters relational learning: Concrete animal names may help children to remember individual locations, but they do not accentuate spatial relations. Nonspatial albeit systematic terms can do just that. We see that successful mapping can occur even if the language domain (number) does not match the conceptual domain (spatial); a systematic set of terms invites cross-domain comparisons.

Common labels The third mechanism is seemingly mundane: *Any* label can invite alignment. The idea is that when two things or events are called by the same name, learners are invited to align them. This is a qualitatively different mechanism from the previous two because it relies on the mere appearance of a word, independent of its meaning. Learners take the cue to compare even when confronted with an unfamiliar label.

To show this mechanism at work, we again use the Relational Match to Sample Task (RMTS): match AA to BB, not CD. Recall that in the classic setup, 2- and 3-year-olds could not perceive the requisite relational similarity and choose randomly between BB and CD. In a novel label version of this task (Christie & Gentner, 2014), we marked AA with a novel label such as *truffet* and asked children which of the two options was also a truffet. Remarkably, without any training whatever, 2-year-olds declared the relational match BB to be a truffet. A novel label invited comparison.

The simplicity of this mechanism (it is independent of semantics!) means that it is available early—evidently at least as early as 2 years. The only prerequisite for this mechanism to effect comparison is that learners must have a prior assumption that like labels refer to like things. For this reason, I believe this labeling mechanism is one of the primary drivers of multiple exemplar learning in young children.

Social Relational Learning

Humans learn from and about relational concepts in almost every domain. To name a few examples, verbs and prepositions are relational components of language learning, numbers and arithmetic are intrinsically relational, spatial concepts—as experienced in daily life and confirmed by physicists (Einstein) and philosophers (Mach)—are all relative.

Likewise, the social domain is filled with relational concepts. To navigate the social world well, learners have to understand a plethora of relations—kinships, friendships, alliances, social hierarchies—because these relations govern behavior and have far-reaching consequences (friends help; foes do not). A learner needs to recognize multiple exemplars of a *social relation*: It is not enough to know that *mom* refers to my mother; I must also know that another woman is *mom* to my friend and not to me. The inherently relational nature of the social domain makes it an intriguing ground on which to inspect the problems and solutions for relational learning.

The Problem

Do learners really have difficulty recognizing multiple exemplars of social relations? We know this is the case in other domains of relational learning: Preschoolers do not spontaneously perceive identical relational matches (Christie & Gentner, 2014) or have difficulties learning verbs (e.g., Childers et al., 2015). In the social domain, however, the literature does not directly chart the difficulties of relational learning and, in fact, many recent studies suggest that humans learn social relations precociously. For example, 6-month-old infants infer a dominance relation based on group size (Pun, Birch, & Baron, 2016), while 5-month-olds prefer "helpers" over "hinderers" (Hamlin, Wynn, & Bloom, 2007). Already by 15 months, infants can organize perceived relations into a larger structure. For instance, they expect that two babies who are comforted by the same adult will affiliate with each other (Spokes & Spelke, 2017). On the other hand, social relations form such an intricate web that social learners still have plenty of opportunities to encounter the challenges and pitfalls of relational learning. For example, even 4-year-olds have difficulty differentiating siblings from friends; they were equally likely to indicate that siblings and friends have a grandmother in common (Spokes & Spelke, 2016). As another example that may be challenging even for the adult readers of this text, think about acquiring the concept *mother of firstborn child* that plays an important role in many polygamous cultures.

It is therefore useful to precisely characterize the problems of social relational learning. What is difficult to learn and why?

Problem 1: Lack of Domain Knowledge

Even if we accept that some social knowledge is innate (such as core knowledge in the social domain; see Spelke & Kinzler, 2007), most social relations will not be in place from the very beginning. Rather than cataloguing the social relations to be learned (kinship, friendship, schadenfreude, etc.), here I will focus on a few sample relations: relational sameness, goals, and false beliefs.

Relational sameness Two- and three-year-olds do not spontaneously select the relational match to the identity relation in the classic RMTS task, in which two inanimate, simple geometric shapes depict identity (Christie & Gentner, 2014). Does the perception of relational sameness change if identity is depicted by social entities such as faces? My lab has begun to answer this question: We are testing 2- and 3-year-olds on a social version of RMTS—match AA to either BB or CD—in which each letter represents a face. Preliminary results at the time of writing suggest that 2-year-olds are better in perceiving the relational match in this task than they were in its nonsocial predecessor. If the preliminary findings bear out, a new intriguing question will emerge: Can social relations provide a springboard for understanding relations in other domains? I will return to this question in section "Problem 2: Relational Versus Object Matches"

Goals Goals are relational and fit squarely within the social domain: They can only be attributed to social agents (see Woodward, Sommerville, Gerson, Henderson, & Buresh, 2009 for review). Infants as young as 6 months understand goals (Woodward, 1998), but this understanding develops over time. As an example, the same 6-month-old infants do not perceive the goal of an action that involves a tool, for example, using a clasp to reach for a toy; that understanding comes later at

12 months (e.g., Cannon & Woodward, 2012; Sommerville & Woodward, 2005). Here, comparison and structure mapping come in handy: When 7-month-olds were given a comparison of the action, they were able to understand the goal (Gerson & Woodward, 2012). In effect, comparison allowed these infants to "skip ahead" a few months in their understanding of goals.

False belief Comparison also benefits children's false belief understanding—a complex task, which is necessary to function in a social environment. Children younger than 4 years of age often fail in a standard false belief task (Wellman, Cross, & Watson, 2001; though see recent evidence of infants passing a simpler false belief test, e.g., Baillargeon, Scott, & He, 2010). Hoyos, Horton, and Gentner (2015) asked whether comparison can help young children to understand false beliefs earlier. They found that 3-year-olds who compared two false belief events and aligned them could then understand a false belief in a later test; their peers who saw the same events sequentially did not (see also Pham, Bonawitz, & Gopnik, 2012).

Problem 2: Relational Versus Object Matches

Even if children know a relation, they may not recognize its multiple exemplars because they find object matches more salient. A prototype of this situation is the Gentner and Christie (2006) study where preschoolers matched AA to AC rather than BB. A propensity to attend to objects is perhaps even more likely to occur in the social domain because children naturally find social entities salient, for example, preferring faces over other, equally complex objects. To my knowledge, no developmental study thus far has directly pitted relational versus object matches in the social domain. Doing so would be very useful for several reasons.

First, object matches seem to play a prominent role in social reasoning. As a heuristic example, when seeking information about China, many people would rather turn to a US-born Asian-American than, say, to an African-American professor of sinology. These types of choices are common in adults' social life; quite possibly, they are also faced frequently by young social learners. Characterizing social reasoning as a (possibly shifting and context-dependent) preference for relational versus object matches can potentially transform our understanding of many social learning processes.

Second, social categorization phenomena can often be understood in terms of a tension between object and relational matching. Five- to six-year-old children take the most minimal cues (such as t-shirt color) as signaling in-group membership (Dunham, Baron, & Carey, 2011; Dunham & Emory, 2014; Spielman, 2000). But if we temporarily set aside our knowledge of social status, we will discover that seemingly fundamental cues in adult social grouping—gender, race, nose shape in Rwanda—are, in essence, object-matching cues. Viewing social categorization as an arena in which object matches and relational matches vie for our attention and preference opens up a large area for future studies—to explore how social groups and their perception originate from relational learning.

One example of such a study is Kurzban, Tooby, and Cosmides (2001), who showed that in adults, a particular relational similarity—coalition—can trump social categorization based on race (object similarity). Participants were given a Memory Confusion Protocol, where they were exposed to a series of photographs and sentences (each pictured individual said a number of sentences). After this exposure, participants had to recall who said what. The premise is that one is more likely to make within-category error than between-category error. For example, a soccer fan is more likely to confuse Messi and Neymar (a famous duo that played for Barcelona) than the archrivals Messi and Ronaldo (playing for two different clubs). The results of the experiment showed that although both cues for social categorization were available (coalitional vs. race), subjects made more errors based on the coalitional cues. Kurzban, Tooby, and Cosmides see this result as supporting their theory that coalitional cues form the evolutionary basis for how humans carve social groupings in their representations. If this is true, we should expect young children to *start out* with noticing coalitional cues—relational cues—in their learning of social groups rather than progressing toward it from initially noticing surface cues such as race. But, this is not what happens in other domains, where young children's preference shifts in time from object matches to relational matches. Is the direction reversed in the social domain? So far, no developmental works have directly pitted race (object match) versus coalition (relational match).

My lab has begun to look at the tension between relational and object matches in the social domain. As an initial test, we are using a social variant of the RMTS task: Match the pair of faces AA to faces BB or faces AC. Unlike the classic RMTS task (where 4-year-olds overwhelmingly choose object matches over relational ones (Christie & Gentner, 2010), here both outcomes are possible. Four-year-olds may prefer object matches because two identical faces are just so compelling, but the social nature of the faces may also work to make the identity relation more obvious (think twins).

The latter outcome would mean a fundamental change in our understanding of relational reasoning—that children initially favor object matches and only later shift to relational matches (Gentner, 1988; Halford, 1992). Instead, we may need to posit a new account where attention to relations is initially limited to the social domain and only later colonizes other domains. In this view, children could use social relational concepts to learn analogous relations in other domains. At present, this is a hypothesis; we do not yet have evidence to substantiate this view.

The Solution: How Does Structure Mapping Work in the Social Domain?

If understanding the social world is a relational learning problem, then structure mapping ought to provide a solution. I have discussed two developmental studies in which structural alignment and comparison helped young children to perceive relational commonalities. In the first one, Gerson and Woodward (2012) showed that 6-month-olds could learn about goals only when they were exposed to comparison learning. In the second study, Hoyos et al. (2015) showed that 4-year-olds who initially had not passed a false belief task succeeded in a later test only when they compared false belief scenarios in the interim. More studies of this kind are needed to give a complete account of structure mapping in the social domain.

The following sections discuss the next set of questions, which are relevant for social relational learning.

Language for Learning Relational Concepts

What is the role of language in learning social relational concepts? Children's earliest vocabulary contains many social relational concepts, particularly relating to kinship: for example, *mother, brother, uncle, and grandma*. Do children possess these relational concepts prior to learning their labels, or does labeling the concepts help children to learn their relational meaning? For example, do children know that "mother" is a relational concept that can be applied in widely different parent-offspring contexts—to *my* mother, a friend's mother, or Mother Goose? One classic study by Keil (1989) suggests that this is not the case. Preschoolers thought that *uncle* referred to a man wearing a hat and smoking a pipe rather than one's father's brother—a clear example of the tension between object and relational matches. To the extent that this (mis)understanding is widespread for kinship terms, their acquisition is an excellent ground for studying how language promotes learning relational terms. The template from nonsocial domains suggests this mechanism: Common labels invite comparison, which fosters relational abstraction (Christie & Gentner, 2014). On day 1, a child may hear her friend say "that's my uncle," and on day 2, she may hear Mom say, "give your uncle a hug." The common label "uncle" may prompt the child to compare the two events. Despite the differences in perceptual features between days 1 and 2—the uncles in question may well be different men—comparison can highlight the common relational structure.

Does Object Similarity Help?

Recall that children can learn from progressive alignment: advancing from comparing highly similar exemplars to more distant ones. For example, children learned the concept *four* more easily when they compared two quadruples of similar objects (e.g., four red balls and four pink balls) than when they compared sets of four dissimilar objects (four red balls and four cats) (Mix, 1999, 2008). Assuming progressive alignment also applies in the social domain, we should expect children to recognize social relational concepts earlier when they are instantiated by more similar-looking people or groups. For example, do children apply the concepts *parent* or *enemy* more easily to people of the same age, gender, or race?

Social Comparison and Object Versus Relational Similarity

Comparison is ubiquitous in the social domain; people constantly compare themselves to others, as well as comparing others to one another (social comparison; Festinger, 1954). Structure mapping theory (SMT) offers an account of this process.

First, SMT predicts that the earliest mappings happen between closely similar exemplars—social entities that share featural similarities, for example, people of the same gender, race, and age. If young children are indeed likely to restrict their initial social comparison to superficially similar people or groups, it may have a far-reaching impact on their everyday social reasoning. For example, a child who restricts her initial social comparisons to children of the same race will not derive the benefit of comparison—noticing deeper relational commonalities—in her social learning about other races. Without intervention from the outside, a child may lag behind in recognizing—or, indeed, never recognize—the full scope of what she and children of other races have in common.

Second, structure mapping shifts learners' perception from surface similarities to favoring relational commonalities. Thus, advanced learners are expected to engage in comparison more easily when the subjects of comparison share a meaningful relation rather than a simple surface feature. There is evidence for this among adults. Mussweiler and Gentner (2007) asked adult participants to compare a target character to either a surface-similar standard or to a relationally similar standard. For example, if Bob (the target) was a sophomore highly dedicated to sports, the feature-match Adam would share surface characteristics (same gender, disinterest in cultural events, athletic build) while the relational match Melissa would combine deeper commonalities with Adam (a high dedication to some pursuit such as music) with superficial differences (different gender, interests, hobbies). Adults preferred to compare the target to the relationally similar standard.

We do not yet know whether this finding holds across ages, contexts, and cultures. It is quite possible that young children will show the opposite pattern, favoring social comparisons involving object-matching standards over relational ones. Adults may also reverse their choices when conducting comparison under time pressure: In casual, quick social interactions, they may only get through stage 1 of the structure mapping process—mapping object matches. An extra layer of complexity is that cultural context affects which object and relational similarities appear salient in the comparers' representation. For example, in gender-segregated cultures, a female teacher target may be more likely compared to another female than to another teacher. Indeed, in a society where gender sets strict boundaries on one's relational role (e.g., designating a woman as the primary caregiver, mother, and wife), perhaps we should consider gender as a relational rather than a feature cue. This example illustrates that in the social domain, boundaries between relational and object matches can be more fluid and debatable than in other domains. Much exciting research waits to be done in order to understand how we use surface and relational matches in our social comparisons. What commonalities and what differences occur between predominant social comparison styles in different cultures?

Alignable Differences

Structure mapping theory predicts that it is easier to find meaningful differences between alignable events than between non-alignable ones. For example, it is easier to list the differences between hotel H and motel M than between hotel H and cat C. This is because the characteristics of a hotel (price, parking lot, reception) are amenable to being mapped onto the characteristics of a motel; we call such features alignable dimensions. The dimensions of a hotel and a cat, however, do not align. Sagi, Gentner, and Lovett (2012) have shown that adults declare two events to be different faster when they are non-alignable but can list more differences between two events when they are alignable. Likewise, preschoolers could find more meaningful differences between alignable items such as forks and spoons than they could between the non-alignable forks and cats (Gelman, Raman, & Gentner, 2009). Greater similarity affords a more meaningful contrast.

Alignable differences can potentially explain many phenomena in the social domain because its main entities—people—are by nature alignable. Moreover, they tend to be *more* alignable than exemplars in other domains so that differences between people are easier to discern than differences between other entities. We found evidence for this when we contrasted inductive generalization reasoning in the social versus the biological domain (Noyes & Christie, 2016). We asked 5-year-olds to find out whether *all* children like toy Y by inquiring either with two girls (a narrow sample) or with a boy and a girl (a broad sample). The broader sample (boy-girl) provides the larger scope for generalization (the inductive principle; Osherson, Smith, Wilkie, Lopez, & Shafir, 1990), and our 5-year-old subjects opted for this correct choice. However, when the same question was asked about animals (narrow: a zebra and a horse; broad: a zebra and a mouse), they chose at chance. Children's inability to use the broader sample to generalize in the animal kingdom was consistent with prior results (Gutheil & Gelman, 1997; Li, Cao, Li, Li, & Deak, 2009; Rhodes & Brickman, 2010), but the finding that they use the broader sample when reasoning about people was novel—the first demonstration that 5-year-olds are capable of applying the inductive principle. We believe that the inductive principle was activated in the social domain because the latter enabled them to rank the variabilities of the samples—to perceive that a boy and a girl really are more diverse than a pair of girls. When reasoning about animals, children likely perceived the zebra-horse pair to be just as diverse as the zebra-mouse pair.

Another potential outcome of using alignable differences in our reasoning is the out-group homogeneity effect: that people perceive members of the out-group to be less different from one another than members of the in-group (Messick & Mackie, 1989; Park & Rothbart, 1982; Quattrone, 1986). A likely explanation is that people within one's in-group are more alignable (at least to oneself) so that differences among them are easier to perceive. On an anecdotal level, similar cultures and countries tend to see greater divisions or conflicts than countries that are more distinct, and people's dislikes or rivalries run stronger when the rivals have more in common. The big question is whether and when young learners start using alignable difference-based reasoning in their calculations about social relations. It is possible that with every multiple example that they encounter, the alignable exemplars give rise to understanding many social relational concepts involving differences.

Summary

Noticing relations is hard. Given multiple exemplars, learners can overcome this difficulty by comparing the surface features of exemplars so as to distill their relational commonality. This relational abstraction allows learners to subsequently recognize multiple exemplars of the same relations: from recognizing relational nouns, verbs, spatial relations, and prepositions to potentially understanding complex relations in the social domain. Structure Mapping Theory (Gentner, 1983) provides core explanations for the learning process: Learners who compared multiple exemplars will favor relational commonalities more than featural commonalities; to compare means to *align* the exemplars, not just merely have two (or more) multiple exemplars.

But learners do not always align multiple exemplars as evidenced by young children having difficulties recognizing exemplars of relational concepts. Even the basic relation such as identity poses a problem to young children (Christie & Gentner, 2010, 2014). Explicit comparison given by "teachers" (adults, caregivers, experimenters) can foster learners' relational thinking. But other learning tools, which are readily available in the learner's everyday environment, can also invite alignment: (i) object matches or overall similarity and (ii) common labels. The critical question for further research is what prompts learners to spontaneously make use of these learning tools. Why do some people align and some do not?

Learning about the social world requires learning about multiple exemplars of social relations: Who counts as friends, foes, kins, and superiors? Even if children are precocious social learners, they still have to learn many relational concepts in the social world. Given this same problem of relational learning, it may be fruitful to parameterize the problems and solutions from the structure mapping perspective. For example, learners may also face the tension between object and relational matches in choosing standards for social comparison. At the same time, aligning and comparing people may result in learners noticing relational commonalities—such as same roles (e.g., we're both parents), favoring such commonality over mere appearances (e.g., we're both white). Much research is to be done to understand social learning as relational learning.

Acknowledgment I thank Bartłomiej Czech and Jane Childers for suggestions and insights in writing this chapter. Thanks to Shuai Shao and Yang Qisen for assistance in data collection and figure drawing.

References

Baillargeon, R., Scott, R. M., & He, Z. (2010). False-belief understanding in infants. *Trends in Cognitive Sciences, 14*(3), 110–118. https://doi.org/10.1016/j.tics.2009.12.006

Barner, D., & Snedeker, J. (2008). Compositionality and statistics in adjective acquisition: 4-year-olds interpret tall and short based on the size distributions of novel noun referents. *Child Development, 79*(3), 594–608.

Cannon, E. N., & Woodward, A. L. (2012). Infants generate goal-based action predictions. *Developmental Science, 15*(2), 292–298. https://doi.org/10.1111/j.1467-7687.2011.01127.x

Casasola, M. (2005). Can language do the driving? The effect of linguistic input on infants' categorization of support spatial relations. *Developmental Psychology, 41*, 183–192.

Childers, J. B., Parrish, R., Olson, C. V., Burch, C., Fung, G., & McIntyre, K. P. (2015). Early verb learning: How do children learn how to compare events? *Journal of Cognition and Development, 17*(1), 41–66. https://doi.org/10.1080/15248372.2015.1042580

Christie, S. (2010). *How simple is same: Perception of relational similarity in young children* (Doctoral dissertation). Retrieved from ProQuest Dissertations Publishing.

Christie, S., & Gentner, D. (2007). Relational similarity in identity relation: The role of language. In S. Vosniadou & D. Kayser (Eds.), *Proceedings of the Second European Cognitive Science Conference*.

Christie, S., & Gentner, D. (2010). Where hypotheses come from: Learning new relations by structural alignment. *Journal of Cognition and Development, 11*(3), 356–373. https://doi.org/10.1080/15248371003700015

Christie, S., & Gentner, D. (2014). Language helps children succeed on a classic analogy task. *Cognitive Science, 38*(2), 383–397.

Dunham, Y., Baron, A. S., & Carey, S. (2011). Consequences of "minimal" group affiliations in children. *Child Development, 82*(3), 793–811.

Dunham, Y., & Emory, J. (2014). Of affect and ambiguity: The emergence of preference for arbitrary ingroups. *Journal of Social Issues, 70*(1), 81–98.

Ferry, A. L., Hespos, S. J., & Gentner, D. (2015). Prelinguistic relational concepts: Investigating analogical processing in infants. *Child Development, 86*, 1386–1405.

Festinger, L. (1954). A theory of social comparison processes. *Human Relations, 7*(2), 117–140. https://doi.org/10.1177/001872675400700202

Gelman, S. A., Raman, L., & Gentner, D. (2009). Effects of language and similarity on comparison processing. *Language Learning and Development, 5*(3), 147–171. https://doi.org/10.1080/15475440902824079

Gentner, D. (1983). Structure-mapping: A theoretical framework for analogy. *Cognitive Science: A Multidisciplinary Journal, 7*(2), 155–170.

Gentner, D. (1988). Metaphor as structure mapping: The relational shift. *Child Development, 59*, 47–59.

Gentner, D., & Christie, S. (2006). Using numbers to structure space. In R. Sun & N. Miyake (Eds.), *Proceedings of the twenty-eight annual meeting of the Cognitive Science Society.*

Gentner, D., & Christie, S. (2010). Mutual bootstrapping between language and analogical processing. *Language and Cognition, 2*(2), 261–283.

Gentner, D., & Medina, J. (1998). Similarity and the development of rules. *Cognition, 65*(2), 263–297.

Gentner, D., & Namy, L. L. (1999). Comparison in the development of categories. *Cognitive Development, 14*, 487–513.

Gentner, D., & Namy, L. L. (2006). Analogical processes in language learning. *Current Directions in Psychological Science, 15*, 297–301.

Gentner, D., & Rattermann, M. J. (1991). Language and the career of similarity. In S. A. Gelman & J. P. Byrnes (Eds.), *Perspectives on thought and language: Interrelations in development*. London: Cambridge University Press.

Gerson, S. A., & Woodward, A. L. (2012). A claw is like my hand: Comparison supports goal analysis in infants. *Cognition, 122*(2), 181–192. https://doi.org/10.1016/j.cognition.2011.10.014

Gervain, J., Berent, I., & Werker, J. F. (2012). Binding at birth: The newborn brain detects identity relations and sequential position in speech. *Journal of Cognitive Neuroscience, 24*(3), 564–574. https://doi.org/10.1162/jocn_a_00157

Gomez, R., & Gerken, L. (2000). Infant artificial language learning and acquisition learning. *Trends in Cognitive Science, 4*, 178–186.

Gomez, R. L. (2002). Variability and detection of invariant structure. *Psychological Science, 13*(5), 431–436. https://doi.org/10.1111/1467-9280.00476

Gutheil, G., & Gelman, S. A. (1997). Children's use of sample size and diversity information within basic-level categories. *Journal of Experimental Child Psychology, 64*(2), 159–174.

Halford, G. S. (1987). A structure-mapping approach to cognitive development. *International Journal of Psychology, 22*(5–6), 609–642.

Halford, G. S. (1992). Analogical reasoning and conceptual complexity in cognitive development. *Human Development, 35*(4), 193–217.

Hamlin, J. K., Wynn, K., & Bloom, P. (2007). Social evaluation by preverbal infants. *Nature, 450*, 557–559.

Hoyos, C., Horton, W. S., & Gentner, D. (2015). Analogical comparison aids false belief understanding in preschoolers. In D. C. Noelle, R. Dale, A. S. Warlaumont, J. Yoshimi, T. Matlock, C. D. Jennings, & P. P. Maglio (Eds.), *Proceedings of the 37th annual conference of the Cognitive Science Society* (pp. 1–6). Austin, TX: Cognitive Science Society.

Jee, B., Gentner, D., Uttal, D., Sageman, B., Forbus, K., Manduca, C., et al. (2014). Drawing on experience: How domain knowledge is reflected in sketches of scientific structures and processes. *Research in Science Education, 44*(6), 859–883. https://doi.org/10.1007/s11165-014-9405-2

Keil, F. C. (1989). *Concepts, kinds, and conceptual development*. Cambridge, MA: MIT Press. https://doi.org/10.2307/415569

Keil, F. C. (1998). Cognitive science and the origins of thought and knowledge. In R. M. Lerner (Ed.), *Theoretical models of human development. Handbook of child psychology* (5th ed.), Editor-in-Chief: William Damon. New York: Wiley.

Kotovsky, L., & Gentner, D. (1996). Comparison and categorization in the development of relational similarity. *Child Development, 67*, 2797–2822.

Kurzban, R., Tooby, J., & Cosmides, L. (2001). Can race be erased? Coalitional computation and social categorization. *Proceedings of the National Academy of Sciences, 98*, 15387–15392.

Li, F., Cao, B., Li, Y., Li, H., & Deak, G. (2009). The law of large numbers in children's diversity-based reasoning. *Thinking & Reasoning, 15*(4), 388–404.

Loewenstein, J., & Gentner, D. (2005). Relational language and the development of relational mapping. *Cognitive Psychology, 50*(4), 315–353.

Loewenstein, J., Gentner, D., & Hung, B. (2007). Comparison facilitates children's learning of names for parts. *Journal of Cognition and Development, 8*(3), 285–307.

Loewenstein, J., Thompson, L., & Gentner, D. (1999). Analogical encoding facilitates knowledge transfer in negotiation. *Psychonomic Bulletin & Review, 6*(4), 586–597.

Marcus, G. F., Vijayan, S., Rao, S. B., & Vishton, P. M. (1999). Rule learning by seven-month-old infants. *Science, 283*(5398), 77.

Markman, A. B., & Gentner, D. (1993). Splitting the differences: A structural alignment view of similarity. *Journal of Memory and Language, 32*, 517–535.

Messick, D. M., & Mackie, D. M. (1989). Intergroup relations. *Annual Review of Psychology, 40*(1), 45–81.

Mix, K. S. (1999). Similarity and numerical equivalence: Appearances count. *Cognitive Development, 14*(2), 269–297.

Mix, K. S. (2008). Surface similarity and label knowledge impact early numerical comparisons. *British Journal of Developmental Psychology, 26*(1), 13–32.

Mussweiler, T., & Gentner, D. (2007). On apples and oranges: Structural alignment in the selection of social comparison standards. *Journal of Cognitive Science, 8*(1), 1–38.

Namy, L. L., & Gentner, D. (2002). Making a silk purse out of two sow's ears: Young children's use of comparison in category learning. *Journal of Experimental Psychology: General, 131*, 5–15.

Namy, L. L., Gentner, D., & Clepper, L. (2007). How close is too close? Alignment and perceptual similarity in children's categorization. *Cognition, Brain, Behavior, 11*(4), 647–659.

Noyes, A., & Christie, S. (2016). Children prefer diverse samples for inductive reasoning in the social domain. *Child Development, 87*(4), 1090–1098.

Oakes, L. M., & Ribar, R. J. (2005). A comparison of infants' categorization in paired and successive presentation familiarization tasks. *Infancy, 7*, 85–98.

Osherson, D. N., Smith, E. E., Wilkie, O., Lopez, A., & Shafir, E. (1990). Category-based induction. *Psychological Review, 97*(2), 185–200.

Park, B., & Rothbart, M. (1982). Perception of out-group homogeneity and levels of social categorization: Memory for the subordinate attributes of in-group and out-group members. *Journal of Personality and Social Psychology, 42*(6), 1051.

Pham, K., Bonawitz, E., & Gopnik, A. (2012). Seeing who sees: Contrastive access helps children reason about other minds. In *Proceedings of the 34th annual conference of the Cognitive Science Society* (pp. 2180–2185).

Pun, A., Birch, S. A., & Baron, A. S. (2016). Infants use relative numerical group size to infer social dominance. *Proceedings of the National Academy of Science, 113*(9), 2376–2381.

Quattrone, G. A. (1986). On the perception of a group's variability. *Psychology of intergroup relations, 2*, 25–48.

Quinn, P. C. (1994). The categorization of above and below spatial relations by young infants. *Child Development, 65*(1), 58–69.

Quinn, P. C., Cummins, M., Kase, J., Martin, E., & Weisman, S. (1996). Development of categorical representations for above and below spatial relations in 3- to 7-month-old infants. *Developmental Psychology, 32*, 942–950.

Rhodes, M., & Brickman, D. (2010). The role of within-category variability in category-based induction: A developmental study. *Cognitive Science, 34*(8), 1561–1573.

Sagi, E., Gentner, D., & Lovett, A. (2012). What difference reveals about similarity. *Cognitive Science, 36*(6), 1019–1050. https://doi.org/10.1111/j.1551-6709.2012.01250.x

Sandoval, M., & Gómez, R. L. (2013). The development of nonadjacent dependency learning in natural and artificial languages. *WIREs Cognitive Science, 4*, 511–522.

Sommerville, J. A., & Woodward, A. L. (2005). Pulling out the intentional structure of action: The relation between action processing and action production in infancy. *Cognition, 95*(1), 1–30.

Spelke, E. S., & Kinzler, K. D. (2007). Core knowledge. *Developmental Science, 10*(1), 89–96.

Spielman, D. A. (2000). Young children, minimal groups, and dichotomous categorization. *Personality and Social Psychology Bulletin, 26*(11), 1433–1441.

Spokes, A. C., & Spelke, E. S. (2016). Children's expectations and understanding of kinship as a social category. *Frontiers in Psychology, 7*, 440.

Spokes, A. C., & Spelke, E. S. (2017). The cradle of social knowledge: Infants' reasoning about caregiving and affiliation. *Cognition, 159*, 102–116.

Waxman, S. R., & Hall, D. G. (1993). The development of a linkage between count nouns and object categories: Evidence from fifteen-to twenty-one-month-old infants. *Child Development, 64*(4), 1224–1241. https://doi.org/10.2307/1131336

Wellman, H. M., Cross, D., & Watson, J. (2001). Meta-analysis of theory-of-mind development: The truth about false belief. *Child Development, 72*(3), 655–684. https://doi.org/10.1111/1467-8624.00304

Woodward, A. L. (1998). Infants selectively encode the goal object of an actor's reach. *Cognition, 69*(1), 1–34. https://doi.org/10.1016/s0010-0277(98)00058-4

Woodward, A. L., Sommerville, J. A., Gerson, S., Henderson, A. M., & Buresh, J. (2009). The emergence of intention attribution in infancy. *Psychology of Learning and Motivation, 51*, 187–222. https://doi.org/10.1016/s0079-7421(09).51006-7

Chapter 12
Epilogue: Comparing Comparison Theories: What Can We Gain?

Jane B. Childers

Abstract This book brings together researchers who use different theories to test infant and children's learning about concepts and language. To do this, chapters were invited from researchers who focus on how infants (see Chaps. 2, 3, 4, 5, and 6) and children learn from multiple examples (see Chaps. 7, 8, 9, 10, and 11). A concerted effort was made to include people who relied on different theories to explain unsupervised learning across stimuli to examine how these theories structured their studies, which mechanisms were posited in each theory, and what the evidence was for each theory in a domain. This final chapter discusses five main themes that can be drawn across these chapters including the following: (1) the ability to use information across exemplars can be seen early in development, (2) comparison abilities develop, (3) there are underlying mechanisms in each view that can be compared, (4) do particular procedures keep researchers from considering alternate theories?, and (5) what future directions are suggested? In conclusion, the most amazing part of learning from multiple examples is that it is powerful! By collecting ideas from a variety of thinkers who have tackled similar questions using different tools, we hope new insights will emerge for individual researchers and for the field.

The goal of this book was to bring together in one place researchers who use different theories to test infant and children's learning about objects and words. To do this, chapters were invited from researchers who focus on how infants (see Chaps. 2, 3, 4, 5, and 6) and children learn from multiple examples (see Chaps. 7, 8, 9, 10, and 11). A concerted effort was made to include people who relied on different theories to explain unsupervised learning across stimuli to examine how these theories structured their studies, which mechanisms were posited in each theory, and what the evidence was for each theory in a domain. (Recall that not all traditions are represented, but many key theories are present in this book.) A central idea in each of these theories is that examining evidence across examples aids in learning, leading to a greater and deeper understanding of the phenomena at hand. To this end, the

J. B. Childers (✉)
Department of Psychology, Trinity University, San Antonio, TX, USA
e-mail: jchilder@trinity.edu

J. B. Childers (ed.), *Language and Concept Acquisition from Infancy Through Childhood*, https://doi.org/10.1007/978-3-030-35594-4_12

goal was to create a single book in which readers could actively compare theories and evidence from one chapter to the next. I hope your ability to compare across these chapters has led to new insights—as predicted by these theories! I believe this process has been productive and, in the next section, I will consider what can be learned by reflecting on some key themes across these chapters.

Theme 1: The ability to use information across examples can be seen early in development

Although timing varies some across the infant chapters, all describe evidence that infants benefit from comparing multiple exemplars within the first year of life across several different domains. Specifically, Johnson shows newborns can process multiple examples of (simple) stimuli that is visual or auditory, and Theissen adds that infants can use statistical learning for attending to VOT and acquiring phonemes early. Hespos et al. describe processing of a relation across examples by 3 months, and Casasola and Park describe infants' understanding of spatial relations emerging within the first year. Sobel et al. also show that as early as 5 months, infants can compare across trials to attend to co-occurrences between faces or dynamic arrows and locations. Additionally, Graham et al. show that 9- and 11-month-olds can learn and generalize properties of object categories, but that they are influenced by the familiarity of the category, their age, and by whether they see a single or multiple exemplars (multiple exemplars promote learning). Christie reminds us that infants perceive social relations in the first 12–15 months but that fully understanding social categories (e.g., the category "mom") can be difficult. Thus, the infant chapters show an early ability to compare across multiple exemplars can aid in learning speech sounds, objects and categories, and words and relations. They also demonstrate that these studies can be motivated by, and explained, using different theories of comparison.

Theme 2: Comparison abilities develop

In the infant chapters, development is described either later in the first year or between the first and second year, and in the chapters with older children, development is described across the preschool years/early elementary ages. A key theme is that later in development, children can process more stimuli, more complex stimuli, and more varied stimuli in a domain than is demonstrated at younger ages.

Specifically, in Johnson's chapter, infants can process sets of more complex stimuli (e.g., visual sequences of shapes) using statistical learning, increasing in their ability over the first year. In Casasola and Park, infants' ability to compare diverse exemplars increases such that early in development, their learning is more fragile and context bound than it is later in development. Both Casasola and Park, and Theissen, show the perceptual characteristics of stimuli and number of exemplars, as well as variety across exemplars, affects infants' learning. Hespos et al. note that specific object familiarity interferes with (relational) processing at younger ages, with less interference over time. And in Graham et al., at 9 and 11 months, infants have difficulty learning about unfamiliar animals shown in dynamic events if they are only shown a single example. There is a common thread across these accounts and across domains with younger infants needing simpler stimuli that are less varied, and perhaps repeated, to learn from a set of examples and older infants or children being able to process more complex varied stimuli. Whether this is due to greater cognitive abilities, specific experiences, or both is unclear. Interestingly, this pattern appears across theories, ages, and domains.

What about in the chapters including children older than 2 years? Sandhofer and Schonberg show that younger or less experienced learners learn words better when there is a common context across examples or between the instances present during learning and new extension objects; older children are less reliant on context. Younger children also benefit from multiple, correlated cues when aggregating (or comparing) instances, whereas older children do not need this support as much. Imai and Childers show that object similarity can influence 3-year-old children's ability to extend a new verb to a new event if given only a single example before test but that experience with high-similarity comparisons can help children learn how to compare. Lapidow and Walker use infants' and young children's early ability to do exploratory causal learning as the basis for a later "Search for Invariance" hypothesis that is based on producing multiple examples of a phenomenon that support a child's current hypothesis. In Christie, as children develop the ability to compare relations across examples, including social relations, they can further develop their social knowledge (e.g., their false belief understanding) and understanding of social categories. Predicting from structure mapping theory to this new domain of social categories, what should help children compare is exposure to new relational words, experience with examples with high similarity, and a focus on alignable differences. Thus, in this chapter, focusing on older children, early abilities underlie later contexts in which children can compare.

Theme 3: There are underlying assumptions or mechanisms that are inherent in each theory. (Are they different from each other?)

The chapters in this book describe how a specific theory has framed research for a specific problem in development and specific age range. Here, I explore the mechanisms that have been described in order to compare these mechanisms to each other.

Recall that the key mechanism in statistical learning is attention to co-occurrences across stimuli. Now, consider two main points from Erik Thiessen. One is that we can contrast conditional probability (from one stimuli to the next) with distributional probability (processing across the set; cross-situational processing). This is an important contrast that perhaps could be extended in other areas. What would conditional probability look like if applied to verb learning, or spatial processing, as compared to distributional probability processing? As a verb researcher, conditional probabilities could be computed within a single event, while distributional probabilities could be computed across the set of events linked to a verb, but (to our knowledge) no verb study has directly compared these types of tasks/processing. In terms of mechanisms, Theissen argues that memory structures in the brain and the mind are critical to comparison, which is an important assertion. Often, infants and children are comparing stimuli they are experiencing with stimuli from the past, and thus memory is a key part of the task of comparing, and memory stores and abilities could be key mechanisms in either statistical learning or structure mapping processing. Scott Johnson poses a different question which is whether infants are storing transitional probabilities (TPs) across examples or are storing chunks (sets of examples). He proposes that perhaps they are storing TPs on the way to forming chunks. A similar notion is described in Sobel et al. in that infants and toddlers may be forming units across co-occurrences of speaker + label + object, which are then judged to be reliable or unreliable.

Sandhofer and Schonberg also link memory to statistical learning in older children. They remind us that infants and children hear many repetitions of a word before learning it. That what helps in memory, and in aggregating examples or comparison, is having similar contexts and correlated cues. Also forgetting and being reminded of a word helps children learn a word, as is found in memory experiments. This is supported by evidence from their lab showing that spaced practice produces the best performance when there is a delayed (15 minutes later) test (Vlach, Ankowski, & Sandhofer, 2012). They also note that there is a tension between the similarity of examples and variation across a set of examples, which is an important point across chapters. Multiple researchers note that similarity leads to easier but more restricted learning, whereas varied examples are more difficult to aggregate but lead to greater generalizations. Lapidow and Walker argue that a phenomenon once seen as a cognitive error, only testing or creating tests of a causal relation that support a current hypothesis, actually does help learners because these tests give the learner multiple examples of contexts in which a particular causal relation holds. This also helps learners discern how to generalize a new causal relation.

In structure mapping, infants and children who are comparing across examples can begin by finding all possible matches (see Gentner, 2010), and then build up these matches into an alignment of specific elements across the examples based on their common relational structure. Thus, both statistical learning and structure mapping have a bottom-up component to the processing, but in structure mapping, the mechanism then leads to the consideration of structure. Most chapters in this volume only consider one theory or the other, but see Casasola and Park for a consideration of how different theories could apply to their studies of spatial learning. Hespos et al. argue for a mechanism for relational processing that is unique to human infants, not seen in other animals. Would statistical learning produce a pattern of results confined to human infants?

In sum, both statistical learning and structure mapping have bottom-up processing, both lead to the (possible) creation of larger units of information (either chunks or units, or representations with a structure), and both can include higher-order information either because this information is also considered (e.g., rational real-world knowledge; see Sobel et al.) or because higher-order insights are promoted.

Theme 4: Do the procedures used in a topic area keep researchers from considering alternative theories (or vice-versa)?

When I initially envisioned this book, it seemed possible that researchers could be constrained by the procedures they use, only using specific procedures to test a particular theory. For example, in object permanence, using the violation of expectation procedure (e.g., Baillargeon's work) leads to different conclusions than does the direct enactment procedure used by Piaget. To explore this question, I asked researchers to describe the procedures commonly used in their domain and ask whether other procedures would lead to different results.

In the infant chapters, researchers are using the same set of procedures regardless of theory, which is a strength. For example, Johnson discusses studies using habituation and so does Casasola and Park, Theissen, and Hespos and collaborators. At times, additional procedures are mentioned, but none seem critical for shaping one outcome versus another. For example, Johnson mentions studies of preferential looking, eye tracking, fMRI, and ERP, but these could be used in Casasola and Park or in Theissen's studies. Hespos et al. add animal studies and computer simulations, and Graham et al. describe studies in which toddlers manipulate objects at test, but these procedures also could be applied in other chapters.

In the chapters with older children, again, researchers appear to be using common procedures across theories. For example, Sandhofer and Schonberg describe word learning tasks, often focusing on forced-choice tasks. These types of tasks are also present in Imai and Childers' chapter which uses a different theory and word type (verbs). Lapidow and Walker describe studies of children's scientific reasoning using tasks commonly employed in that domain while Christie describes forced-choice tasks that can be used across domains.

Thus, there does not seem to be a procedural choice that is dividing researchers, which is important. This should make the comparison of results across studies with different theoretical foundations easier, as similar types of data should be available from different theoretical camps. It is thus interesting that we are not doing more to contrast theories within a single paper or area of study or examining whether different theoretical mechanisms could be integrated with each other.

Theme 5: What do these chapters tell us about future directions?

One future direction that stands out is how to *distinguish* statistical learning from structural alignment in terms of the *results* of studies or the data. This is a difficult question to answer, and I feel emboldened in stating that this is a difficult question by again noting that Casasola and Park could apply both theories to their data on infant spatial reasoning. On the one hand, Hespos et al., Christie, and Imai and Childers note that being caught up in object properties (seen usually in younger participants) is one hallmark of structural alignment, but could statistical learning explain this? Perhaps in statistical learning, learners are attending to specific stimuli and their co-occurrences with other stimuli and thus should be *even more caught up in* specific features of specific stimuli than is seen in structural alignment. What about progressive alignment? In structural alignment, closer comparisons are easier to compare and help less experienced learners. Could close comparisons also be easier to correlate in statistical learning? Is it that statistical learning happens first, earlier in development, and then structural alignment comes in later? If so, what should we make of Hespos et al.'s findings that infants can do some structural alignment or relational processing by 3 months? Further studies with infants incorporating structure mapping is one area that needs attention. We need to consider what data would distinguish these views or create a view in which both mechanisms are at work and describe how they work together.

Another key future direction raised across chapters is the call to further examine the mechanisms that underlie a specific theory. For example, Johnson notes that the mechanisms that underlie statistical learning are not fully described, while Thiessen and Sandhofer and Schonberg explore how memory systems could be one of those mechanisms. Extending these ideas, memory abilities and memory phenomena (e.g., the spacing effect, in Sandhofer and Schonberg) also likely underlie structural alignment processing, but their role needs to be examined in this area. Imai and Childers add that contrasting examples with each other—either for learning an individual verb or learning verbs within an overall system—is useful. How comparison and contrast interact in these views needs further explanation. In structure mapping, contrast should include attending to alignable differences, but there is a different goal when contrasting two examples than comparing them. Sandhofer and Schonberg discuss statistical learning in terms of aggregating information across examples. Is there a role for contrastive information in this theory? The power of contrastive information can be seen in word learning (e.g., Waxman and colleagues) and other domains, but how it fits with comparison theories has not been fully explicated yet.

In conclusion, the most amazing part of learning from multiple examples is that it is powerful! Unique insights arise, as I hope you have seen within each chapter and in this concluding chapter that compares these chapters to each other. Doubtless, if you have read multiple chapters, you have been able to draw your own deeper insights into how comparison works using your comparison abilities. Attention to multiple examples also leads to important transfers to new examples, and we hope you will be able to transfer some ideas to your own research, teaching, and/or interests. Keep comparing! And keep extending your knowledge in new ways. As one chapter noted, "variety is the spice of life." By collecting a variety of thinkers who have tackled similar questions using different tools, I hope we have added to your toolbox.

References

Gentner, D. (2010). Bootstrapping the mind: Analogical processes and symbol systems. *Cognitive Science, 34*(5), 752–775.

Vlach, H. A., Ankowski, A. A., & Sandhofer, C. M. (2012). At the same time or apart in time? The role of presentation timing and retrieval dynamics in generalization. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 38*(1), 246–254. https://doi.org/10.1037/a0025260

Index

A
Abstract rule learning, 17, 18
Accuracy information, 180
Acquired distinctiveness, 66
Active learning
  "self-teaching" model, 205
Aggregation
  background context, 162, 163
  category learning, 164, 165
  conditions, 163, 164
  correlated cues, 165, 166
  shared context, 163, 164
  training and testing context, 164
Alignable differences, 241, 242
Analogical abilities, 81, 83
Analogical learning, 82
Analogical reasoning, 17
Analogy, 230, 238
Autism spectrum disorder (ASD), 20

B
Bimodal distribution, 65
Brain-based methods, 19

C
Categorization
  context-dependent, 124
  familiar animal
    eleven-month-olds, 119, 120
    familiarization, 119
    generalizations, 119
    investigations, 118
    same- and different-breed conditions, 119
    sound properties, 122
  familiar *vs.* unfamiliar, 122
  familiarization, 122, 123
  generalization, 124 (*see also* Inductive reasoning)
  information-processing capabilities, 124
  learning, 124
  link object properties, 120
  memory capacity, 124
  multiple exemplars, 122, 123
  preschoolers, 123
  single exemplars, 121
  unfamiliar animal
    ability to link properties, 115
    "basic-like" categories, 116
    familiarization, 116
    learning, 118
    online, 121
    sound properties, 116
    training task, 117, 118
Category learning, 24, 168, 172
Causal inference, 212
Causal knowledge, 207, 208
Causal learning, 198, 203–205, 209, 213, 214
  accuracy, 211
  children, 206
  interventionism, 207–209
  invariance, 207–209
  research, 206
  self-directed, 206, 214
  SI hypothesis (*see* Search for Invariance (SI) hypothesis)
Causal relationships, 207, 208

J. B. Childers (ed.), *Language and Concept Acquisition from Infancy Through Childhood*, https://doi.org/10.1007/978-3-030-35594-4

Children's causal reasoning, 180
Children's spatial reasoning, 49
Children's social learning
  accuracy information, 186
  dual-process model, 193
  familiarization, 185
  gaze-monitoring paradigm, 185
  hypothesis, 186
  "robust bias to trust", 188
  selective word learning, 183
  spatial cueing accuracy, 188
  statistical regularities, 186, 188
Chunking/clustering models, 22–24
Coarticulation, 60
Coarticulatory environment, 59
Cognitive development, 20, 21, 24
Cognitive science, 11
Communicative Development Inventories (CDI), 144
Comparison abilities development, 248, 249
Comparison, multiple exemplars
  artificial languages, 228
  dependency, 228
  explicit and implicit comparison, 229
  in Gomez's study, 229
  language invites, 233–235
  object similarity, 230–232
  progressive alignment, 232
  relational commonalities, 227
  relational learning, 227
  RMTS, 228, 229
  similarity, 227
  three-element strings, 228
Complementary learning system (CLS), 3, 68, 69
Component-level comparison, 51
Confirmation, 210–213
Confirmation bias, 7
Contingency theory, 226
Contrast, verb learning
  categorization task, 142
  comparison, 140, 142
  events, 139, 141
  explicit contrast, 139–141
  familiar action *vs.* a novel action, 140
  familiar test scene *vs.* a new test scene, 140
  implicit contrast, 139–141
  individual verb meanings, 139
  set of objects, 141
Correlated cues, 166
Cross-situational statistics, 64

D
Declarative and nondeclarative memory systems, 25
Decontextualization
  abstraction, 166
  aggregation, 167, 169
  category development, 168
  exemplars, 168
  support condition, 167
  variability, 168
Developmental disabilities, 2, 20–22, 25
Disambiguation paradigm, 191
Distributed coding system, 70
Distributional statistics, 64–66, 70, 71

E
Ease of learning
  criteria, 151
  entropy, 151
  lexical development, 151
  referents, 152
  regression analyses, 153
  semantic domain, 152
  semantic properties, 152, 153
  word learning, 153
Electrophysiological methods, 19
Entropy, 151
Epistemic trust, 188
Event-related potentials (ERPs), 2, 19
Exemplar *vs.* prototype models, 3
Exemplars
  conceptual scaffolding, 46
  familiarization, 32, 42
  infants' experience, 32
  learning, 45
  memory models, 66, 67
  perceptual similarity, 33
  relational learning, 35
  robust generalization, 47
  spatial reasoning, 53
  spatial relation, 40, 43, 46
  variability, 66–68
Experience-based learning, 1
Experimentation skills development, 200
Exploration, 197–200, 202, 206, 209, 211, 214
Eye-tracking methods, 19

F
Familiarization phase, 190
Familiarization trials, 189
Forced-choice tasks, 251

G
Generalization, 31–33, 36, 40, 42, 45, 47, 49, 52, 106–108, 113, 116, 118–120, 122–124
"Generating effects" theories, PTS, 205
"Generating evidence" theories, PTS, 205
Gentner's structure-mapping theory, 80

H
Habituation, 14, 16, 38, 40, 42, 43, 51, 52, 251
Head-turn method, 17
Hippocampal dynamic theory, 3
Hippocampal formation (HF), 68
Human learning, 197
Human mind's ability, 47
Hypothesis testing, 198
  "ascending numbers", 201
  enumeration process, 201
  and exploration, 209
  goal *vs.* method, 204
  intentional form, 203
  negative tests, 198
  Popper's prescription, 203
  prescriptions, 203
  PTS, 199
  and reasoning, 201
  in scientific reasoning, 210
  variable of interest and ask participants, 200
  variables, 198

I
Implicit contrast, 139
Individual differences scaling (INDSCAL), 5, 148
Induction, 106, 109, 122
Inductive reasoning, 4
  animal's mouth movement, 114
  category-based inductive reasoning, 106
  extension trials, 115
  familiarization phase, 114, 115
  human cognition, 105
  infancy (*see* Infancy)
  motor and cognitive skills, 113
  preschoolers, 106, 107
  unfamiliar and familiar categories, 115
Infancy
  baseline condition, 108
  category labels, 110, 111
  imitation, 108, 109, 112
  interactive task, 112, 113
  predicted condition, 109
  real-world objects, 111
  shape similarity, 110
  shared category labels, 111
  shared count noun labels, 112
  single exemplar, unfamiliar category, 112
  symbolic artifacts, 111
  target object, 108
  touchstone, 110
  violated-expectation condition, 108
Infant attention, 24
Infant development, 8
Infant learning
  discrimination, 12
  explicit (articulable) knowledge, 12
  spatiotemporal visual sequences, 15
  TPs (*see* Transitional probabilities (TPs))
  visual SL, 14
Infant spatial categorization
  abstract representation, 34, 52
  analogical reasoning, 34
  analogies, 47
  cognitive skills, 33, 34
  containment events, 36
  degree of consistency, 46
  dependent variable, 41
  developmental changes, 36
  discrimination, 33, 46
  event perception, 35
  exemplars, 32, 37, 40, 42, 47
  familiarization, 38–41, 44
  learning, 52, 53
  locative terms, 34
  minimal variability, 40
  monochromatic symbols/geometric forms, 39
  object cognition, 35
  occlusion event, 41
  perceptual and cognitive abilities, 35, 36
  perceptual category, 34
  physical reasoning, 34, 45
  relational commonality, 47
  representational change, 48
  semantic category, 41
  static *vs.* dynamic instances, 38
  stimuli, 37, 38
  test events, 44
  young children, 48
Integrative themes, 2
Intermodal preferential-looking paradigm (IPLP), 6, 182, 190
Intervention
  causal invariance, 207–209
  causal learning, 204
  and inference, 214
  informative, 206, 214

Intervention (*cont.*)
primary strategy, 205
types, 205
Interventionism
and invariance, 207–209
Invariance
component, 208
concept, 207
counterfactual dependence, 208
description, 207
and interventionism, 207–209
knowledge, 207, 208
SI (*see* Search for Invariance (SI) hypothesis)

K
Krogh-Jespersen and Echols' procedure, 182

L
Language acquisition, 4, 12, 24, 59, 61–64
Language invites comparison
comparative words, 233
labels, 234, 235
systematic structure, 233, 234
Language, social relational learning, 239
"Learner-as-scientist" approach, 214
Learning
behaviour, 205
contexts, 31
Lexical system
carrying/holding, 142, 143
challenges, 132
ease of learning, 153–154
language-specific linguistic features, 134
noun learning, 132
object saliency, 154
object similarity, 154
semantic dimensions, 142
semantic domain, 133
semantic features, 133
semantic structure, domain, 144
similarity structure, semantic domain, 146–151
three age groups, children *vs*. adults, 145, 146
verb representations, 134
verb types, 145
vocabulary growth, 144
word learning, 132
Lexicon, 133, 142, 146, 154, 155
Limit testing, 211
Linguistic development, 59
Linguistic organization, 61
Looking-time task, 85

M
Match-to-sample (MTS), 89, 91, 97
Maturational development, 59
Medial temporal lobe (MTL), 68
Memory abilities, 252
Memory-based approach, 71
Memory systems, 172
Memory trace, 170
Monomodal distribution, 65
Multidimensional scaling (MDS), 5, 146
Multimodal similarity, 135, 136
Multiple exemplars of relations
challenges, 222
idempotent/chirality, 222
object *vs.* relational matches, 223, 224
relational labels, 222, 223
relational similarity, 223
SMT (*see* Structure mapping theory (SMT))

N
Natural language, 63
"Near-optimal" learning strategy, 204
Neural processing, 3
Novel noun generalization task, 163
Novel object trials, 182

O
Object/feature similarity, 223
Object *vs.* relational similarity, 240
Ordinal information, 16
"Outcome-focused" theories of PTS, 202–203

P
Phoneme identity, 61
Phonology, 65, 71
Positive testing strategy (PTS), 7
accounts and overlapping evidence, 205
and causal learning (*see* Causal learning)
characteristics, 214
definition, 199
"generating effects" theories, 205
means of generating evidence, 203–205
in rule learning, 201
in scientific reasoning, 199–200
young learners, 199
Preferential looking, 115

Premise-conclusion similarity, 4
Preschool-aged children, 49
Preschool children's mapping, 50
Progressive alignment, 48, 231, 232
Propositional logic, 212
Prototype model, 66

R
Relational commonalities, 50
Relational labels, 222, 223
Relational learning, 2, 3, 33, 35, 227, 228, 230, 234
  abstract relations, 85, 87
  alignment, relational structure, 86, 88
  analogical learning, 86
  assumption, 89
  attention, 82
  breadth of *alignable*, 88
  common labels, 83
  generalization, 88
  habituation trials, 86
  idea, 89
  measure sameness, 91
  MTS, 89
  progressive alignment, 88
  and reasoning, 83
  RMTS, 89–91
  same-different task, 90, 91
  social (*see* Social relations)
  structural alignment, 88
Relational match-to-sample (RMTS), 82, 89–91, 97, 227–229, 231, 235, 236, 238
Relational transfer tasks, 50
Reliability judgments, 191
Retention
  effects, 172
  interval, 172
  memory, 169
  supports, 170
  time points, 169
  word learning, 173
  word-object, 169, 170
Rule learning
  PTS, 201

S
Same/different discrimination
  human repertoire, 84
  infants, 95
  match-to-sample task, 90
  nonhuman primates, 90
  nonhuman species, 90
  RMTS, 91
  sameness and difference, 90
  species, 90
Scientific reasoning, 202, 203, 210
  PTS, 199–200
Search for Invariance (SI) hypothesis
  causal relationships, 198
  causal variable, 209
  characteristics, 209
  "engineering desirable outcomes", 210
  learners' tendency, 213
  negative tests, 209
  PTS (*see* Positive testing strategy (PTS))
  "seeking confirmatory evidence", 211–213
  self-directed learners, 214
  sufficiency, 211–213
  testing (*see* Hypothesis testing)
Selective learning
  accuracy information, 181, 182
  associative mechanisms, 181
  environment, 189
  familiarization, 189, 190
  interpretation, 181
  IPLP, 182, 190
  Krogh-Jespersen and Echols' procedure, 182
  manual forced-choice procedure, 183
  novel object trials, 182
  objects-related, 183
  rational mechanism, 193
  situational factors, 193
  speakers' accuracy, 181
Self-directed behavior
  "science *vs*. engineering models", 202
Self-directed learning
  characteristics, 198
"Self-teaching" model, active learning, 205
Semantic domain
  Common Space, 148, 149
  dimensions, 148, 150
  Individual Space, 150
  INDSCAL, 148
  MDS, 146, 147, 151
  progressive alignment, 148
  semantic features, 147
  similarity structure, 146
  types of objects, 148
Sensory modalities, 13
Sequential learning engine (SEQL), 95, 96
Shared context, 162–164
Similarity-based activation, 70
Social categorization, 237, 238
Social comparison, 240

Social domain, 7, 8, 238–242
Social interactions, 191
Social reasoning, 237
Social relations, 249
  alignable differences, 241, 242
  domains, 235
  false belief, 237
  goals, 236, 237
  language, 239
  vs. object matches, 237–238
  object similarity, 239
  relational sameness, 236
  social comparison and object *vs.* relational similarity, 240
  social domain, 235
Sound symbolism, 135, 136
Spatial cognition, 33, 52, 53
Spatial learning, 52, 53
Spatial relations, *see* Infant spatial categorization
Spatial skills, 33, 34, 52, 53
Speaker-label-referent association, 189, 190
Speaker reliability, 190
Speakers' accuracy, 181
Speech perception, 60, 61
Statistical learning (SL), 2, 3, 6, 8, 51, 248
  abstract rule learning, 16
  acoustic cues and phonemic identity, 60
  capacities, 181, 185, 189
  chunking models, 22–24
  cognitive development, 20, 21
  cross-situational experiments, 161
  definition, 160
  developmental disabilities, 20, 21
  discrimination, 16
  distributional information, 11
  distributional statistics, 64–66
  encoding and generalization, 68, 69
  environmental structure, 11
  exemplars, 60
  independent of surface features, 17
  infants (*see* Infant learning)
  informative, 60, 61
  language acquisition, 61–64
  mechanisms, 180, 188, 249, 250
  memory, 69–71, 250
  ordinal information, 16
  perceptual primitive, 17
  probabilistic sequences, 16
  real-world knowledge, 193
  reduplication pattern, 17
  referential ambiguity, 161
  regularities, 160, 161
  segmentation tasks, 18
  social cues, 16
  speech stream, 17
  and structure mapping, 250
  testing methods, 19, 20
  utterances, 59
  variable exemplars, 66–68
  word-referent pairs, 161
Statistical *vs*. chunking models, 23
Structural alignment, 2, 47–49, 51, 251
Structure mapping, 250, 252
Structure mapping theory (SMT), 240
  alignment, 226
  arbitrary relations, 225
  artificial grammar learning, 94, 95
  commonality, 224
  critical far test, 94
  description, 92
  early development, 92
  exemplars, 225
  findings, 94
  language learning, 92, 93
  learning trials, 94
  learning verb, 93
  learning, 224, 225
  methodology, 226
  progressive alignment, 92–94
  same-different learning, 95
  SEQL, 95
  SME, 95
  social domain, 238–242
  structural alignment, 92, 93
  teaching trials, 96
  vehicles, 225
  visual alignment, 96
Structured sequence, 14
Structure-level comparison, 51
Structure-mapping engine (SME), 4, 95
Structure-mapping theory, 2, 3, 8
Study-phase retrieval theory, 6
Summary statistics, 64
Surface-level similarity, 48, 49
Symbolic simplicity, 50

T

TP-structured *vs*. pseudorandom sequences, 16
Transitional probabilities (TPs), 2, 62–64
  description, 13
  frequency-balanced design, 13
  habituation, 14
  infants, 12
  perceptual and cognitive functions, 13

Transition-finding models, 22
Truncated recursive autoassociative chunk extractor (TRACX), 22

V
Verb learning, 4, 5
  contrast (*see* Contrast, verb learning)
Verb lexicon, 133
Verb meanings
  argument structure, 135
  lexical system (*see* Lexical System)
  multimodal similarity, 135, 136
  object similarity, 136–138
  syntactic cues, 135
Verbs
  lexical system (*see* Lexical system)
Visual statistical learning, 14
Vocabulary
  acceleration, 168
  acquisition, 160
  and experience, 173
  development, 168
Voice onset time (VOT), 3

W
Wason selection task, 201
Williams syndrome, 21
Word learning, 6, 65, 132, 153
  abstraction, 166–169
  aggregation, 162–166
  generalization, 160
  retention and memory, 169–173
  statistical learning (*see* Statistical learning)
  tasks, 251
Word segmentation, 62