The Development of Collaborative Active Learning and Coordination in Early School-Aged Children

By

Andrew G. Young

A dissertation submitted in partial fulfillment of
the requirements for the degree of

Doctor of Philosophy
(Psychology)

at the
UNIVERSITY OF WISCONSIN-MADISON
2016

Date of final oral examination: 8/15/2016

This dissertation is approved by the following members of the Final Oral Committee:
Martha W. Alibali, Professor, Psychology
Charles W. Kalish, Professor, Educational Psychology
Timothy T. Rogers, Professor, Psychology
Karl S. Rosengren, Professor, Psychology
Vanessa R. Simmering, Assistant Professor, Psychology

Acknowledgements

My graduate training was supported in part by a fellowship through the Interdisciplinary Training Program in the Education Sciences. Thanks to the children, parents, and teaching staff who made this research possible, especially Stephanie Murphy and the teachers of the YMCA of Dane County Afterschool Programs. I would also like to thank the research assistants who helped me conduct and think about this work: Marika Dobos, Erin Ortiz, Bailey Waters, Jessica Foley, Felicia Garland, Sarah Krier, and Taylor Rozman.

With great enthusiasm, I would like to sincerely thank my advisors Martha Alibali and Chuck Kalish for their mentorship. I am grateful for the guidance, patience, and friendship you’ve shown me in both my professional and personal life.

Thanks to Tim Rogers, Vanessa Simmering, and Karl Rosengren for serving on my committee and for your support and advice throughout graduate school.

I am grateful for the friendship and general awesomeness of my fellow labmates of past and present: Shanta, Sunae, Pooja, Noelle, Annie, Amelia, Sarah, and Andrea.

Thank you to my parents and brother for your love and encouragement in all of my endeavors. Mega thanks to my wife Anna for your music, friendship, and love.

Finally, thank you to the great city of Madison, Wisconsin. It’s been real.

Table of Contents

Abstract

Does collaboration affect how children generate and learn from evidence? Despite the importance of collaborative learning in major theories of child development, there has been surprisingly little research on the development of collaborative learning skills. This dissertation explores early school-aged children's collaborative and individual learning in an active category-learning task. Younger (5- to 6-year-olds) and older (7- 8-year-olds) children played a novel board game in which they queried a continuous two-dimensional category space in order to learn the categorical preferences of a puppet (e.g, liking small toys, not large toys). In two collaborative conditions, dyads jointly selected learning examples. In an individual condition, children completed the game alone. Collaboration improved older children's learning over individual learning. However, collaboration did not improve (and sometimes hindered) younger children's learning. Investigation of children's collaborative interactions revealed developmental differences in children's coordination of shared representations (Chapter 3), information gathering actions (Chapter 4), and verbal communication (Chapter 5). Together, these data demonstrate children's ability to profit from collaborative learning rapidly develops in the early school years. Moreover, these results provide new insight into the development of specific skills that support productive collaboration learning.

Chapter I: Introduction

Much of what humans accomplish is done with others. We achieve goals that would be impossible as individuals such as moving heavy furniture, erecting buildings, and carrying out scientific research. We also learn from and with each other in species-unique ways, allowing for the transmission and creation of culture (Tomasello, 2016). The idea that children learn from collaborative experience is widely held across disciplines, theoretical perspectives, educational practitioners, and the public (Hmelo-Silver, Chinn, Chan, & O'Donnell, 2013). Indeed, use of collaboration as a means to effective learning (vs. individual learning) is among the most dominant and long-lasting pedagogical practices throughout the world and across the lifespan (Johnson & Johnson, 2009). In addition, collaboration has recently been characterized as a "21st century skill" essential to children's success in their adult lives (Dede, 2010; Kuhn, 2015; Trilling & Fadel, 2009). Thus, understanding when and how working with others can support effective learning is of substantial theoretical and practical importance (Dillenbourg, 1999).

The focus of this dissertation is collaborative learning between child peers. Though theorists from different traditions emphasize different potential mechanisms of collaborative learning, within the field of cognitive development, collaborative learning has long been viewed in terms of achieving shared thinking (Dillenbourg, Baker, Blave, & O'Malley, 1996). Sociocultural theorists emphasize collaborative interactions with experienced partners as opportunities for guided exposure to more advanced cultural practices and tools (Vygotsky 1978; Gauvain, 2001; Rogoff, 1998). Constructivist theorists emphasize cognitive conflict and consequent resolution processes, such as disagreement and dialogue, among partners as opportunities for the co-construction of novel cognitive and behavioral forms (Piaget, 1926;

Ames & Murray 1982; Howe, 2009). In both cases, children in pursuit of shared learning goals coordinate and modify their own actions and perspectives to construct and maintain shared understandings of problems (Dillenbourg, 1999; Roschelle & Teasley, 1995; Tomasello, Kruger, & Ratner, 1993). The subsequent internalization of such co-constructions serves as the basis for cognitive growth.

Working with partners towards shared goals has been shown to benefit, among other things, children's planning (Radziszewska & Rogoff, 1991), executive functioning (Ou, 2011), mathematical, scientific, and spatial problem-solving (Phelps & Damon, 1989; Teasley, 1995; Tudge & Winteroff, 1993; Schwartz; 1995), literacy (Brown & Campione, 1990), categorization (Fawcett & Garton, 2005), memory (Sommerville & Hammond, 2007), and conceptual change (Howe, 2009). However, research has not always supported the benefits of collaborative learning (e.g., Gauvain & Rogoff, 1989; Sampson & Clark, 2009; Schwarz, Neuman, & Biezunger, 2000). Early research aimed at identifying the conditions under which collaboration improves learning generated positive results that often failed to generalize beyond specific tasks and participant characteristics (e.g., age and levels of prior knowledge; Dillenbourg et al., 1996). In response, researchers shifted focus to the types of interactions between learners, in order to infer mechanisms that could explain why some collaborators learned better than others (Dillenbourg et al., 1996). A majority of this research explores how cognitive-linguistic processes (e.g., argument, negotiation, explanation, mutual-regulation, conflict resolution) function in establishing and maintaining shared problem and solution understandings in middle and high school-aged students (e.g., Baker, 2015; Howe, 2009; Kuhn, 2015; Roschelle & Teasley, 1995; Schwartz, 1995).

Emphasis on late developing cognitive-linguistic processes has led to an empirical blind spot on collaborative learning in early school-aged children. Additionally, influential theoretical proposals suggest effective collaborative learning likely only emerges around 7 years of age, when children's second-order theory of mind and verbal abilities allow for recursive and reflective peer dialogue (Piaget, 1977; Tomasello et al., 1993). Thus, despite the importance of peer collaboration in major theories of child development and learning, there has been surprisingly little research on the development of collaborative learning skills (Flynn, 2010). However, recent years have seen a abundance of research on the development of children's collaboration more generally, both to further understand human cognitive and social development from early infancy (Brownell, 2011) and to try and understand what differentiates humans from our nearest primate relatives (Tomasello, Carpenter, Call, Behne, & Moll, 2005). These lines of research demonstrate children are surprisingly capable of engaging in productive and sophisticated peer collaborative problem-solving by 4 or 5 years of age, suggesting study of the development of collaborative learning should be extended to early school-aged children.

Young Children's Collaborative Problem-Solving Skills

Collaborative problem-solving requires the simultaneous coordination of several different behaviors. These include motivations and commitments to act collaboratively (e.g., Butler & Walton, 2013; Rekers, Haun, & Tomasello, 2011) and social-cognitive capacities to understand when one has joint attention, joint goals, and joint commitments with a partner (Tomasello & Hamann, 2012). Perhaps the most critical component of mature collaborative activity is the ability to form and maintain a shared-intentional representation of a collaborative task (Bratman, 1992; Tomasello et al., 2015). That is, collaborators must coordinate their actions by simultaneously representing their own actions and partner's actions in relation to an overarching shared goal. Such a "bird's-eye view" representation comprises the whole collaboration from an agent-neutral external viewpoint (Tomasello, 2014; Toumela, 2007).

Partners operating with a “bird’s-eye view” of a task can maintain a collaboration of interrelated roles in the face of perturbation. For example, collaborators should be able to switch roles if needed and or adjust their own action plans in response to a partner’s changing actions. By 4 to 5 years of age, children seem to readily form and maintain shared representations that facilitate such flexible collaborative activity. Preschool-aged children will spontaneously coordinate complementary roles and support each other towards shared problem-solving goals (Hamann, Warneken, & Tomasello, 2012; Warneken, Steinwender, Hamann, & Tomasello, 2014), simultaneously represent their own and a partner’s roles in relation to joint goals (Fletcher, Warneken, & Tomasello, 2012; Rakoczy, Gräfenhain, Clüver, Dalhoff, & Sternkopf, 2014), and incorporate a partner’s role into their own action plans (Milward, Kita, & Apperly, 2014; Saby, Bouquet, & Marshall, 2014).

Because preschool and early-school aged children are successful collaborative problem solvers, might they also be effective collaborative learners? Two recent studies suggest young children’s collaborative skills do foster learning within problem-solving tasks.

Children’s learning from collaborative problem-solving. Representing collaboration from a “bird’s eye view” should enable partners to flexibly switch roles, as carrying out one role entails mental representation, and therefore, understanding of a partner’s role. Researchers have looked for evidence of role-reversal learning, or learning about a partner’s role while performing one’s own role, as a test for such agent-neutral representation of collaborative activity. Fletcher, Warneken, and Tomasello (2012) found evidence of role-reversal learning in a problem-solving task requiring the coordination of two complementary roles. The task involved sending a marble from one side of an apparatus to the other via a trap door (Role A) and then back via a spring launcher (Role B) until the marble finally exited. Children's role efficiency was measured during their initial experience with their respective roles, and then in a second round after switching roles. The critical comparison concerned whether children in the switched role demonstrated greater efficiency than their partner had during the initial round, suggesting learning about a partner's role from engaging in collaboration. The data revealed 5-year-olds, but not 3-year-olds, demonstrated greater efficiency upon changing roles, and therefore learning about a partner's role.

Productive collaboration must be also flexible. One needs to be able to modify one's own action plans in response to a partner's changing action opportunities in a manner that maintains pursuit of a joint goal, otherwise collaborations would simply fall apart. Warneken, Steinwender, Hamann, and Tomasello (2014) examined whether young children (3- and 5-year-olds) could plan their own actions on the basis of anticipating what actions a partner might be able to take. In one condition, both partners were presented with boxes containing two tools. Both tools were required to retrieve a prize from an apparatus across the room, and thus, success required partners choosing different tools to bring to the prize box. In another condition, one child received a box containing only one of the tools, and thus, success required the other child to reason about their partner's constrained option and choose the alternative tool. In both conditions, 5-year-olds successfully solved the problem. Importantly, children's planning communication (e.g., announcing one's choice or suggesting a tool to one's partner) predicted success. This effect of verbal planning resembles those found with older children across a variety of collaborative learning tasks (e.g., Dillenbourg & Traum, 2006; Rogoff, 1998; Teasley, 1995).

In the same task, 3-year-old dyads failed to retrieve the prize above chance levels in either condition. However, 3-year-olds did become successful in a particular two-phase version of the task. Three-year-olds who first experienced the both-tool condition and then participated in the constrained condition still failed to successfully coordinate their actions, but children who participated and failed in the constrained condition later performed successfully in the both-tool condition (5-year-olds also benefitted). More experience itself did not improve children's performance, but exposure and response to a partner's constrained action possibilities did subsequently facilitate children's task understanding and performance.

A need to coordinate with a partner's constrained actions has been show to facilitate collaborative learning in other tasks. A related result was observed for the 3- and 5-year-olds in the role-switching task described earlier (Fletcher et al., 2012). A second study constrained the task by making the complementary roles temporally dependent upon one another (i.e., once Role A launched the ball, Role B had to act within a certain temporal window or else the ball would roll back). In this version of the task, 5-year-olds demonstrated greater role-reversal learning than on the non-temporally dependent apparatus and 3-year-olds demonstrated some role-reversal learning (compared to none in the in the unconstrained task).

Young Children's Collaborative Learning

Preschool and early school-aged children can effectively coordinate and learn within collaborative problem-solving tasks. To what extent do these skills generalize to tasks with shared learning goals? There are several recent studies suggesting young children have difficulty forming and maintaining shared representations of tasks when learning with partners.

Asymmetric role representation and learning. Successful collaborators should represent and learn about their partner's actions at the same time they represent and learn about their own (i.e., a "bird's-eye view" representation and role-reversal learning; Fletcher et al., 2012). Kushnir, Wellman, and Gelman (2009) found 4-year-olds' role representation and learning in a causal learning task were asymmetric. When faced with probabilistic evidence arising from joint actions (i.e., two buttons being pressed simultaneously by a child and an experimenter), 4-year-olds errantly inferred their own actions were responsible for causal effects, even when prior evidence strongly supported a partner's actions were responsible. Preschoolers did not exhibit this error when they observed identical evidence, nor when evidence was deterministic. This self-agency bias, being more influenced by one's own contributions to a collaborative learning task, can result in systematically inaccurate learning, potentially leading to detrimental effects of collaboration.

Riggs and Young (2016) found evidence of similar asymmetries in children's learning of game rules. Younger children (4- and 5-year-olds) better learned and were more likely to enforce game rules that governed their own role in a collaborative game than rules governing a peer partner's role. However, older children (6- and 7-year-olds) demonstrated symmetrical learning and enforcement of game rules across roles. Thus, children's abilities to form and maintain "bird's-eye view" representations of learning tasks continue to develop into the early school years.

Learning from collaborative joint action. In an initial study that laid the groundwork for this dissertation, Young, Alibali, and Kalish (under review) examined whether early school-aged children's causal learning from collaborative joint action differs from their learning from independent action or observation. To address this question, they asked kindergartners and first graders to perform equivalent causal interventions on novel machines, either jointly with an adult partner or by themselves, or to observe an adult partner perform the interventions. They found collaborative joint action improved first graders' learning over acting alone and observing a partner. Kindergartners demonstrated a different pattern of results. Collaborative joint action impaired kindergartners' inductive causal inference, compared to acting alone and observing a partner. Children in both grades demonstrated similar learning when acting alone and observing a partner, but first graders showed vastly superior learning to kindergartners in the collaborative condition (75% vs. 25% accuracy). Interestingly, these effects seemed to arise from developing abilities to coordinate learning from joint actions, rather than a diminishing self-agency bias.

Task limitations. While the studies above suggest younger children are less effective collaborative learners than collaborative problem solvers, the tasks reflect somewhat limited forms of collaboration (Kushnir et al, 2012; Riggs & Young, 2016; Young et al., under review). For example, the studies cited all had the goal of equating the information available to collaborators and non-collaborators. Thus, the causal learning tasks involved highly scripted interactions with adult experimenters that required minimal coordination between partners, as the action possibilities were entirely fixed. Furthermore, the interactions were devoid of linguistic coordination between partners. The limited interaction and coordination demands of these tasks potentially masked some of young children's collaborative learning abilities. It is also possible children's failure to integrate and learn from the outcomes of their own and partner's actions might be specific to the domain of causal learning, as representation of causal relations are thought to be action-based (Lagnado & Sloman, 2004; Sobel & Kushnir, 2006; Styvers, Tenenbaum, Wagenmakers, & Blum; 2003). Finally, none of tasks involved children controlling their own learning experience (with exception of one condition in the rule learning study; Riggs & Young, 2016). In other words, the examples children experienced were fixed. Research with school-aged and adolescent students suggests peer interaction is more likely to benefit learning when active exploration and hypothesis testing is possible (e.g., Schwarz et al., 2000). It is possible younger children will demonstrate more effective collaborative learning in tasks where they too are able to control their own learning experiences.

Current Research

The overarching goal of this dissertation is to better understand when and how collaboration is effective for learning in the early school years. What collaborative learning skills develop between age 5, when children are successful collaborative problem-solvers, and ages 7 and 8, when collaboration has demonstrated benefits over individual learning (Flynn, 2010; Young et al., under review)? When children are in control of their own collaborative activity, what interaction and coordination behaviors contribute to successful (or unsuccessful) collaborative learning?

To study early school-aged children's developing collaborative learning and coordination, I adapted a two-dimensional category-learning paradigm recently used to investigate relations between individual adults' active information sampling behavior and learning (Markant & Gureckis, 2014). Participants in the task select and query individual items for category membership (i.e., ask if an item belongs to category A or B) and then perform a categorization task, revealing inferred category rules. This is an active learning (or self-directed learning) task, because participants select their own learning examples. Although self-directed procedures were common in early research on category learning (Bruner, Goodnow, & Austin, 1956; Huttenlocher, 1962), they have garnered little recent attention despite their relevance to classroom learning. Active versus passive learning is central a very influential topic in the education and learning sciences (Bruner, Jolly, & Sylva, 1976; Montessori, 1964; National Research Council, 1999; Piaget, 1930), and an emphasis on active learning underlies motivations behind instructional methods such as discovery, experiential, and inquiry learning (Bruner, 1961; Klahr & Nigam, 2004; Kolb, 1984; Papert, 1980).

This dissertation presents children with a two-dimensional active category-learning task in which children select and ask about individual items in order learn about a character's categorical preferences. The task provides a number of useful features for purposes of studying children's collaborative learning. First, it affords a reasonable division of labor into complementary roles. That is, members of a dyad can be assigned to control different dimensions. Second, collaborating children necessarily have to coordinate their actions to select learning items. Thus, in addition to learning outcomes, the task provides behavioral data on children's coordination, communication, and information sampling. Third, the task can be implemented with a variety of category structures (e.g., unidimensional and two-dimensional categories) that provide unique opportunities to investigate children's collaborative and individual learning. Finally, learning in similar tasks has been investigated extensively, and thus there are a number of computational tools to characterize participants' learning (e.g., Ashby & Maddox, 2005; Markant & Gureckis, 2014).

I investigate children's collaborative learning in two dyadic conditions. To effectively learn, partners individually responsible for specifying the level of a single dimension need to coordinate their respective selections to query informative category labels. In one condition (*dyad)*, children are free to complete the task as they please. In a second condition (*constrained dyad*), children follow a turn-taking policy, such that partners alternate the dimension on which they make the first selection for each query (i.e., on one query, partner A goes first and selects a level on their own dimension. On the next query, partner B goes first and selects a level of their assigned dimension). Thus, each partner has a structured opportunity to wait for, attend to, and adjust to his or her partner's selection action. I included this condition because prior research suggests children's representation of and learning within collaborative problem-solving tasks can be facilitated by the presence of action and temporal constraints (Fletcher et al., 2012; Warneken et al., 2014).

With this task, I also investigate individual children's learning in order to evaluate whether working with a partner improves learning, and to examine the characteristics of collaborative activity that promote superior learning. In addition, developmental differences occurring in individual learning can be used to benchmark potential developmental differences in collaborative learning (Kuhn, 2015).

Chapter Overview

In Chapter 2, I describe the general experimental and analytical methods.

In Chapter 3, I investigate children's learning in the task. Specifically, I compare dyad, constrained dyad, and individual learning in younger (5 and 6-year-old) and older (7 and 8-year-old) children. I additionally examine whether the extent to which collaborating children form a "bird's-eye view" representation (as opposed to an asymmetric representation) of the task affects learning outcomes.

In Chapter 4, I examine children's information sampling behaviors and relations between their information sampling and learning. One way collaboration could help (or hinder) learning is by influencing the kind of evidence children generate. I specifically evaluate whether measures of sample informativeness and sample complexity vary among dyads, constrained dyads, and individuals, and whether they vary as a function of children's age. Furthermore, I investigate whether the predictive relations between sample informativeness and sample complexity vary among age groups and conditions.

In Chapter 5, I examine dyads' verbal communication and relations between verbal communication and learning. Specifically, I evaluate whether younger and older children differ in their production of verbal planning and evidence interpretation. Furthermore, I investigate whether the predictive relations between verbal communication and learning are stable or vary across development.

Thus, the present research directly investigates relations between children's collaborative coordination skills, information sampling, and learning.

Chapter II: General Methods

Participants

One hundred and three younger children (5- to- 6-year-olds; M = 6 years, 2 months) and 110 older children (7- to- 8-year-olds; M = 7 years, 10 months) were recruited from afterschool programs, summer programs, and research participant databases in a mid-sized Midwestern city. No formal demographic measures (other than age and gender) were collected.

Children participated in one of three conditions: *individual* (younger: N = 21; older: N = 26), *dyad* (younger: N = 42; older = 42), and *constrained dyad* (younger: N = 40; older: N = 42). Previous research on children's collaborative learning suggests dyads composed of friends produce more sophisticated problem solving strategies and discussions than dyads composed of acquaintances (Azmitia & Montgomery, 1993). Because the present study sought to evaluate early school-aged children's abilities to coordinate collaborative learning, children in the dyadic conditions participated with a friend or frequently playmate. Children recruited from childcare programs were assigned by teachers into same age-group pairs of friends or frequent playmates. These pairs were then randomly assigned to condition such that pairs of children assigned to one of the dyadic conditions (i.e., dyad and constrained dyad) participated together and pairs of children assigned to the individual condition participated individually. Children recruited from databases were randomly assigned to condition prior to contact. Parents of children assigned to the dyadic conditions assisted in recruiting same age-group friends to participate with their children.

Stimuli & Materials

Game boards. A grid-based board game served as the primary task. Two unique stimuli sets were defined by two-dimensional feature spaces (see Figure 2-1). One set of stimuli varied by size (radius of a circle) and number of spikes. The other stimuli set varied by arm length and pattern density (i.e., total area of a green dot pattern). Pilot testing revealed children most frequently described the pattern density dimension in terms of being mostly green or mostly white. Within stimuli sets, each dimension was divided into 7 equally spaced levels (e.g., 7 circle sizes) and then combined orthogonally to generate 49 unique items. Stimuli of this type have been used in studies of perceptual classification (e.g., Nosofsky, 1989; Ashby & Maddox, 2005), and previous work suggests such dimensions are largely separable. Learning game boards were physically instantiated on laminated 36 cm x 36 cm boards.

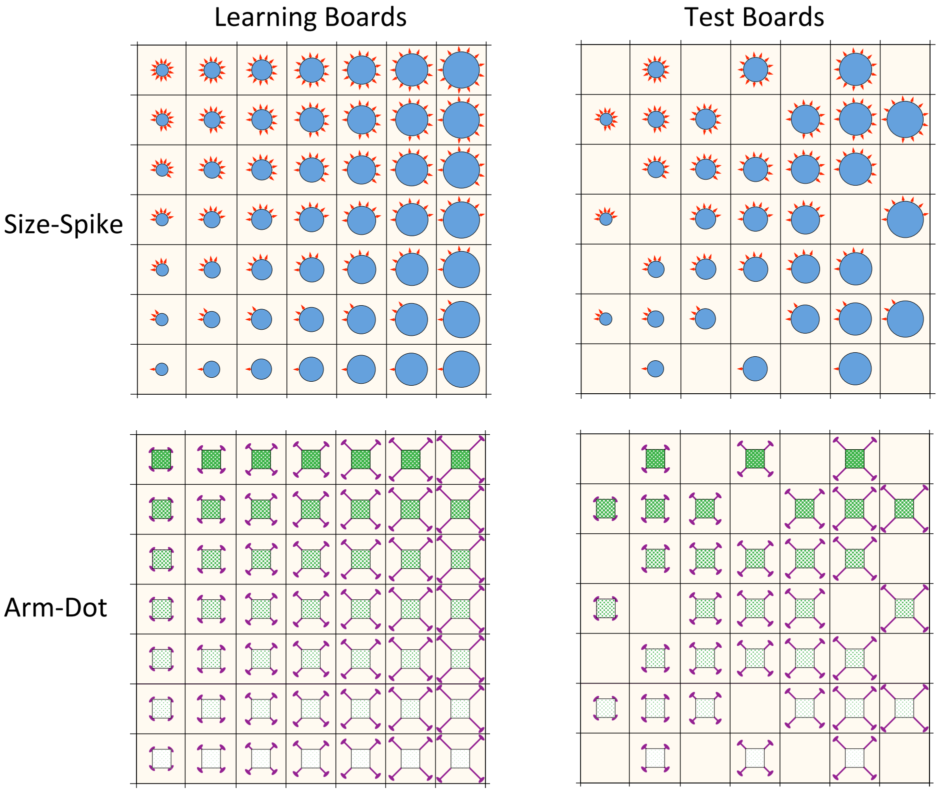

Figure 2-1. Learning and test board stimuli sets.

Other materials. A children's dinosaur puppet (named Dobo) served as the expert informant in the task. White and black tokens that affixed to the game boards served as markers of Dobo's responses during the learning phase and children's categorization responses during the test phase. Two transparent question tools allowed children to select an entire row or column. When both tools were placed on the board, they overlapped to highlight an item for querying. Finally, a practice game board consisted of 8 items varying along a single dimension of twistiness (i.e., the proportion of a line segment that was straight vs. zigzagged).

Procedure

Introduction and familiarization. Children participated with two experimenters at their school or in a laboratory. All sessions were videotaped. Children and experimenters sat such that they were oriented toward the game boards. The primary experimenter (E1) began by telling children they were going play games that involved figuring out what kinds of toys Dobo liked and didn't like. E1 then introduced Dobo, operated by the second experimenter (E2/Dobo). E1 again stated the goal of the games was to determine what kinds of toys Dobo likes and doesn't like, and noted that a good way to figure out what someone likes and doesn't like is to ask them questions.

E1 then presented a practice game board that depicted dinosaur fruits and introduced the items in terms of the dimension end points (i.e., "Some of the fruits are very twisty, and some of them aren't very twisty"). E1 then stated that Dobo liked to eat some of the fruits on the board, but not others. Children were then instructed to ask Dobo about 4 of the fruits on the board in order to learn about what he liked and didn't like. E1 also explained to children that they should try to ask "smart" questions to learn about both what Dobo likes and doesn't like, since they only get to ask half of the fruits.

Children proceeded to ask about individual items sequentially. On the first question, Dobo responded yes by nodding. E1 then placed a white "yes" token on the selected item. For the remaining questions, Dobo responded yes or no depending on whether the item was located in the same half or other half of the practice board. In response items on the negative half of the board, Dobo responded by shaking his head, and E1 then placed a black "no" token on the item.

After the fourth question, E1 highlighted the 4 remaining items and asked children to guess about whether Dobo would like or not like each of them.

Learning phase – 1D trial. Children were presented with the first learning board in which Dobo's categorical preference was structured by a unidimensional linear boundary (Figure 2-2). Stimuli set, relevant dimension and preference (e.g., larger toys, not smaller toys), and board orientation were counterbalanced across condition. There were 8 possible rules: big, small, spiky, not spiky, long, short, green, and white. In the dyadic conditions, one partner was randomly assigned to control the relevant dimension, leaving the other partner in control of the irrelevant dimension.

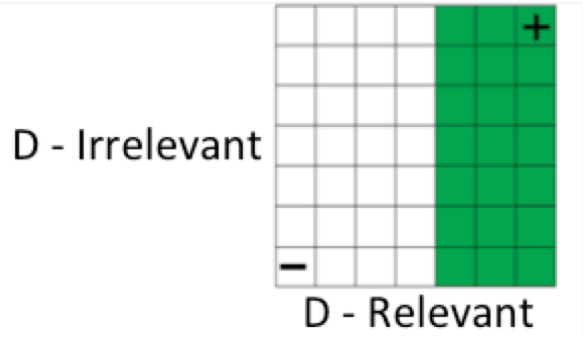

Figure 2-2. Category structure used in the 1D category trial. Green squares represent positive items and white squares represent negative items. Plus and minus signs represent the first and second experimenter questions respectively.

E1 introduced the stimuli by describing the toys in terms of the end points of each dimension (e.g., "Some of these toys are big and some of them are small. And some of these are spiky and some of them are not spiky. Dobo likes some of these toys, but he doesn't like others”). After restating the goal of the game (i.e., “Let’s ask Dobo about some of these toys and try to learn about what he likes and doesn’t like”), E1 began the task by asking about an extreme positive example and then the opposite negative example (e.g., the largest least spiky toy and then the smallest most spiky toy; see Figure 2-2). In doing so, E1 demonstrated the use of the question tools. To choose a particular level of a dimension, a tool was placed on top of the corresponding row. Placement of both tools overlapped and highlighted the toy to be queried (see Figure 2-3).

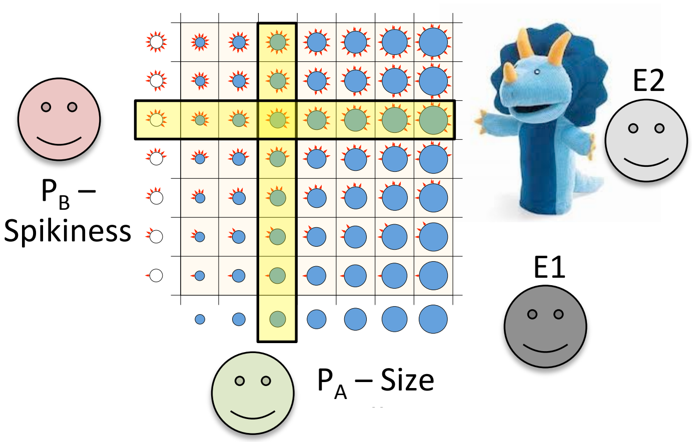

Figure 2-3. General learning phase set-up for dyads.

E1 then transitioned game play to the children. In the dyadic conditions, E1 explicitly assigned children to their respective dimensions and question tools (e.g., “I’m going to give this question tool to [*Child Name*]. That means, [*Child Name*], you are in charge of picking how spiky toys will be”). E1 then explained the following, “Now it is your turn to ask Dobo some questions. You’ll need to work together and talk to each about the toys you want to ask Dobo about. Together you get to ask Dobo 14 more questions. Remember to try and learn about the toys he likes and doesn’t like, because we will ask you about both kinds of toys later.”

Children in the dyad condition proceeded to choose and ask Dobo about 14 toys. After each question Dobo responded with respect to the true category boundary and E1 assisted in placing yes/no tokens to mark Dobo's responses. E1 intermittently updated children with the number of remaining questions and kept children on task when necessary.

Children in the constrained condition were given additional turn-taking instructions prior to beginning. Specifically, E1 told partners that they would need to alternate which partner uses their tool first on each question. When necessary, during game play E1 reminded children which partner was to use their question tool first.

Children in the individual condition were responsible for both dimensions and were given both question tools (i.e., "I'm going to give you both question tools so you can pick how spiky and big the toys will be"). Additionally, following Teasley (1995), E1 encouraged individuals to talk if they'd like to (i.e., "You can also talk while you are thinking and choosing, it may help you figure out what toys Dobo likes and doesn't like.")

Test phase – 1D trial. Dobo's response to children's 14th question concluded the learning phase. Children then individually with E1 or E2 completed the following tasks in order.

Categorization task. In the primary outcome task, children were presented the appropriate test game board in an orientation that matched their learning phase perspective. The experimenter introduced the task as follows, "Here's a board with the same kind of toys you got to ask Dobo about. You need to decide which of these toys Dobo likes and which ones Dobo doesn't like. We want to give Dobo all the toys he likes, but we don't want to give him any toys he doesn't like. Here are some yes/no tokens like we used earlier. You should put a *yes* token on all the toys Dobo will like and an *no* token on the toys Dobo won't like." Because pilot testing revealed that labeling all 49 items was too demanding, the test boards contained only 33 items (see Figure 2-1). Experimenters let children complete the task in their own fashion. If children stopped before categorizing all 33 items, experimenters highlighted and solicited responses for the unmarked items.

Forced choice task. Children completed a generalization item forced choice task presented on a computer or tablet. The structure of the 1D forced choice items is presented in Figure 2-4. Four "inner" items were generated using novel dimension levels within the stimuli space of the learning board. Four "outer" items were generated using dimension levels more extreme than those found on the learning board. Children answered 12 forced choice questions in random order. Four direct-contrast questions paired items whose relationship in the category space crossed category boundary perpendicularly (2 inner pairs and 2 outer pairs). Four diagonal-contrast questions paired items whose relationship in the category space crossed the true category boundary diagonally (2 inner pairs and 2 outer pairs). To keep the task general to all possible 1D rules, 4 no-contrast questions paired the remaining inner and outer items whose spatial relationship in the category space did not cross the category boundary (i.e., both items were negative or positive). For each question, children were asked to choose the item Dobo would like more.

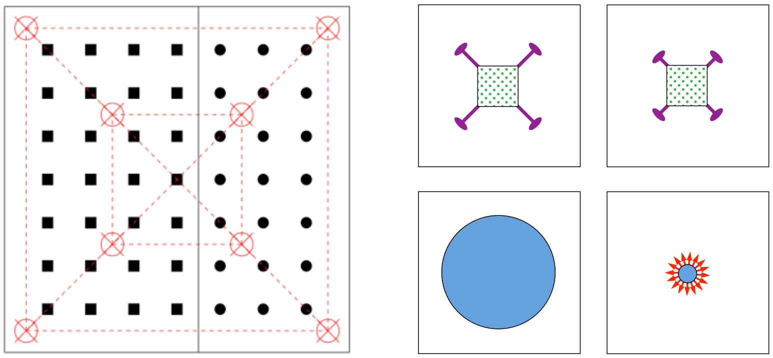

Figure 2-4. Left: Forced choice item structure for the 1D category trial. Black circles and squares represent positive and negative learning item locations. Black line represents true category boundary. Red symbols represent novel forced choice items. Dashed red lines link inner and outer items paired in forced choice questions. Top Right: Example inner direct-contrast. Bottom Right: Example outer diagonal-contrast.

Free response questions. Experimenters asked children the following question, “What kind of toys do you think Dobo liked? Can you use your words to describe them?” If children failed to provide a response, experimenters would ask about whether each dimension mattered (e.g., “Do you think Dobo cared about how big/small the toys were?”).

Learning phase– 2D trial. Upon completion of the first test phase, children played the game a second time. Children were presented with the remaining learning board in which Dobo’s categorical preference was structured by a two-dimensional linear boundary (Figure 2-5). Preference rule and board orientation were counterbalanced across dyads and individuals. There were 8 possible rules: big & spiky, big & not spiky, small & spiky, small & not spiky, long & green, long & white, short & green, short & white. In the dyadic conditions, children were randomly assigned to dimension. Note that since both dimensions were relevant to the 2D boundary, these dimensions did not correspond to dyadic roles (i.e., relevant or irrelevant).

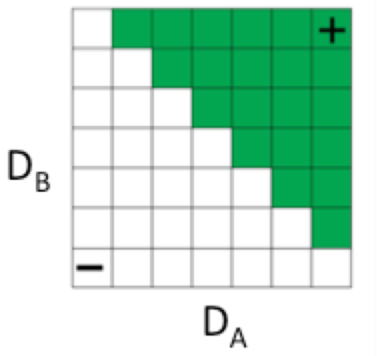

Figure 2-5. Category structure used in the 2D category trial. Green squares represent positive items and white squares represent negative items. Plus and minus signs represent the first and second experimenter questions respectively.

Game play in the 2D trial proceeded exactly as game play in the 1D trial.

Test phase – 2D trial. At the conclusion of the 2D learning phase, children completed the categorization task, forced choice task, and free response questions as they had in the 1D trial. Note, the forced choice item structure was modified to match the 2D category boundary, but maintained the critical properties of the 1D forced choice structure (see Figure 2-6).

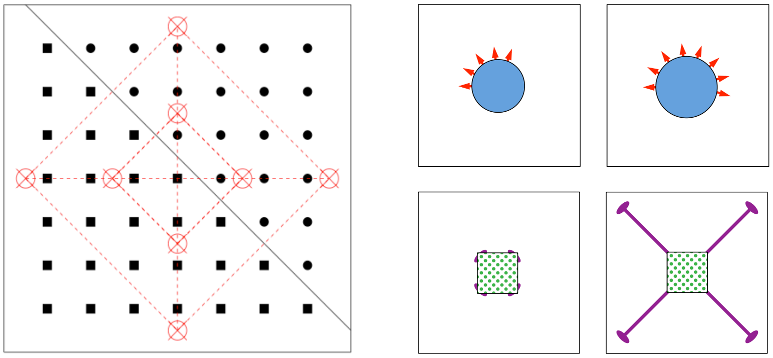

Figure 2-6. Left: Forced choice item structure for the 2D category trial. Black circles and squares represent positive and negative learning item locations. Black line represents true category boundary. Red symbols represent novel forced choice items. Dashed red lines link inner and outer items paired in forced choice questions. Top Right: Example inner direct-contrast. Bottom Right: Example outer diagonal-contrast.

Data Analysis – Learning Measures

Categorization accuracy. Children's 33 test board responses on a given trial were coded as correct or incorrect with respect to the true category boundary.

Model-based strategies. An advantage of using this task is that there are computational models (i.e., decision boundary models) that can characterize a child's categorization responses in terms of more general rule-governed strategies (Huang-Pollak, Maddox, & Karalunas, 2011; Maddox & Ashby, 1993; Markant & Gureckis, 2014). This is valuable as multiple different strategies can yield similar accuracy rates (see Figure 2-7).

Six models were applied to each child's 1D trial data: (1) a unidimensional model along the relevant dimension that assumed the optimal decision boundary, (2) a unidimensional model along the relevant dimension that estimated the decision boundary from the data, (3) a unidimensional model along the irrelevant dimension that estimated the decision boundary from the data, (4) a general linear classifier (GLC) model (i.e., linear 2D model) that estimated the decision boundary slope and intercept from the data, (5) a conjunctive model (i.e., AND 2D model) that estimated a decision boundary for each dimension from the data, and (6) a random responder model that estimated the probability of a yes response from the data. All models included a noise parameter.

Similarly, six models were applied to each child's 2D trial data: (1) a general linear classifier (linear 2D model) that assumed the optimal decision boundary slope and intercept, (2) a GLC model (i.e., linear 2D model) that estimated decision boundary the slope and intercept from the data, (3) a conjunctive model (i.e., AND 2D model) that estimated a decision boundary for each dimension from the data, (4 & 5) a unidimensional model along each dimension that estimated the decision boundary from the data, (6) a random responder model that estimated the probability of a yes response from the data. All models included a noise parameter.

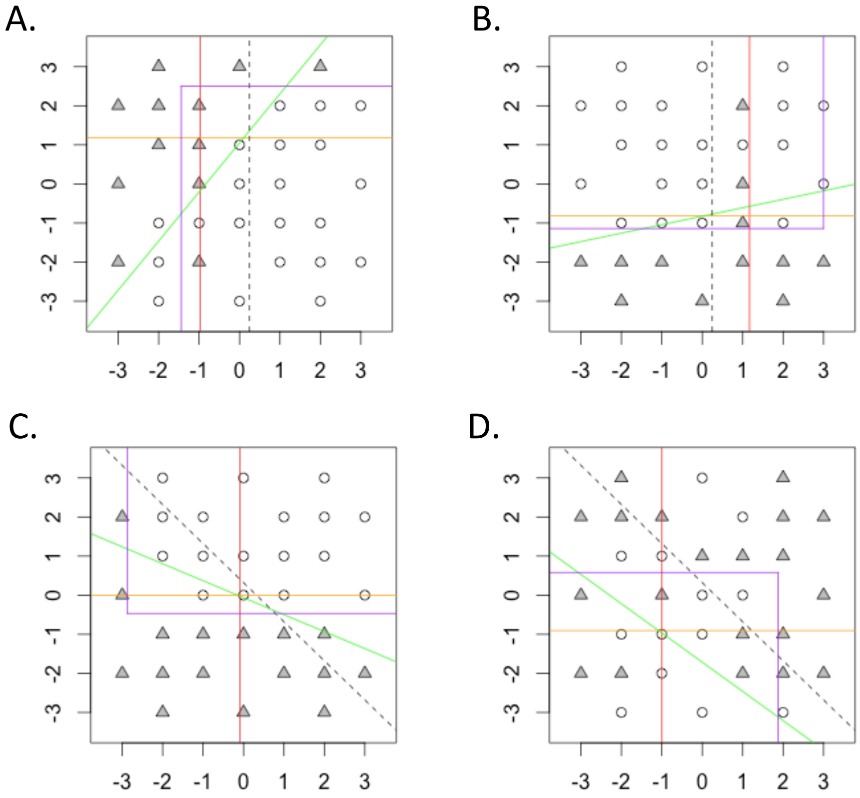

Figure 2-7. Example decision boundary models of four participating children. Dashed lines depict optimal boundaries, vertical red lines and horizontal orange lines depict unidimensional models, green lines depict linear 2D models, and purple lines depict conjunctive 2D models. Categorization responses are represented as white circles (toys Dobo likes) and black triangles (toys Dobo dislikes). Example A was best-fit by the linear 2D model. Example B was best-fit by the horizontal unidimensional model. Example C was best-fit by the conjunctive 2D model. Example D was best-fit by a random responder model.

Model parameters were estimated using maximum likelihood (Maddox & Ashby, 1993). I identified each child's best fitting models using AICc, a finite sample size correction of AIC (Akaike, 1974; Burnham & Anderson, 2002). This criterion assesses the goodness-of-fit of models that differ in the number of free parameters, and selects the model that provides the most parsimonious account of the data (i.e., the model with the smallest AICc value).

For purposes of data analysis, model fits were coded with respect to the correct 1D and 2D trial rules (see Figure 2-7). For the 1D trial, best fitting models along the relevant dimension (either the optimal or estimated boundary model) were coded as correct strategies, best fitting models along the irrelevant dimension were coded as irrelevant dimension strategies, best fitting GLC and conjunctive models were coded as 2D strategies, and best fitting random responder models were coded as random strategies. For the 2D trial, best fitting GLC and conjunction models were coded as correct strategies, best fitting unidimensional models were coded as unidimensional strategies, and best fitting random responder models were coded as random strategies. For children in the dyadic conditions, unidimensional strategies were further coded as own dimension and partner dimension strategies contingent upon whether the child or their partner was in charge of the given dimension during the learning phase. Table 2-1 summarizes children's model-based strategies in the task.

Table 2-1

Percentage of children best fit by model-based strategies by age group, condition, role, and category

		Younger Children					Older Children				
		Dyad		Constrained Dyad		Individual	Dyad		Constrained Dyad		Individual
		Relevant Partner	Irrelevant Partner	Relevant Partner	Irrelevant Partner		Relevant Partner	Irrelevant Partner	Relevant Partner	Irrelevant Partner	
1D Category	Correct Dimension	28.6	38.1	55	25	47.6	61.9	61.9	52.4	57.1	42.3
	Irrelevant Dimension	19	14.3	5	10	4.8	9.5	0	4.8	0	3.8
	2D	9.5	23.8	15	45	28.6	23.8	28.6	28.6	19	23.1
	Random	42.9	23.8	25	20	19	4.8	9.5	14.3	23.8	30.8
2D Category	Correct 2D	26.3	57.9	35.3	47.1	62.5	90	55	60	57.9	54.2
	Own Dimension	36.8	15.8	17.6	17.6	25	10	20	30	21.1	33.3
	Partner's Dimension	10.5	15.8	5.9	11.8		0	15	5	15.8	
	Random	26.3	10.5	41.2	23.5	12.5	0	10	5	5.3	12.5

Note: Individuals in the 2D trial were in charge of both dimensions, and thus the own/partner distinction of incorrect 1D strategies is not applicable.

Forced choice. Children's responses to direct-contrast and diagonal-contrast forced choice questions were coded as rule-consistent or not with respect to the true category boundary (i.e., choose the item Dobo would like more given the rule). No-contrast questions were not analyzed.

Free response questions. Children's free responses were coded for accuracy. However, the reliability of this coding has yet to be assessed; therefore, I do not present analyses of these data.

Data Analysis – General Statistical Methods

Unless otherwise noted, I used binomial generalized linear mixed models (GLMMs) for to analyze the data. GLMMs were fit with *parsimonious* random-effects structures, following the procedure recommended by Bates, Kliegl, Vasishth, and Baayen (2015). This procedure involves performing a principal component analysis on the variance components of a model fitted with maximal random-effects structures (Barr, Levy, Scheepers, & Tilly, 2013) in order to detect over-parameterization and removing random-effects that are not supported by the data. Each model began with a maximal by-participant, by-dyad, and by-rule random-effects structure and was simplified via the above procedure when appropriate. Note by-rule random effects were used in place of full by-item random effects. Preliminary analyses revealed including by-item random effects resulted in several convergence issues, arising from the 33 items being crossed with the 16 possible category rules. However, preliminary analyses revealed by-rule random effects served a similar function and often provided a better fit.

Inference for fixed effects was carried out via likelihood-ratio test (LRT) model comparison (Barr et al., 2013; Bates, Maechler, Bolker, & Walker, 2015). Specifically, I conducted factorial type 3 sums of squares comparisons (Singmann, Bolker, & Westfall, 2015). Simple comparisons (i.e., differences between cell means) are reported in terms of odd-ratios (OR). Inference for simple comparisons was carried out via Wald 95% confidence intervals (Bates et al., 2015). Preliminary analyses revealed no effect of child gender or location (i.e., laboratory vs. school), and thus, these variable were no included in the final statistical analyses.

Chapter III: Collaborative and Individual Active Category Learning in Early School-Aged Children

This chapter summarizes children's learning in the task. The analyses address the following questions and predictions.

Do Younger and Older Dyads, Constrained Dyads, and Individuals Exhibit Different Learning?

Prior research and theoretical proposals suggest effective collaborative learning might not emerge until around age 7 (Flynn, 2010; Tomasello et al., 1993). For example, in a highly scripted causal learning task (Young et al., under review), working with a partner improved first graders' learning compared to working alone, but hindered kindergartners' learning compared to working alone. Furthermore, first graders and kindergartners demonstrated similar learning when working alone, but very different learning when working with a partner (i.e., 75% vs. 25% accurate). I expect the present study will align with these results. Specifically, I predict older dyads will demonstrate better learning than older individuals and younger dyads. Additionally, I predict younger dyads will not demonstrate better learning than younger individuals. Because children's individual learning on comparable active learning tasks has only recently begun to be explored (e.g., Sim, Tanner, Alpert, & Xu, 2015), it is unclear whether older individuals will demonstrate superior learning to younger individuals.

Research on children's problem-solving has found conditions that demand coordination with a partner's constrained actions can facilitate collaborative learning in young children (e.g., Fletcher et al., 2012; Warneken et al., 2014). The manipulation of turn-taking in the constrained dyads was meant to serve a similar function, providing a structured opportunity to wait for, attend to, and adjust to a partner's actions. Thus, I expect constrained dyads' learning to be greater than that of the non-constrained dyads. This might be particularly true for younger children, as they will likely have more difficulty forming and maintaining a "bird's-eye view" representation of the task (e.g., Kushnir et al., 209; Riggs & Young, 2016).

Does Dyadic Learning Vary by Role?

Collaborative learning is unlikely to be effective unless partners coordinate their own actions and perspectives to form and maintain a shared problem representation (Dillenbourg, 1999; Roschelle & Teasley, 1995; Tomasello, et al., 1993; Tomasello, 2016). Asymmetric role representation (i.e., favoring one's own role or actions) has been shown to hinder young children's collaborative learning on causal inference and rule learning tasks (Kushnir et al., 2009; Riggs & Young, 2016). To what extent might asymmetric role representation explain potential developmental differences in the task?

In the 1D trial, I examine role representation by comparing the learning of children assigned to relevant dimensions to that of children assigned to irrelevant dimensions. Several lines of research suggest assignment to an irrelevant dimension might hinder learning compared to assignment to a relevant dimension. As described in Chapter I, preschool-aged children have difficulty separating the effects of their own actions from a partner's in inductive causal learning (Kushnir, et al., 2009). Additionally, instructions to focus on irrelevant dimensions in passive category learning tasks have been shown to disrupt adult learning (Grimm & Maddox, 2013). Finally, successful active category learning in adults is highly contingent on attention to and availability of candidate hypotheses (Markant & Gureckis, 2014; Markant, 2016). Agency over a dimension would likely increase both attention and availability.

If early school-aged children demonstrate poor collaborative learning because of asymmetric role representation, I predict irrelevant role children will demonstrate poorer 1D trial learning than relevant role children. In addition, when considering children's model-based categorization strategies (see Chapter 2), I predict irrelevant role children will use irrelevant-1D and 2D strategies more frequently than relevant role children, as control over an irrelevant dimension might inhibit learning that the dimension is not diagnostic. Finally, I expect these effects will be most pronounced in younger dyads, as I predict children's ability to form symmetrical task representations improves both with age and with exposure to the constrained dyad manipulation.

In the 2D trial, there is little reason to believe children's assignment to a relevant or irrelevant dimension in the 1D trial should affect learning (i.e., both children are in charge of a diagnostic dimension in the 2D trial). However, learning in the 2D trial requires integration across both dimensions. Thus, if early school-aged children demonstrate poor collaborative learning because of asymmetric role representation, I predict children will be more likely to use unidimensional strategies corresponding to their own dimensions than unidimensional strategies corresponding to a partner's dimension. Again, I predict this effect will be most pronounced in younger dyads compared to older dyads and in non-constrained dyads compared constrained dyads.

Results

Did Children's Learning Vary by Condition and Age Group?

I formally examined children's learning using three measures: categorization accuracy, use of correct model-based strategy, and rule-consistent forced choice responses.

Categorization accuracy. Figure 3-1 summarizes children's overall categorization accuracy across condition, age group, and trial.

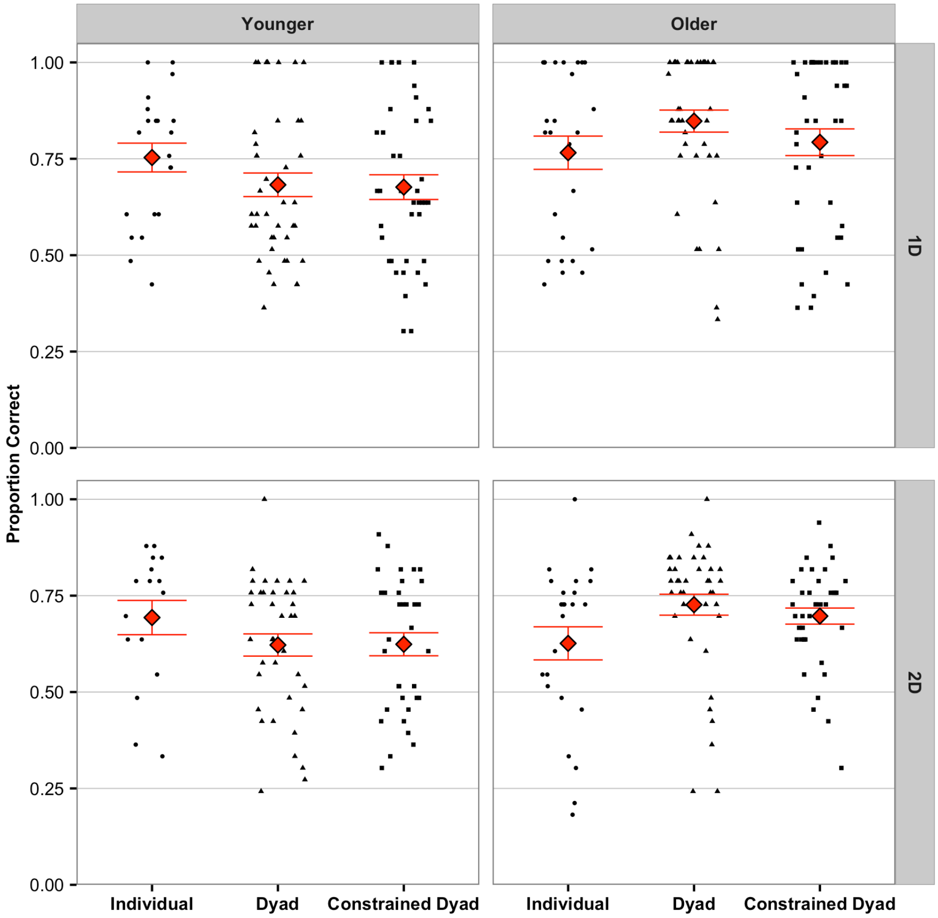

Figure 3-1. Raw categorization accuracy (proportion out of 33 items) by condition, age group, and category. Jittered points represent individuals (two per child), boxes represent means, and error bars represent ± SE.

I first examined children's learning by estimating a binomial GLMM on children's correct categorization with Condition (*individual, dyad, constrained dyad*), Age Group (*younger, older*), Category (*1D, 2D*), and their interactions as fixed effects, and a parsimonious by-participant, by-dyad, and by-rule random effects structure (Bates et al., 2015; see Chapter 2 for more details). The data are presented in Figure 3-2. There was an interaction between Condition and Age Group, LRT $X^2(2) = 6.65$, $p = .036$. Older dyads were more accurate than older individuals, *OR* = 1.58, 95% CI [1.11, 2.78], $p = .016$, and similarly accurate as older constrained dyads, *OR* = 1.28, 95% CI [.83, 1.96], *p* = .263. Older constrained dyads and individuals were similarly accurate, *OR* = 1.37, 95% CI [.86, 2.18], *p* = .172. In contrast, younger children demonstrated similar accuracy across conditions. In addition, older children were more accurate than younger children in the dyad condition, *OR* = 1.96, 95% CI [1.24, 3.06], *p* = .003 and marginally more accurate in the constrained dyad condition, *OR* = 1.53, 95% CI [.98, 2.41, *p* = .063. Older and younger individuals were similarly accurate, *OR* = .84, 95% CI [.49, 1.42], *p* = .506

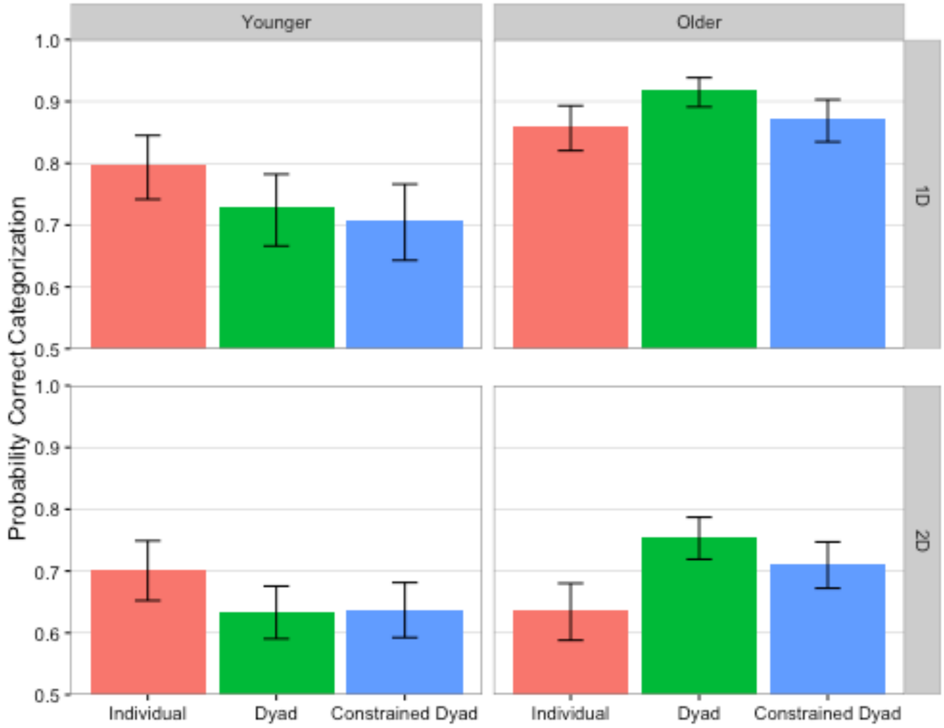

Figure 3-2. Estimated probabilities of correct categorization by condition, age group, and category. Error bars represent ± SE.

The GLMM also yielded a interaction between Age Group and Category, LRT $X^2(1)$ = 27.86, *p* < .001. Older children demonstrated greater accuracy than younger children on the 1D trial, *OR* = 2.04, 95% CI [1.89, 3.57], *p* < .001, but not the 2D trial, *OR* = 1.26, 95% CI [.94, 1.70], *p* = .115.

Correct model-based strategy use. As detailed in Chapter 2, decision-boundary models provide a useful method to characterize the types of categorization rules children used during categorization, as many different strategies can yield similar accuracy rates. Did children's tendency to use a correct categorization strategy (i.e., a correct unidimensional and two-dimensional boundary model in the 1D and 2D category trials) vary by condition and age group? I estimated a binomial GLMM on children's correct strategy use (1 = correct strategy, 0 = other strategy) with Condition (*individual, dyad, constrained dyad*), Age Group (*younger, older*), Category (*1D, 2D*), and their interactions as fixed effects, and a parsimonious by-participant, by-dyad, and by-rule random effects structure random effects structure.

As seen in Figure 3-3, there was an interaction between Condition and Age Group, LRT $X^2(2) = 7.11$, $p = .029$. Older dyads were more likely to use a correct strategy than older individuals, *OR* = 1.74, 95% CI [1.05, 4.89], $p = .036$, but not older constrained dyads, *OR* = 1.57, 95% CI [.79, 3.11, $p = .196$. Older constrained dyads and individuals demonstrated similar correct strategy use, *OR* = 1.22, 95% CI [.68, 3.07], $p = .337$. However, younger children demonstrated similar correct strategy use across conditions. In fact, younger individuals were marginally more likely to use a correct strategy than younger dyads, *OR* = 2.04, 95% CI [.88, 4.73], $p = .094$. In addition, older children were more likely to use a correct strategy than younger children in the dyad condition, *OR* = 3.56, 95% CI [1.76, 7.21], $p < .001$, and marginally more likely in the constrained dyad condition, *OR* = 1.97, 95% CI [.99, 3.93], $p = .055$. Older and younger individuals demonstrated similar correct strategy use, *OR* = .77, 95% CI [.31, 1.87], $p = .556$.

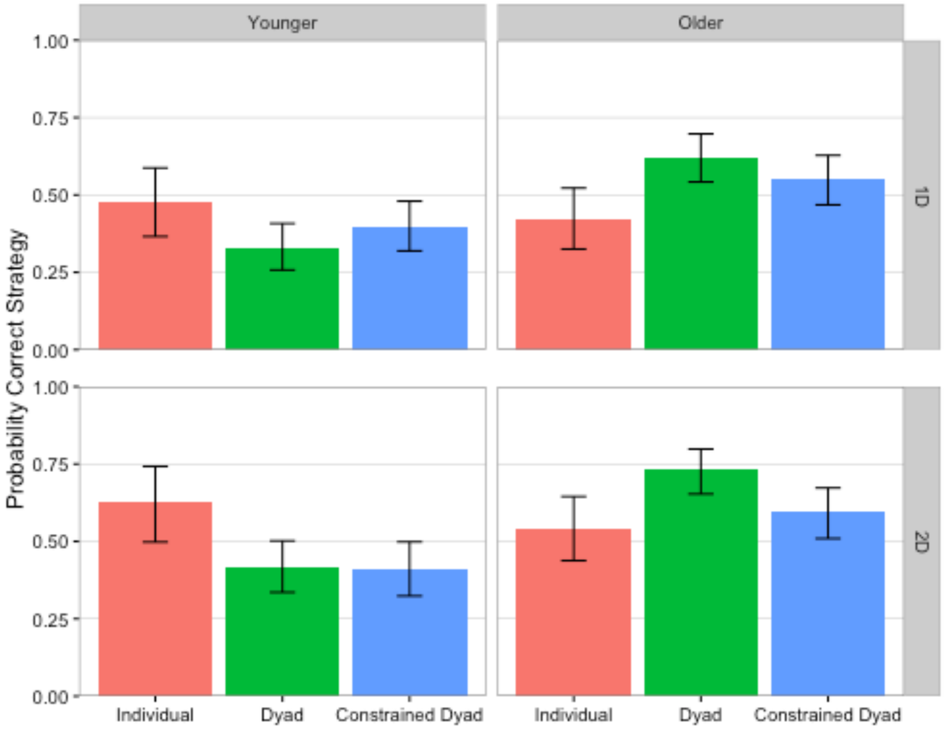

Figure 3-3. Estimated probabilities of correct strategy use by condition, age group, and category. Error bars represent ± SE.

Forced Choice Performance

Children's categorization accuracy and correct strategy outcomes suggest older dyads demonstrated superior learning compared to older individuals and younger dyads. However, these outcomes potentially depend on the spatial scaffolding of the test boards and memory for queried items. Does children's learning in the task generalize to novel items without spatial scaffolding?

To examine children's performance on the forced choice task, I estimated a linear mixed-effects model (LMM) on the number of rule-consistent choices (see Chapter 2) with Condition (*individual, dyad, constrained dyad*), Age Group (*younger, older*), Category (*1D, 2D*), Forced Choice Type (*direct contrast, diagonal contrast*) and their interactions as fixed effects, and a parsimonious by-participant, by-dyad, and by-rule random effects structure. Figure 3-5 shows children's performance was above chance across all factors, suggesting learning in the task was not limited to measures potentially influenced by spatial scaffolding.

There was an interaction between Condition and Age Group, LRT $X^2(2) = 6.14$, $p = .046$. Older children made a similar number of rule-consistent choices across conditions. However, younger individuals made more rule-consistent choices than younger dyads, $\beta = .52$, 95% CI [.05, .98], p = .031, and marginally more than younger constrained dyads, $\beta = 0.42$, 95% CI [-.06, .90], p = .090. In addition, older dyads made more rule-consistent choices than younger dyads, $\beta = .62$, 95% CI [.22, 1.02], $p = .003$. Individuals and constrained dyads were similar across ages.

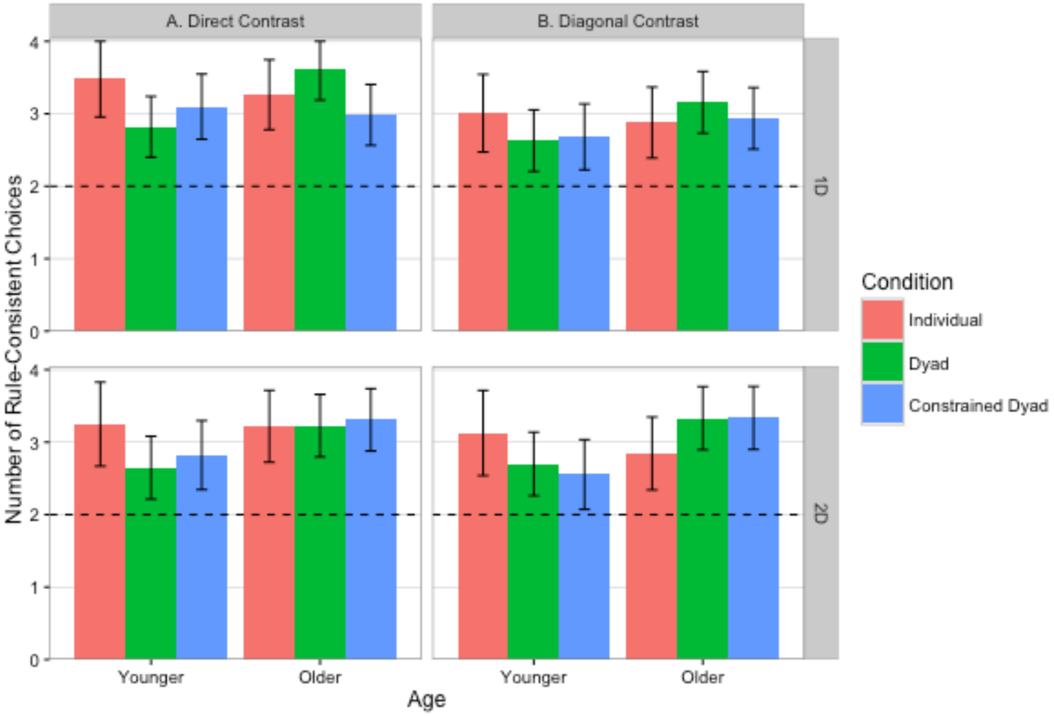

Figure 3-4. Number of rule-consistent choices by forced choice type, condition, age group, and category. Dotted lines represent chance performance and error bars represent 95% CIs.

Did Dyadic Learning Vary by Role?

Categorization accuracy. Here I examine whether role assignment affected children's categorization accuracy. I predicted children assigned to irrelevant dimensions in the 1D trial would demonstrate poorer learning that children assigned to relevant dimensions and that this effect would likely be greatest for younger children in the dyad condition. For the 2D trial, I did not expect role assignment in the previous trial to affect learning on the 2D trial.

1D trial. I estimated a binomial GLMM on children's correct categorization on the 1D trial with Condition (*dyad, constrained dyad*), Age Group (*younger, older*), Role (*relevant, irrelevant*), and their interactions as fixed effects, and a parsimonious by-participant, by-dyad, and by-rule random effects structure. As seen in Figure 3-5 (top row), there was a three-way interaction between Condition, Age Group, and Role, LRT $X^2(1) = 4.41$, $p = .036$. Older children were similarly accurate across conditions and roles. However, younger children assigned to the irrelevant dimension were unexpectedly more accurate than children assigned to the relevant dimension in the dyad condition, $OR = 2.53$, 95% CI [1.04, 6.17], $p = .041$, but not the constrained dyad condition, $OR = .67$, 95% CI [.28, 1.61], $p = .372$.

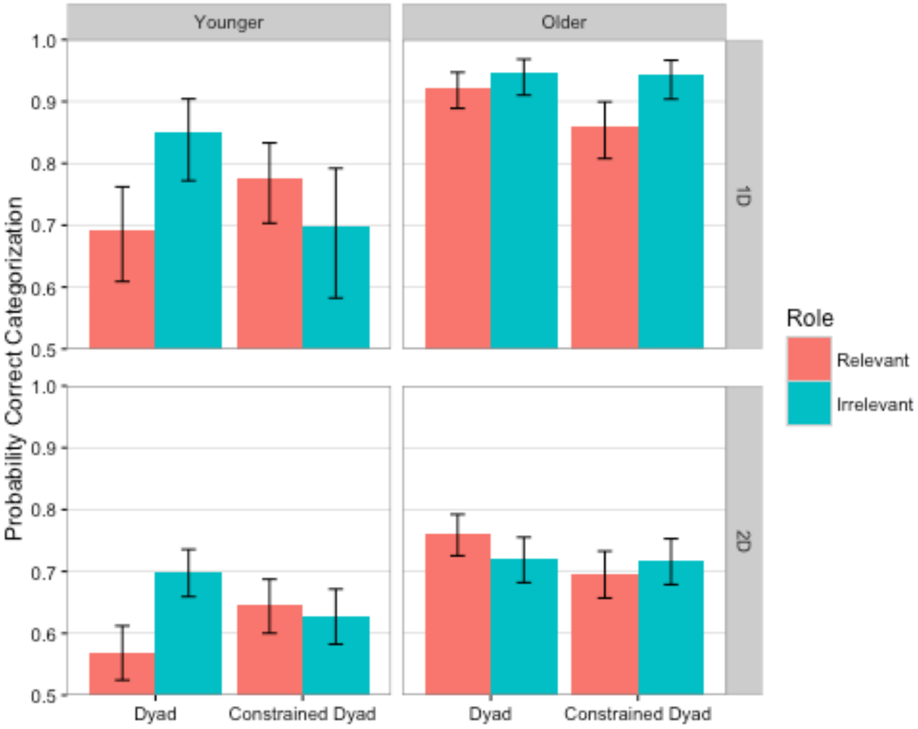

Figure 3-5. Estimated probabilities of correct categorization by condition, age group, and role for 1D and 2D trials. Error bars represent ± SE.

2D trial. I estimated a binomial GLMM on children's correct categorization on the 2D trial with Condition (*dyad, constrained dyad*), Age Group (*younger, older*), Role on the 1D trial (*relevant, irrelevant*), and their interactions as fixed effects, and a parsimonious by-participant, by-dyad, and by-rule random effects structure. As seen in Figure 3-5 (bottom row), there was a three-way interaction between Condition, Age Group, and Role, LRT $X^2(1) = 7.76$, $p = .005$. Older children were similarly accurate across conditions and roles. However, younger children assigned to the irrelevant dimension in the prior trial (i.e., the 1D trial) were unexpectedly more accurate than children assigned to the relevant dimension in the dyad condition, *OR* = 1.76, 95% CI [1.28, 2.41], $p < .001$, but not the constrained dyad condition, *OR* = .67, 95% CI [.28, 1.61], $p = .372$. The unexpected advantage of irrelevant role children in the dyad condition is possibly a carry over effect from their learning on 1D trial (see below).

Dyadic model-based strategy distributions. Here I examine whether role assignment affected children's model-based strategy use, including incorrect strategies. Table 3-1 summarizes children's best fitting strategies.

Table 3-1

Percentage of children best fit by model-based strategies by age group, condition, role, and category

		Younger Children					Older Children				
		Dyad		Constrained Dyad		Individual	Dyad		Constrained Dyad		Individual
		Relevant Partner	Irrelevant Partner	Relevant Partner	Irrelevant Partner		Relevant Partner	Irrelevant Partner	Relevant Partner	Irrelevant Partner	
1D Category	Correct Dimension	28.6	38.1	55	25	47.6	61.9	61.9	52.4	57.1	42.3
	Irrelevant Dimension	19	14.3	5	10	4.8	9.5	0	4.8	0	3.8
	2D	9.5	23.8	15	45	28.6	23.8	28.6	28.6	19	23.1
	Random	42.9	23.8	25	20	19	4.8	9.5	14.3	23.8	30.8
2D Category	Correct 2D	26.3	57.9	35.3	47.1	62.5	90	55	60	57.9	54.2
	Own Dimension	36.8	15.8	17.6	17.6	25	10	20	30	21.1	33.3
	Partner's Dimension	10.5	15.8	5.9	11.8		0	15	5	15.8	
	Random	26.3	10.5	41.2	23.5	12.5	0	10	5	5.3	12.5

Note: Individuals in the 2D trial were in charge of both dimensions, and thus the own/partner distinction of incorrect 1D strategies is not applicable.

I predicted children assigned to irrelevant dimensions in the 1D trial would be more likely to use irrelevant 1D and 2D strategies than children assigned to relevant dimensions. For the 2D trial, I expected children to be more likely to use a unidimensional strategy associated with their own dimension than a unidimensional strategy associated with their partner's dimension. I additionally expected both effects to be most pronounced in younger dyads.

1D trial. To examine whether role assignment affected older and younger dyads' rule learning on the 1D trial, I estimated a multinomial logistic regression on individuals' best fitting model-based strategy with Condition (*dyad, constrained dyad*), Age Group (*younger, older*), Role (*relevant, irrelevant*), and their interactions as predictors.

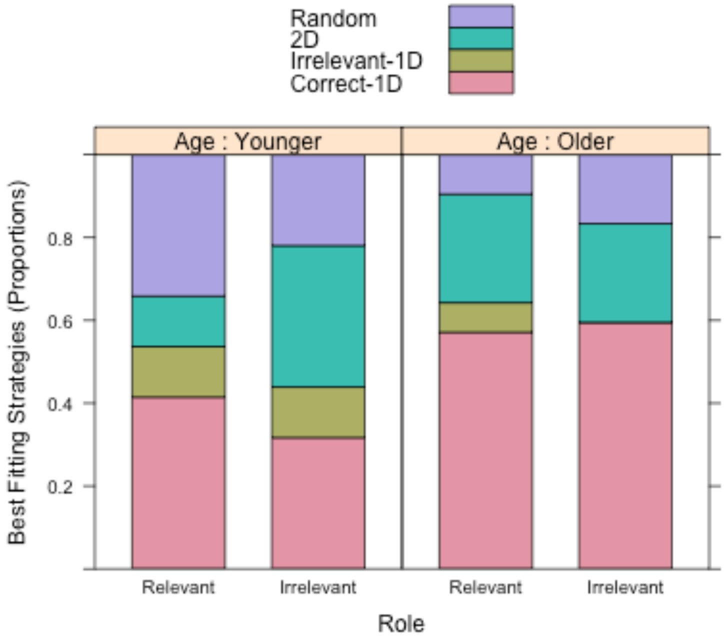

Figure 3-6. Best fitting model-based strategy by age group and role on the 1D trial.

The model yielded an interaction between Age Group and Role, LRT $X^2(3) = 8.29$, $p = .040$, and no effects involving Condition. As seen can be seen in Figure 3-6, older children's strategy distributions were similar across roles with one exception: irrelevant role partners never used the irrelevant-1D strategy. Estimation of this effect in the multinomial regression was unstable due to the empty cells, however a binomial test revealed relevant role partners' use of the irrelevant-1D strategy was similarly infrequent, $p = .250$.

Younger children showed a more robust strategy difference across role. More specifically, irrelevant role children were more likely to use a 2D strategy than relevant role children, *OR* = 4.17, 95% CI [1.24, 14.02], $p = .021$. However, irrelevant role children were no more likely than relevant role children to use the irrelevant-1D strategy, *OR* = 1.31, 95% CI [.32, 5.49], *p* = .714. Thus irrelevant role children integrated their own irrelevant dimension with the relevant dimension, as opposed to completely discounting the relevant dimension.

Finally, comparing children's strategy distributions across age groups revealed two differences. First, younger children assigned to the relevant dimension were more likely to use a random strategy than relevant role older children, *OR* = 5.89, 95% CI [1.44 24.10], *p* =.014. Second, older children assigned to the irrelevant dimension were more likely to use the correct 1D strategy than irrelevant role younger children, *OR* = 3.67, 95% CI [1.29,10.42], *p* = .015.

2D trial. To examine whether role assignment affected older and younger dyads' rule learning on the 2D trial, I estimated a multinomial logistic regression on individuals' best fitting model-based strategy with Condition (*dyad, constrained dyad*), Age Group (*younger, older*), Role on the 1D trial (*relevant, irrelevant*), and their interactions as predictors. The model yielded a marginal interaction between Age Group and Role, LRT $X^2(3) = 7.48$ *p* = .058, and no effects involving Condition. As seen can be seen in Figure 3-7, older children's strategy distributions were similar across roles. However, younger showed one strategy difference across role. Younger irrelevant role children were more likely to use a correct 2D strategy than relevant role children, *OR* = 3.45, 95% CI [1.01, 11.81], *p* = .048. This effect is consistent with the unexpected accuracy advantage younger irrelevant role children demonstrated in the 2D trial.

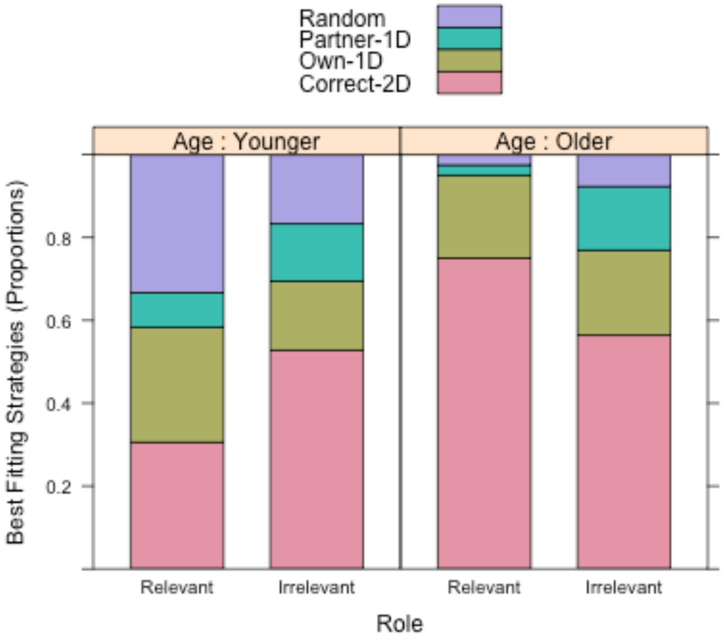

Figure 3-7. Best fitting model-based strategy by age group and role on the 2D trial.

Comparing children's strategy distributions across age groups revealed older and younger children assigned to the irrelevant dimension in the 1D trial had similar strategy distributions. However, older children assigned to the relevant role in the 1D trial were more likely to use the correct 2D strategy than younger relevant role children, *OR* = 7.24, 95% CI [2.40 ,21.89], *p* < .001, and less likely to use a random strategy, *OR* = .039, 95% CI [.003, .40], *p* = .007.

Finally, I conducted a planned contrast to test if children were more likely to use an incorrect 1D strategy corresponding to their own dimension or partner's dimension. Indeed, children were more likely to use a strategy focusing on their own dimension, *OR* = 4.50, 95% CI [1.24, 16.28], *p* = .022. This effect did not interact with age group, role, or condition.

Discussion

Children's Collaborative and Individual Learning Across Age Groups

A primary goal of the present study was to test whether early school-aged children's collaborative and individual learning differ in an active category learning task, in which children were responsible for coordinating their own information sampling and learning. As predicted, older dyads demonstrated better learning than older individuals and younger dyads. This was true for measures of overall categorization accuracy and correct strategy use. In addition, younger dyads did not demonstrate better learning than younger individuals. If anything, collaboration hindered younger childrens' learning (i.e., forced choice responsponding and correct strategy use). Interestingly, age group differences only existed for dyads; younger and older individuals demonstrated similar learning. These findings demonstrate collaboration can promote chlidren's active learning by age 7. Furthermore, these findings add to evidence suggesting children's abilities to profit from collaboration develop considerably in the early school years.

Children's performance in the constrained dyad condition was not as expected. Motivated by examples in the problem-solving literature suggesting action and temporal constraints can facilitate task learning, the turn-taking manipulation used here was meant to facilitate children's collaborative learning. However, children's learning in the constrained dyad condition failed to diffrentiate itself from dyad and individual learning in both age groups. In addition, age group differences for the constrained dyads were less robust than age group differences for dyads. One possible explanation for the manipulation's lack of effectiveness is that it simply failed to impose strong enough action and temporal constraints. Another possibility is that the turn-taking manipulation may have disrupted other productive components of children's collaborative interactions. For example, forced turn-taking may have altered children's information sampling behaviors and or communication. I suspect both factors are likely at work and revisit these possibilities in Chapters 4 and 5.

Dyadic Role Representation

Shared problem representations provide the foundation for effective collaborative learning. In the present task, I tested for asymmetric role representation (i.e., bias towards one's own role or actions) by evaluating the learning of children assigned to different roles. In both the 1D and 2D trials, older children in the dyadic conditions displayed similar categorization accuracy across roles. Furthermore, older children's categorization strategy use did not systematically vary by role. However, older children did display one instance of role asymmetry; those who used an incorrect unidimensional strategy in the 2D trial were more likely to categorize along their own dimension than partner's dimension. In general, older children seemed able to form and maintain a "bird's-eye view" representation of the task, thereby facilitating collaborative learning.

In contrast, younger children in the dyadic conditions displayed a number of role asymmetries. On the 1D trial, irrelevant role children were, as predicted, more likely to use an inaccurate 2D strategy than relevant role children (though no more likely to use the irrelevant-1D strategy). If irrelevant role children are biased towards their own action, they likely experienced difficulty learning their own role was non-diagnostic (e.g., Kushnir et al., 2009). As with older children, younger children who used an incorrect unidimensional strategy in the 2D trial were more like to categorize along their own dimension than partner's. I did not expect children's 1D trial role assignment to affect their 2D trial learning, however irrelevant role children were more likely to use the correct 2D strategy than relevant role children. Thus, it seems irrelevant role children's errant belief that both dimensions mattered on the 1D trial generalized to the 2D trial. The presence of this carry-over effect further suggests younger children experienced difficulty forming a shared "bird's- eye view" representation of the task. The result also provides a partial explanation for why irrelevant role children in the dyad condition had greater categorization accuracy than relevant role children in the 2D trial. However, the finding that irrelevant role children also had greater categorization accuracy in the 1D trial is difficult explain given these children were more likely to use an incorrect 2D strategy. This discrepancy highlights children's learning of the right dimensions and specifics of boundary locations can sometimes differ.

Effective Collaborative Learning

The present findings suggest early school-aged children's changing ability to profit from collaborative learning is partially underpinned by changing abilities to form shared agent-neutral representations. Specifically, younger children often operated with asymetric representations of the task, treating their own role and partner's role differently. This demonstration is noteworthy, as related results have been generated from tasks that lacked self-directed peer coordination (Kushnir et al., 2009; Riggs & Young, 2016; Young et al., under review).

The present findings also demonstrated collaboration improved learning over working alone in the older age group. Why might collaborative learning provide an advantage over individual learning? In the remainder of this dissertation I examine children's behaviors and interactions within the task and whether they explain the findings presented in the current chapter. In Chapter 4, I investigate whether children's information sampling behaviors vary across condition and age group. In addition, I examine the relations between these sampling behaviors and learning. In Chapter 5, I investigate potential developmental differences in dyads' verbal communication and how various types of talk may affect learning.

Chapter IV: Children's Information Sampling and Learning

Recent research on active category learning suggests adults and children, under certain circumstances, can select data in non-random and useful ways that provide advantages over passively receiving learning examples (Castro, Kalish, Nowak, Qian, Rogers, & Zhu, 2008; Markant & Gureckis, 2014; Markant, 2016; Sim et al, 2015). In the current research, collaboration benefitted learning in older children, but not younger children (see Chapter 3). To what extent might these effects be the result of variations in children's information sampling behaviors?

Background

Hypothesis-driven information sampling. Markant and Gureckis (2014) investigated the advantage of selection vs. reception (i.e., active vs. passive) category learning in terms of adults' information sampling behaviors. They found active selection of category examples led to superior learning of unidimensional boundaries (in a two-dimensional space) compared to passive reception of yoked active examples or random examples. They interpreted their findings in terms of hypothesis-driven sampling bias, suggesting active learners gather data that specifically test the hypotheses they have in mind. In their task, active learners systematically sampled closer to the true category boundary over time, as items further from the boundary became less informative (i.e., associated with less uncertainty) after testing. In addition, a measure of overall sample distance from the true boundary was negatively associated with active learners' categorization accuracy (i.e., closer distances predicted greater learning). However, sample distance was not associated with accuracy in passive learners yoked to identical learning examples, suggesting the active selection of examples is critical. In a similar task with category examples varying along a single dimension, 7-year-olds have behaved similarly (Sim et al., 2015). Specifically, children in an active selection condition systematically sampled closer to a criterion and showed an advantage over yoked reception learners. However, Sim and colleagues did test for (or report) an association between sample distance and accuracy.

Markant and Gureckis (2014) also found their effects depended on the complexity of categorization rule. When learning a two-dimensional rule involving integration across both dimensions (similar to the diagonal 2D rule used in the current study), active learners failed to sample increasingly closer to the true category boundary (in comparison to random sampling) and also showed little learning advantage over reception learners. Analyses revealed adults learning the 2D rule acted with biased expectations, generating samples and predictions that were best characterized by unidimensional hypotheses. Thus, it seems the advantage of hypothesis-driven sampling depends on the extent to which one's beliefs are consistent with the true concept. Indeed, in recent research Markant (2016) has manipulated adults' hypothesis generation to facilitate learning of two-dimensional rules and hinder learning of unidimensional rules.

Dyad coordination and sample complexity. Collaborators who fail to coordinate their actions and perspectives are unlikely to learn much in the present task. They may also be more likely to generate unplanned or incoherent sample distributions. In order to select an item for query, individuals in the task necessarily have to attend to each dimension. This is not the case for children in the dyadic conditions, as partners in poorly coordinated dyads can make selections on their dimension without regard to their partners'.

Spatial sampling distributions can be described using a variety of measures. Following the intuition that poorly planned or coordinated sampling should lead to sample distributions that are less coherent, a measure of sample complexity might prove useful. Mathematicians (Li & Vitányi, 2008) and psychologists (Griffiths & Tenenbaum, 2004) have recognized Kolmogorov-Chaitin, or algorithmic, complexity as a useful measure of complexity (or non-randomness). The algorithmic complexity of a pattern is defined as the length of the shortest computer program that produces the pattern and then halts. Patterns that can be described by a shorter procedure contain more structure and are less complex than patterns requiring longer descriptions.

Algorithmic complexity is uncomputable, but recent methods provide a practical way to approximate lower and upper bounds (Gauvrit, Zenil, Delahayne, & Soler-Toscano, 2014; Soler-Toscano, Zenil, Delahaye, & Gauvrit, 2013). Algorithmic complexity is increasingly being applied to measure structure in domains of human cognition and behavior, including memory, perception, and learning (see Gauvrit, Zenil, & Tegner, in press). For example, Kempe, Gauvrit, and Forsyth (2015) used algorithmic complexity to describe structure in children and adults' memory for random dot patterns on 10 x 10 grids. By way of analogy, I suggest using algorithmic complexity to characterize the amount of structure present in children's sampling patterns on the 7 x 7 game boards.

Research Questions and Predictions

Do younger and older dyads, constrained dyads, and individuals generate samples that differ in informativeness? Sampling close to the true category boundary is often an adaptive selection strategy, as items near the boundary should be associated with greater uncertainty over the course of learning than items far from the boundary. Thus, decreasing sample distance from the true category boundary over time and overall mean sample distance have been used as indicators of informative sampling behavior (Markant & Gureckis, 2014; Markant, 2016; Sim et al., 2015). If sample distance predicts children's learning in the present task, then meaningful variation in sampling distance across conditions and age groups might help explain their differential learning.

As prior research has only explored 7-year-old individuals' sampling in a unidimensional category space, this question is largely exploratory with respect to younger children, peer collaboration, and learning 1D and 2D rules in a two-dimensional category space. For the 1D category trial, I do not expect differences across condition, as a single child is responsible for selection along the diagnostic dimension in all three conditions. For the 2D trial, children in the dyadic conditions might have an advantage over individuals. Adults learning 2D rules fail to sample closer to the boundary because they are frequently biased by a unidimensional-hypothesis. If collaborating children are biased towards their own dimensions, dyads may be able to avoid biased sampling along a single unidimensional hypothesis. This might be particularly true of children in the constrained dyads, as turn taking limits a single child from directing sampling along their own dimension.

Does sample informativeness predict learning, and if so, is this relation similar across younger and older dyads, constrained dyads, and individuals? Markant and Gureckis (2014) found mean sample distance from true category boundaries was negatively correlated with active learners' overall categorization accuracy (i.e., close samplers learned more). Prior research on children's active category learning has not examined this association (e.g., Sim et al., 2015).

If sample distance is related to children's learning, variations in the strength of the relation might also explain differential learning across age groups and conditions, even in absence of underlying differences in sample informativeness. While I do not have strong predictions, research demonstrating collaboration can lead to greater task encoding, abstraction, and attention suggests collaboration could augment children's learning from evidence compared to individual learning (Pacherie. 2012; Schwartz, 1995; Shteynberg, 2015; Sommerville & Hammond, 2007). However, other research has shown poor collaboration can diminish learning from equivalent evidence (Young et al., under review). Thus it is unclear whether condition will interact with sample distance to predict learning.

One possibility I specifically explore concerns dyadic role assignment. Research with adults finds learners yoked to active learners' samples fail to demonstrate the relation between sample distance and categorization accuracy observed in the active learners (Markant & Gureckis, 2014). In the current research, irrelevant role children in the 1D trial are in essence yoked to their relevant role partner's selections, as they are not directly responsible for sampling along the diagnostic dimension. Thus, it may be the case that relevant role children demonstrate stronger sample distance to learning relations than irrelevant role children. I do not expect such an interaction for the 2D category trial, as both children selected along diagnostic dimensions.

Do younger and older dyads, constrained dyads, and individuals generate samples that differ in complexity? Measures of complexity have not yet been applied to active category sampling and learning. Algorithmic complexity can be use to describe the amount of structure (or non-randomness) present in a spatial sampling configuration. If poorly planned or coordinated samples have less structure, I predict younger dyads will have greater sample complexity than older dyads, as they likely have difficulties forming shared representations of learning tasks (Chapter 3; Kushnir, et al., 2009; Riggs & Young, 2016). To the extent that the constrained dyad turn-taking manipulation serves to facilitate shared representations, I also expect constrained dyads' sample complexity to be less than that of non-constrained dyads.

Does sample complexity predict learning, and if so, is this relation similar across younger and older dyads, constrained dyads, and individuals? If structure (i.e., non-randomness) in sample distributions is reflective of planned and coordinated sampling, I expect sample complexity to be negatively associated with learning (i.e., more structured samplers learn more). Since the application of sample complexity to active learning is novel, I have no predictions as to whether its relationship with learning will vary across age groups and conditions.

Methods

Sampling Coding

Informativeness. I calculated the orthogonal distance to the true category boundary for children's 14 samples on each trial (i.e., game). I used the mean of these sample distances as the primary measure of sample informativeness.

Complexity. I computed the algorithmic complexity of sampling patterns following methods described in Kempe, Gauvrit, and Forsyth (2015) and code provided by Gaurvrit (personal communication, January 23, 2016). The procedure breaks a 7 x 7 sampling grid into 4 x 4 tiles that overlapped by 1 row and 1 column. Zenil, Soler-Toscano, Delahaye, and Gauvrit (2012) previously computed algorithmic complexity estimates for all possible 4 x 4 patterns. The total pattern complexity (K) is then estimated using the following formula: formula: $\Sigma_p (\log_2(n_p) + K(p))$, where p denotes the different types of 4 x 4 patterns, n_p the number of occurrences of each axially and rotationally invariant pattern, and $K(p)$ the algorithmic complexity of p (Zenil, Soler-Toscano, Dingle, & Louis, 2014). See Figure 4-1 for examples of samples with varying complexity.

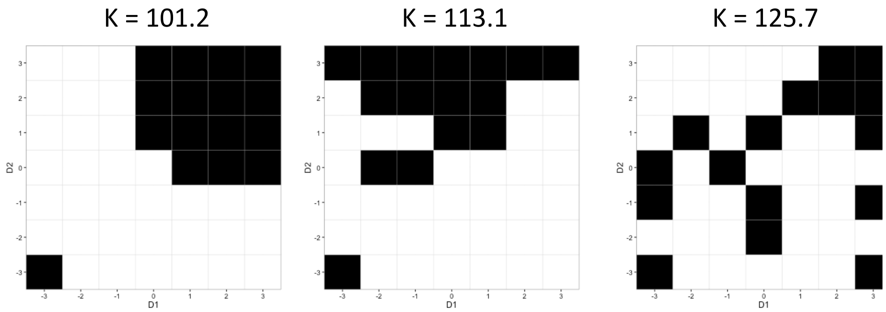

Figure 4-1. Example sampling distributions of low, medium, and high algorithmic complexity, *K*.

Results

Figure 4-2 displays children's aggregate sampling distribution as heatmaps. Notice, children tended to sample positive examples at above chance levels across age group, condition, and category.

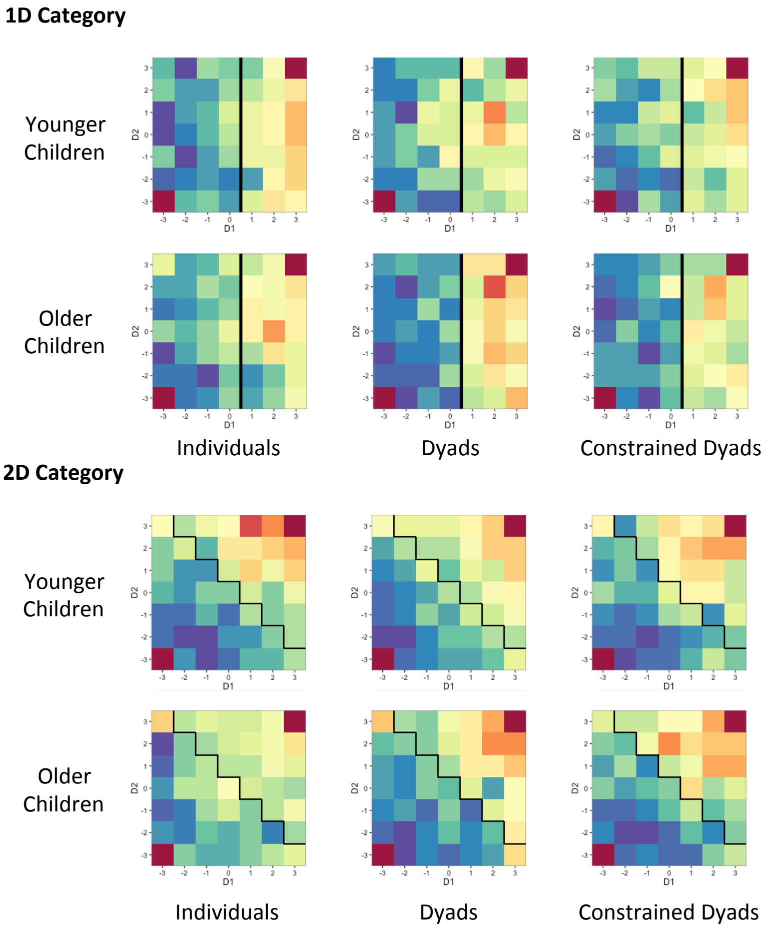

Figure 4-2. Heatmaps of children's aggregate sampling distributions by age group, condition, and category. Warmer colors represent frequent sampling (e.g., red corners were sampled 100% of the time by the experimenter) and cooler colors represent less frequent sampling. Black lines represent true category boundaries.

Did Children’s Sample Informativeness Vary by Condition and Age Group?

To examine potential variations in children’s sample informativeness, I estimated a linear mixed-effects model (LMM) on sample distance with Condition (*individual, dyad, constrained dyad*), Age Group (*younger, older*), Category (*1D, 2D*), and their interactions as fixed effects, and a parsimonious by-group random effects structure. (Note, that the by-group random effect individuates each group the produced a sample. Thus, dyads are treated as individual units.) There were no effects involving Age Group or Condition. Table 4-1 shows children’s sample distance was similar across ages and conditions. An effect of Category revealed sample distance was greater for the 1D category than the 2D category, LRT $X^2(1) = 11.37$, $p < .001$; however this can be attributed to the distance expected from random sampling for the 1D category (1.73) being greater than that of the 2D category (1.56). Descriptively, children sampled closer to the boundary than expected from random sampling in the 1D trial, but not the 2D trial.

Table 4-1

Mean sample distance from true category boundary and sample distance slope over time by age group, condition, and category

		Individual		Dyad		Constrained Dyad	
		Mean (*SE*)	Time β (*SE*)	Mean (*SE*)	Time β (*SE*)	Mean (*SE*)	Time β (*SE*)
1D Category	Younger Children	1.62 (.056) *	.006 (.014)	1.50 (.056) *	.026 (.013) •	1.61 (.058) *	.005 (.014)
	Older Children	1.59 (.051) *	-.016 (.012)	1.55 (.056) *	.001 (.014)	1.54 (.056) *	.016 (. 014)
2D Category	Younger Children	1.50 (.065)	-.003 (.016)	1.41 (.059) *	-.002 (.014)	1.46 (.026)	-.042 (.015) *
	Older Children	1.44 (.053) *	.001 (.013)	1.45 (.058)	-.021 (.014)	1.48 (.058)	.-.026 (014) •

Note: For means, stars denote sample distances smaller than those expected from random sampling (i.e., $p < .05$; circles denote $p < .06$). For time betas, stars represent non-zero slopes (i.e., $p < .05$; circles denote $p < .06$)

Sample distance might only be a useful measure of informative sampling after some learning has occurred (Markant & Gureckis, 2014). Therefore, I also examined sample distance in the second half of the learning phase (i.e., ignoring children's first 7 selections). I estimated a linear mixed-effects model (LMM) on 2nd-half sample distance with Condition (*individual, dyad, constrained dyad*), Age Group (*younger, older*), Category (*1D, 2D*), and their interactions as fixed effects, and a parsimonious by-group random effects structure. There were no effects involving Age Group or Condition. An effect of Category revealed 2nd-half sample distance was greater for the 1D category ($M = 1.61$) than the 2D category ($M = 1.45$), LRT $X^2(1) = 12.60$, $p < .001$.

To examine whether children sample closer to the boundary over time, I estimated a LMM on individual sample distance with Time (i.e., a continuous index of samples 1 through 14), Condition (*individual, dyad, constrained dyad*), Age Group (*younger, older*), Category (*1D, 2D*), and their interactions as fixed effects, and a parsimonious by-group random effects structure. There was a three-way Time by Condition by Category interaction, LRT $X^2(2) = 6.23$, $p = .044$. Because I am interested in whether children's samples became closer to the boundary over time, I examined this interaction by determining whether Time slopes were non-zero across factors. Only three slopes were different from zero (see Table 4-1). In the 1D trial, children's sampling distance did not change over time, with the exception of younger dyads that marginally sampled further from the true boundary over time ($p = .059$). In the 2D trial, older and younger constrained dyads sampled closer to the boundary over time ($p = .015$ and $p = .005$). However, neither older nor younger constrained dyads overall sample distance was less than that expected by random sampling.

In summary, children’s sample informativeness was largely similar across age groups and conditions.

Did Sample Informativeness Predict Children’s Learning?

Sample distance and categorization accuracy. I estimated a binomial GLMM on children’s correct categorization with Sample Distance, Condition (*individual, dyad, constrained dyad*), Age Group (*younger, older*), Category (*1D, 2D*), and their interactions as fixed effects, and a parsimonious by-participant, by-dyad, and by-rule random effects. Sample Distance did not interact with Condition, however there was a three-way Sample Distance by Age Group by Category interaction, LRT $X^2(1) = 10.46$, $p = .001$. As can been seen in Figure 4-3, sampling further from the 1D category boundary predicted greater accuracy in older children (logit $\beta = 2.09$, $SE = .27$, $p < .001$) and marginally greater accuracy in younger children (logit $\beta = .41$, $SE = .22$, $p = .062$). Sample distance did not predict accuracy in the 2D trial. Note, that the positive relation between sample distance and accuracy in the 1D trial stands in contrast to the motivating adult study (Markant & Gureckis, 2014), in which samples closer to the boundary predicted superior learning.

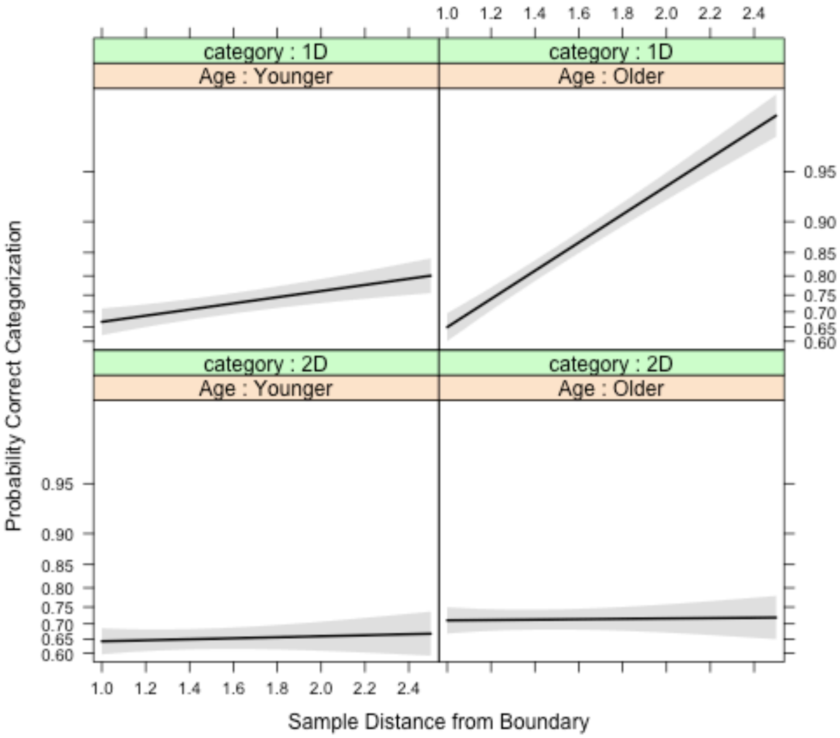

Figure 4-3. Estimated probabilities of correct categorization by sample distance, age group, and category. Shaded regions represent ± SE (pointwise).

To examine whether the informativeness of children's latter samples predicted accuracy, I estimated a binomial GLMM on children's correct categorization with 2nd-Half Sample Distance, Condition (*individual, dyad, constrained dyad*), Age Group (*younger, older*), Category (*1D, 2D*), and their interactions as fixed effects, and a parsimonious by-participant, by-dyad, and by-rule random effects. As can be seen in Figure 4-4, there was an interaction between 2nd-Half Sample Distance and Category, LRT $X^2(1) = 20.8$, $p < .001$, such that sampling further from the 1D category boundary predicted greater accuracy (logit $\beta = .37$, $SE = .13$, $p = .004$) and sampling further from the 2D category boundary predicted less accuracy (logit $\beta = .\text{-}31$, $SE = .13$, $p = .013$). There was also a 2nd-Half Sample Distance by Age Group , LRT $X^2(1) = 22.32$, $p < .001$. As can been seen in Figure 4-4, sampling further from the category boundary predicted greater accuracy in older children (logit $\beta = .41$, $SE = .14$, $p = .003$) and marginally lower accuracy in younger children (logit $\beta = -.20$, $SE = .11$, $p = .060$).

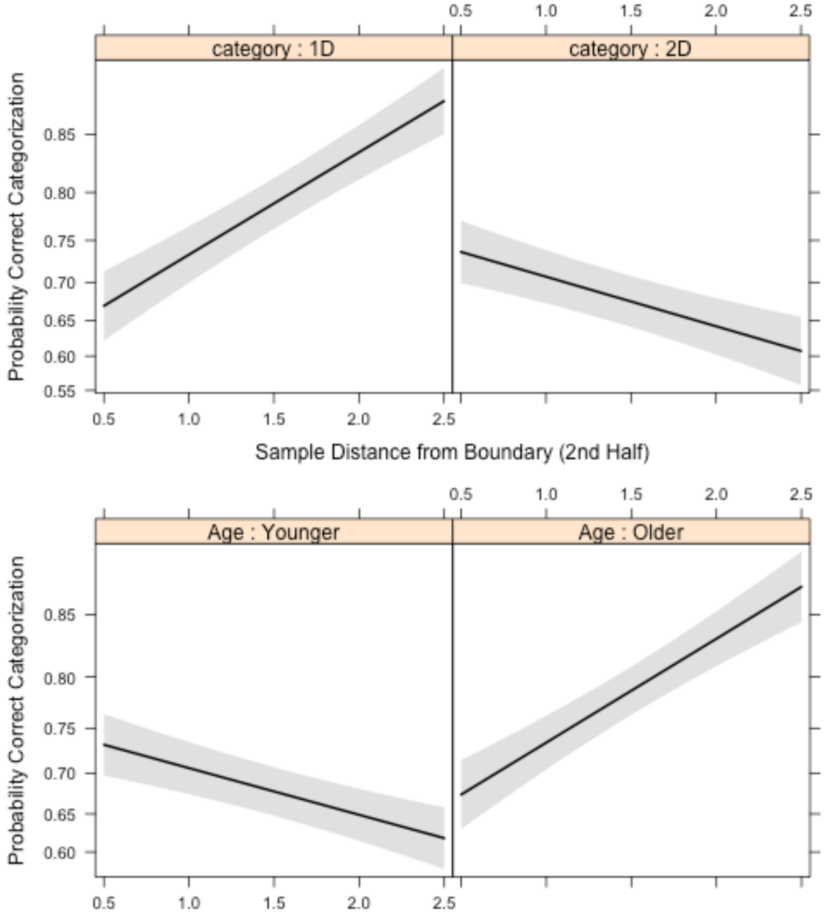

Figure 4-4. Estimated probabilities of correct categorization by 2nd-half sample distance and category (top) and by 2nd-half sample distance and age (bottom). Shaded regions represent ± SE (pointwise).

Sample distance and correct model-based strategy use. I estimated a binomial GLMM on children's correct strategy use (1 = correct strategy, 0 = other strategy) with Sample Distance, Condition (*individual, dyad, constrained dyad*), Age Group (*younger, older*), Category (*1D, 2D*), and their interactions as fixed effects, and a parsimonious by-participant, by-dyad, and by-rule random effects structure random effects structure. Sample Distance did not appear in interactions with Age Group, Condition, or Category. As can been seen in Figure 4-5, there was an overall positive effect of Sample Distance, LRT $X^2(1) = 3.98$, $p = .046$, such that children with greater sample distances were more likely to use a correct strategy across category trials (logit $\beta = 0.99$).

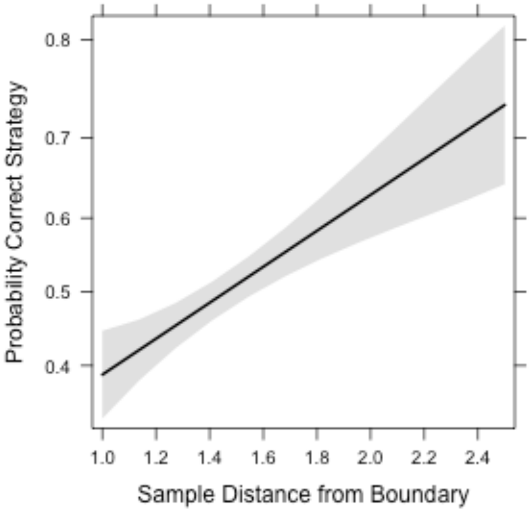

Figure 4-5. Estimated probabilities of correct strategy use by sample distance. Shaded regions represent ± SE (pointwise).

To examine whether the informativeness of children's latter samples predicted learning, I estimated a binomial GLMM on children's correct strategy use (1 = correct strategy, 0 = other strategy) with 2nd-Half Sample Distance, Condition (*individual, dyad, constrained dyad*), Age Group (*younger, older*), Category (*1D, 2D*), and their interactions as fixed effects, and a parsimonious by-participant, by-dyad, and by-rule random effects. Sample Distance in the second half only appeared in marginal interactions with Category, LRT $X^2(1) = 3.11$, $p = .078$, and Age Group, LRT $X^2(1) = 3.22$, $p = .073$ (see Figure 4-6). The relations mirrored those observed for categorization accuracy: sampling further from the boundary was associated with greater correct strategy use in the 1D trial and lower correct strategy use in the 2D trial, however neither slope was robust (*ps* > .10). Similarly, sampling further from the boundary was associated with greater correct strategy use in older children and lower correct strategy use in younger children, however neither slope was robust (*ps* > .10).

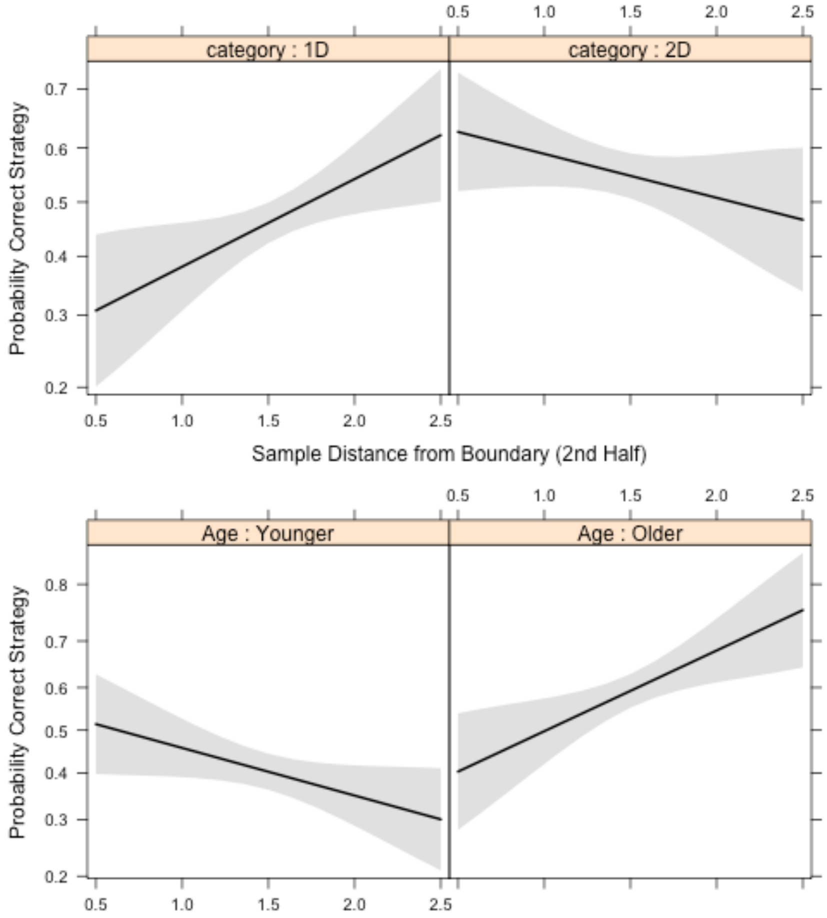

Figure 4-6 Estimated probabilities of correct strategy use by 2nd-half sample distance and category (top) and by 2nd-half sample distance and age (bottom). Shaded regions represent ± SE (pointwise).

Sample distance and dyadic role. To examine whether role assignment in the dyad and constrained dyad conditions interacted with sample distance on children's learning, I estimated binomial GLMMs on correct categorization and correct strategy use for both the 1D trial and 2D trials. Each full model included Sample Distance, Role (*relevant, irrelevant)*, Condition (*dyad, constrained dyad*), and Age Group (*younger, older*) and their interactions as fixed effects. Across the models, Role never appeared in interaction with Sample Distance. Thus, the relation between sample distance and learning did not depend on children's assignment to irrelevant or relevant dimensions.

Did Children's Sample Complexity Vary by Condition and Age Group?

To examine whether children's sample complexity varied across conditions and ages, I estimated a LMM on sample complexity with Condition (*individual, dyad, constrained dyad*), Age Group (*younger, older*), Category (*1D, 2D*), and their interactions as fixed effects, and a parsimonious by-group effects structure. As seen in Figure 4-7, there was a three-way interaction between Condition, Age Group and Category, LRT $X^2(2) = 8.33$, $p = .016$. In the 1D trial, older dyads had lower sample complexity than older constrained dyads, $\beta = -6.08$, 95% CI [-10.43, -1.74], $p = .007$, and marginally lower sample complexity than older individuals, $\beta = -3.53$, 95% CI [-7.66, .60], $p = .095$. However, younger dyads in the 1D trial had higher sample complexity than younger individuals, $\beta = 5.13$, 95% CI [.78, 9.48], $p = .022$. As predicted, younger dyads also had greater sample complexity than older dyads on the 1D trial, $\beta = 7.93$, 95% CI [3.58, 12.28], $p < .001$. Constrained dyads and individuals had similar sample complexity across age groups. In the 2D trial, sample complexity was similar across ages and conditions.

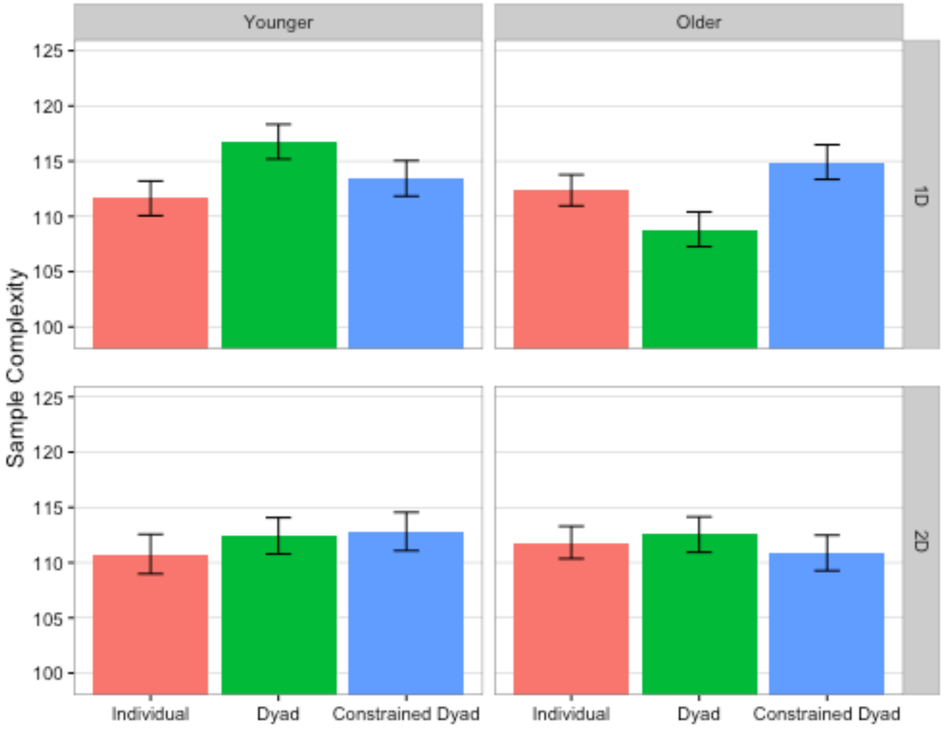

Figure 4-7. Mean algorithmic complexity of samples by age group, condition, and category. Error bars represent ± SE.

Did Sample Complexity Predict Children's Learning?

Sample complexity and categorization accuracy. I estimated a binomial GLMM on children's correct categorization with Sample Complexity, Condition (*individual, dyad, constrained dyad*), Age Group (*younger, older*), Category (*1D, 2D*), and their interactions as fixed effects, and a parsimonious by-participant, by-dyad, and by-rule random effects. Sample Complexity did not interact with Condition, however there was a three-way Sample Complexity by Age Group by Category interaction, LRT $X^2(1) = 4.94$, $p = .026$ (see Figure 4-8). In the 1D trial, more complex samples predicted lower accuracy for older children (logit $\beta = -.05$, $SE = .01$, $p < .001$). Younger children demonstrated a similar relationship, however the effect was not robust (logit $\beta = -.02$, $SE = .01$, $p = .166$). Sample complexity did not predict accuracy in the 2D trial.

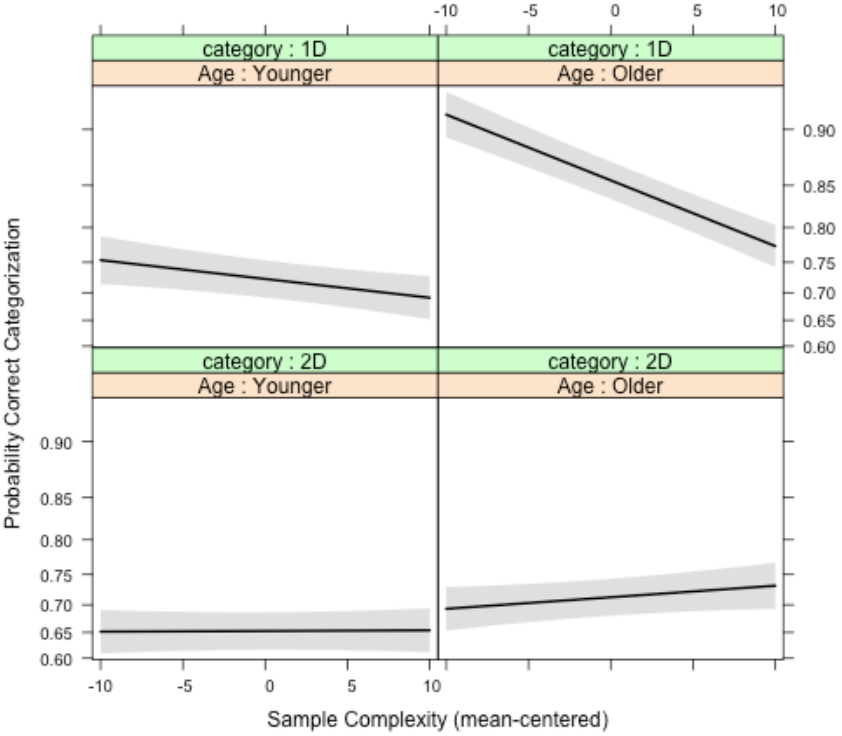

Figure 4-8. Estimated probabilities of correct categorization by sample complexity, age group, and category. Shaded regions represent ± SE (pointwise).

Sample complexity and correct model-based strategy use. I estimated a binomial GLMM on children's correct model-based strategy use with Sample Complexity (mean-centered), Condition (*individual, dyad, constrained dyad*), Age Group (*younger, older*), Category (*1D, 2D*), and their interactions as fixed effects, and a parsimonious by-participant, by-dyad, and by-rule random effects structure. As can be seen in Figure 4-9, there was an overall negative effect of Sample Complexity, LRT $X^2(1) = 5.23$, $p = .022$, such children with that lower sample complexity were more likely to use a correct strategy across category trials (logit $\beta = -.04$). Sample Complexity did not appear in any interactions with other factors.

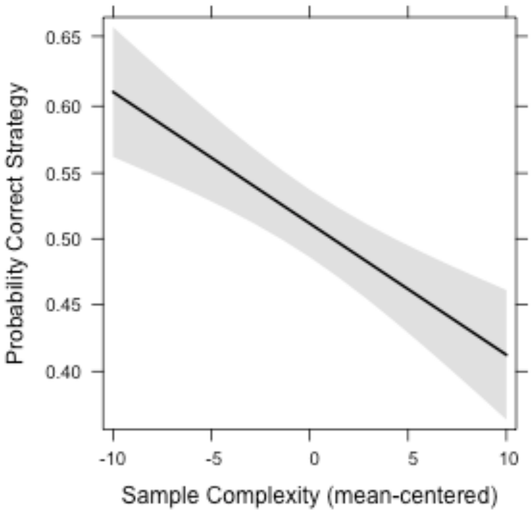

Figure 4-9. Estimated probabilities of correct strategy use by sample complexity. Shaded regions represent ± SE (pointwise).

Robustness of Age Group by Condition Interaction. Sample informativeness and sample complexity both explained variance in children's learning in the above models. However, the Age Group by Condition interaction remained robust in all of models (*ps* < .05).

Discussion

The primary goal of this chapter was to examine relations between children's information sampling behaviors and learning. In addition to generating novel results in the domain of children's active category learning, this research evaluated whether variations in children' information sampling could help explain why collaboration benefits older children, but not younger children.

Sample Informativeness and Learning

Using distance to true category boundaries as a measure of informativeness (e.g., Markant & Gureckis, 2014), I tested whether children's sampling behaviors varied across age group and condition. Children's total and 2^{nd}-half sample distances and tendency to sample closer to the boundary over time were similar across age groups and conditions. Thus, according to this metric, age differences in the effects of collaborative learning in the present task cannot be attributed to younger and older dyads generating more or less informative data. This finding bears some resemblance to studies of collaborative reasoning in scientific reasoning tasks, in which collaborating 4th and 5th graders tend to not generate more informative experiments than individuals (Azmitia & Montgomery, 1993; Teasley, 1995).

Unlike adults, children's sampling did not increasingly approach unidimensional category boundaries over time within the task. More striking is that the relation between children's sample distance and accuracy was frequently positive, as opposed to the negative relation shown in adults (Markant & Gureckis, 2014). Children, particularly the 7 and 8-year-olds, who sampled further from the category boundary were more accurate and likely to use correct strategies. There are a number of differences between the present child task and the adult task used by Markant and Gureckis that might contribute to this reversal of the relationship between sample distance and accuracy. For example, children in our task sampled 14 (16 with experimenter selections) items without replacement on a visible 7 x 7 grid over the course of a single learning block. In contrast, adults sampled 128 items with replacement on a non-visible 16 x 16 grid over the course of 8 learning blocks. Future research could explore whether size of category space, number of samples with respect to size of category space, record of queried items, and number of training blocks affect relations between information sampling and learning.

One difference I suspect played a role was the use of a preference category (i.e., does Dobo like the toy or not) rather than a membership category (e.g., does the antenna pick up channel A or B, Markant & Gureckis, 2014). While use of preference categories can sometimes facilitate children's statistical and causal inference (Sobel & Kirkham, 2012), the fact that Dobo liked certain toys may have biased children towards selecting them. To counteract this possibility, task instructions repeatedly suggested children needed to learn about toys Dobo likes and dislikes to be successful. However, children's above-chance sampling of positive examples suggests these instructions may not have been effective. Of course, children need to know something about a category rule in order to systematically oversample positive examples (e.g., if Dobo likes big toys, let's sample the biggest toys). I suspect this is a primary explanation for why children who sampled further from the boundary demonstrated greater learning

Finally, it is worth noting that measures or characterizations other than sample distance might lead to different conclusions about children's sampling informativeness and learning. For example, the unexpected sample distance relation suggests positive (or negative) sampling might be an indicator of children's hypothesis driven sampling. Additionally, in other work on active learning (Castro et al., 2008; Markant & Gureckis, 2014; Sim et al., 2015), researchers have used passive learners yoked to active learners' examples to disentangle effects of selection decision-making and informativeness. Future research might yoke children (or adults) to the selection data in the present task to better understand whether younger and older dyads, constrained dyads, and individuals generated similarly informative data.

Sample Complexity and Learning

Using algorithmic complexity (Gauvrit et al., in press; Kempe et al, 2015) as a measure of structure in active category sampling is a novel contribution of the present research. I proposed the measure could be used to characterize planful and well-coordinated sampling. Accordingly, this measure differentiated children's sampling in the 1D learning trial, revealing younger dyads generated samples that were more complex than younger individuals and older dyads. These results support the interpretation that unsuccessful collaboration in younger children might arise from difficulties with coordinating actions and perspectives into shared task representations. Furthermore, older dyads generated samples that were less complex than older constrained dyads and marginally less complex than older individuals. The difference between dyads and constrained dyads further suggests the constrained dyad turn-taking manipulation, meant to facilitate joint action plans, was ineffective. Finally, the marginal difference in sample complexity between older dyads and individuals potentially reflects a mechanism of collaborative learning highlighted in older children and adults, namely that collaboration can lead to the generation of more advance problem-solving and reasoning strategies (e.g., Laughlin, 2011). More advanced and coherent strategies might generate samples with more underlying structure (i.e., less complexity).

There was a negative relationship between the algorithmic complexity of children's samples and children's learning. Children, particularly 7 and 8-year-olds, who generated more complex samples demonstrated poorer learning of unidimensional category boundaries. This result, along with the observed age differences in sample complexity, highlights a pathway by which older dyads were more effective learners than younger dyads.

It is important to note that sample complexity itself is unlikely a driver of learning. Rather, processes and behaviors that lead to the generation of less complex samples likely also lead to successful learning. A following prediction is that behaviors that facilitate coordination of joint actions and understandings, such as verbal planning or hypothesis generation, should demonstrate a negative relationship with sample complexity (see Chapter 5).

Information Sampling and Learning: Varying Effects

As the present study is among the first to explore 5- to 8-year-olds' active category learning, there are several findings of note that are peripheral to primary questions of collaborative learning. First, similar to adults, relations between children's information sampling and learning were largely restricted to the 1D trial. In the 2D trial, sample informativeness and complexity were generally not associated with categorization accuracy, sample distances were not closer to the true boundary than expected by random sampling, and sample complexity did not vary across conditions or age groups. Together, these results suggest children's information sampling was poorly matched to learning a diagonal 2D category. Whether an adult-like unidimensional hypothesis sampling bias can explain children's information sampling on the 2D trial is a question for further study.

Second, the relations between children's information sampling and accuracy were considerably stronger in older children (7 and 8-year-olds) than younger children (5 and 6-year-olds). This was true across conditions for both sample informativeness (total and 2nd half) and sample complexity. This pattern suggests the early school-aged years might be a time of important developments in children's active learning.

Finally, both of the above results seem to be qualified by outcome measure. That is, sample distance and sample complexity predicted children's correct model-based strategy use for the 1D trial, but also the 2D trial (unlike categorization accuracy). Studies with adults have used decision boundary models to assess overall condition differences in correct categorization strategy use (Markant & Gureckis, 2014; Markant, 2016), however they have not on reported on relations between information sampling and strategy use as the present research does. Thus, at present it is unclear how to reconcile discrepancies between outcome measures in this research. Future work should clarify the correspondences between categorization accuracy and strategy use in this data set and others.

Summary

Does information sampling help explain the collaborative learning effects observed in this research? The present findings suggest children's information sampling behaviors are sometimes (i.e., on the 1D trial) related to learning and sometimes vary across age group and condition (i.e., older dyads produce less complex samples than younger dyads). Information sampling behaviors explained additional variance in learning beyond the age group by condition interaction reported in Chapter 3, but never enough shared variance to supplant it. Given the current analyses, variations in children's information sampling behaviors are important, but non-primary, mechanisms of the collaborative learning in the current research.

Chapter VI: The Role of Talk in Children's Collaborative Active Learning

Research on successful collaborative learning has stressed the importance of verbal communication that occurs during joint problem-solving (Baker, 2015). Since collaborative learning is thought to emerge from the construction and maintenance of shared problem representations, talk is considered a fundamental resource for coordinating actions and understandings (Roschelle & Teasley, 1995; Tomasello, 2014). More successful collaborators engage in more explanation (e.g., Howe, 2009, Okada & Simon, 1997; Teasley, 1995), argumentation, clarification, and negotiation (e.g., Kuhn, 2015; Kruger & Tomasello, 1986), and discussion of plans, goals, and strategies (e.g., Perlmutter, Behrend, Kuo, & Muller, 1989; Roschelle & Teasley, 1995; Schwarz et al., 2000).

Research on children's collaboration and verbal communication has mostly been focused on later school-aged and adolescent children for two primary reasons. First, a number of the cognitive-linguistic processes of interest are indeed advanced and late developing (e.g., formal argumentation). Second, there has been a general belief that younger children lack the verbal abilities to begin and sustain peer dialogue that supports collaborative learning (e.g., Azmitia, 1988; Flynn, 2010; Piaget, 1977; Rogoff, 2003). Since effective peer collaboration and communication are thought to emerge around 7 years of age (Tomasello, et al., 1993), it is possible that the developmental differences I observed in the benefits of collaboration are due to the emergence of more advanced peer dialogue. Therefore, in this chapter I examine relations between characteristics of dyadic talk and patterns of learning in the task. I consider whether potential age group differences in talk may contribute to explaining the observed developmental difference in collaborative learning.

Background

I base my approach on a study of 4th graders' collaborative learning in a scientific discovery task carried out by Teasley (1995). The task involved individuals and dyads actively generating experiments to test hypotheses about the functioning of a novel machine (i.e., how a "mystery key" worked on spaceship). In addition, Teasley assigned individuals and dyads to talk and no-talk conditions, in order to distinguish the importance of talk from working with a partner. Talking dyads learned more than no-talk dyads and talking individuals. Additionally, Teasley found that 1) overall production of talk was positively associated with performance, 2) verbal planning and talk about evidence (i.e., generating and evaluating predictions and hypotheses) were positively associated with performance, and 3) talking dyads generated more planning and evidence talk than talking individuals. The importance of verbal planning and evidence interpretation to collaborative learning has been demonstrated in other active hypothesis testing tasks with school-aged children and adults (Azmitia & Montgomery, 1993; Howe, 2009; Schwarz et al., 2000). The current study affords an opportunity to examine the extent to which the 5 to 8-year-olds spontaneously generate verbal plans and discuss evidence, and to investigate whether these aspects of talk promote learning in this task.

Recent research on preschool-aged children's collaborative problem-solving suggests peers engage in verbal planning to a greater extent than previously appreciated. For example, Warneken and colleagues found 5-year-olds can prospectively coordinate tool choices by announcing their own action or suggesting what choice their partner should make (Warneken et al., 2014). Additionally, 5 year-olds can present and discuss arguments while making joint decisions with peers (Köymen, Rosenbluam, & Tomasello, 2014). Five year-olds can also modify the informativeness of decision justifications according to the common ground they share with a partner (Köymen, Mammen, & Tomasello, 2016).

Recent research has also shown young children generate and learn from verbal hypotheses and explanations of evidence (Legare & Clegg, 2015). For example, 5 and 6-year-olds learn more when partners generate alternative hypotheses (Young, Alibali, & Kalish, 2012), though it is unclear whether younger peers engage in explicit hypothesis evaluation (Azmitia, 1988). Furthermore, preschool-aged will spontaneously generate explanations in response to evidence that is inconsistent with their prior beliefs (Legare, Wellman, & Gelman, 2010).

Together, these results suggest it is likely that even younger children in our task will engage in and show some benefit from verbal planning and evidence interpretation. However, a primary question for present purposes is whether older dyads generate more or different kinds of talk than younger dyads.

Talk in constrained dyads and individuals. This chapter considers children's talk in the dyad condition, and not the constrained dyad and individual conditions. Preliminary investigations of talk in the constrained dyad condition suggested the turn-taking manipulation altered children's talk in a number of ways. First, the experimenter spoke more frequently in the constrained dyad condition than the dyad and individual conditions, as the experimenter frequently needed to remind children of turn order. Second, children in the constrained dyad condition spoke less frequently than children in the dyad condition. Whether this was the result of greater experimenter participation and interruption or due to the turn-taking manipulation itself is question for further investigation. Additionally, although task instructions encouraged individuals to talk during the learning phase, a majority of individuals did not produce much talk.

In contrast, Teasley (1995) was able to evaluate the effects individuals' talk because they were required to do so.

Research Questions & Predictions

Following the analytical approach of Teasley (1995), this chapter addresses the following questions.

Are there differences between older and younger dyads' talk? Given prior work suggesting effective peer communication emerges around 7 years of age, I expect older dyads will produce more productive talk than younger dyads. I specifically test whether total amount of talk, planning talk, and talk about evidence (including hypothesis generation) vary by age group.

Does the production of talk affect learning? If so, do planning and evidence talk predict learning? In line with prior work suggesting talk is a primary mechanism of collaborative learning, I expect amount of dyad talk to be positively related to learning. Furthermore, as found by Teasley (1995), I expect verbal planning and talk about evidence interpretation (including hypotheses generation) will be positively related to learning.

What are the relations between particular types of talk and information sampling behaviors? Since children's information sampling behaviors co-occurred with their verbal interactions, there are several possible relations I explore. First, since sample informativeness (i.e., distance to boundary) is driven by hypotheses in adults (Markant & Gureckis, 2014), I examine whether there is a relation between sample distance and evidence talk. In adults, I expect the relation would be negative (i.e., the more one verbally interprets evidence the closer one should sample to the boundary). However, since better learners in the present task sampled further from the boundary, the relationship might be positive. Second, in Chapter 4 I proposed children's sample complexity might be an index of poorly coordinated collaborative action during evidence selection. Since verbal plans function to coordinate joint action, I expect children who produce more planning talk will generate less complex samples.

Methods

Coding Talk

Dyads' conversations in the experimental phase were transcribed verbatim and separated into conversational turns. A conversational turn was defined as any segment of time consisting of a single child's utterances (including sounds, sentence fragments, or complete sentences). A new conversational turn began when the other partner spoke or when the experimenter spoke to direct the task. For a a given game (i.e., trial), the total number of child conversational turns was counted as a measure of amount of talk.

Each conversational turn was coded for the presence of five verbal categories derived from the literature: plans, strategies, predictions, hypotheses, and evidence evaluations. Since a given conversational turn could contain multiple categories of talk, coding of these categories was not mutually exclusive. Table 5-1 describes and provides examples of the coding scheme.

Following Teasley (1995), two additional general categories of talk were created. Plans and strategies were counted as planning talk. Predictions, hypotheses, and evidence evaluations were counted as evidence talk. Finally, conversational turns coded as hypotheses were additionally evaluated for content in order to determine the total number of unique hypotheses dyads generated over the course of a game.

Table 5-1

Verbal Coding Categories, Descriptions, and Interrater Reliabilities

Type	Description	Examples	Cohen's κ
Planning Talk	A plan or strategy.	(See Below)	.88
Plans	Explicit statement of one's own selection, suggesting a partner's selection, or suggesting a specific item.	"I want to try this one", "Do this size" "Let's keep going down the row"	.87
Strategies	A plan that is accompanied by a justification or explanation. Note, justifications could refer to evidence.	"I want to double check here" "He didn't like that one, so let's try this"	.82
Evidence Talk	A prediction, hypothesis, or evaluation.	(See Below)	.90
Predictions	Explicit guess or statement of knowing about a specific toy prior to testing.	"I think he's going to like it" "No, he won't like that one"	.80
Hypotheses	Explicit statement about the general rule, kind, or region Dobo likes.	"He likes long arms" "Okay, I think he likes ones with spikes"	.83
Evaluations	Statements about previous evidence or discussion (evaluation) of hypotheses.	"See, he didn't like those" "Yeah, but look at this, they all got shorter arms" "That doesn't work cause this one"	.63

Reliability Coding

A second coder, who was not involved in data collection or verbal transcription, independently coded the transcripts of 9 randomly selected dyads after being trained on an additional transcript. Thus, approximately 20% of the data were coded for reliability. Cohen's Kappa for each verbal category is presented in Table 5-1.

Data Preparation

Because I am interested in how planning and evidence talk guide children's information sampling and learning from evidence, I analyzed the verbal data with respect to children's selections, rather than examining total frequencies or proportion of talk (Teasley, 1995). Specifically, for each of a given dyad's 14 selections, I examined whether at least one statement of each verbal category occurred before Dobo was queried. Thus, for each type of talk, each dyad received a score of 0 to 14, reflecting the number of selections in which children produced a specific type of talk. This method has an additional advantage of making dyads that spoke with varying frequency more comparable.

Results

Differences in Dyad Talk

To examine whether older and younger dyads generated different amounts of talk, both in terms of total number of turns and in terms of each of the coded categories , I estimated linear-mixed models (LMMs) on each talk variable with Age Group (*younger, older*), Category (*1D, 2D*), and their interaction as fixed effects, and a by-dyad random intercept. Descriptive statistics and model results are summarized in Table 5-2. Because the verbal category data might more appropriately be considered count data, I also performed analogous Poisson regressions. Results of the two methods were largely consistent. Below I summarize the observed age and category differences.

Table 5-2
Verbal Descriptive Statistics and Model Summaries

	Younger Children		Older Children		**Age Effect**	**Category Effect**	**Age X Category Interaction**
	Mean (*SD*)	Min., Max.	Mean (*SD*)	Min., Max.			
Planning Talk	4.43 (3.74)	0, 13	5.73 (4.49)	0, 13	$p = .044$	-	-
Plans	3.90 (3.17)	0, 12	5.17 (4.15)	0, 13	$p = .039$	-	-
Strategies	.93 (1.58)	0, 6	1.02 (1.56)	0, 6	-	-	-
Evidence Talk	2.80 (3.05)	0, 11	3.05 (2.98)	0, 10	-	-	-
Predictions	.50 (1.28)	0, 5	.95 (1.32)	0, 4	$p = .081$	$p = .076$	-
Hypotheses	1.68 (2.07)	0, 9	1.80 (2.15)	0, 7	-	-	-
Evaluations	1.28 (1.72)	0, 5	1.24 (1.41)	0, 5	-	-	-
Amount of Talk	33.11 (24.61)	0, 95	29.85 (28.63)	0, 132	-	$p = .009$	-
Unique Hypotheses	1.28 (1.45)	0, 7	1.80 (1.76)	0, 5	-	$p < .001$.	-

Note: Talk categories could take scores 0 through 14. All LMM fixed effects with *ps* < .10 are reported.

Age differences. As summarized in Table 5-2, older dyads produced more verbal plans ($\beta = 1.62$, $p = .039$) and more overall planning talk ($\beta = 1.62$, $p = .044$) than younger dyads. The age difference in planning talk was largely driven by verbal plans, as both groups produced a similarly low number of verbal strategies.

Older and younger dyads produced similar numbers of conversational turns, similar amounts of evidence talk (i.e., predictions, hypotheses, and evaluations), and similar numbers of alternative hypotheses. Thus, older and younger dyads' talk was largely similar, with the exception of older dyads producing more verbal planning.

Category differences. Dyads spoke less in the 2D trial than in the 1D trial ($B = -5.87$), perhaps because they had become accustomed to the task. Dyads also generated more unique hypotheses in the 2D trial than the 1D trial ($B = .80$). There are two likely reasons for this finding. First, children in the 1D trial converged on a correct hypothesis more easily. Second, the complexity of the 2D rule allowed for a greater number of potentially plausible hypotheses.

Relations Between Dyad Talk and Learning

Amount of talk. Did dyads that talked more learn more? To examine this question I estimated binomial GLMM models on children's categorization accuracy and correct strategy use with Amount of Talk (i.e., total number of conversational turns), Age Group (*younger, older*), Category, and their interactions as fixed effects, and by-dyad random intercepts. Amount of Talk did not interact with Age Group, however both models yielded Amount of Talk by Category interactions (Figure 5-1). Amount of talk demonstrated a stronger positive relationship with category accuracy in the 1D trial than the 2D trial, LRT $X^2(1) = 6.24$, $p = .012$. Similarly, amount of talk demonstrated a positive relationship with correct strategy use in the 1D trial, but not the 2D trial, LRT $X^2(1) = 5.03$, $p = .025$. Thus as expected, amount of talk was a positive predictor of children's learning, but primarily for the 1D trial.

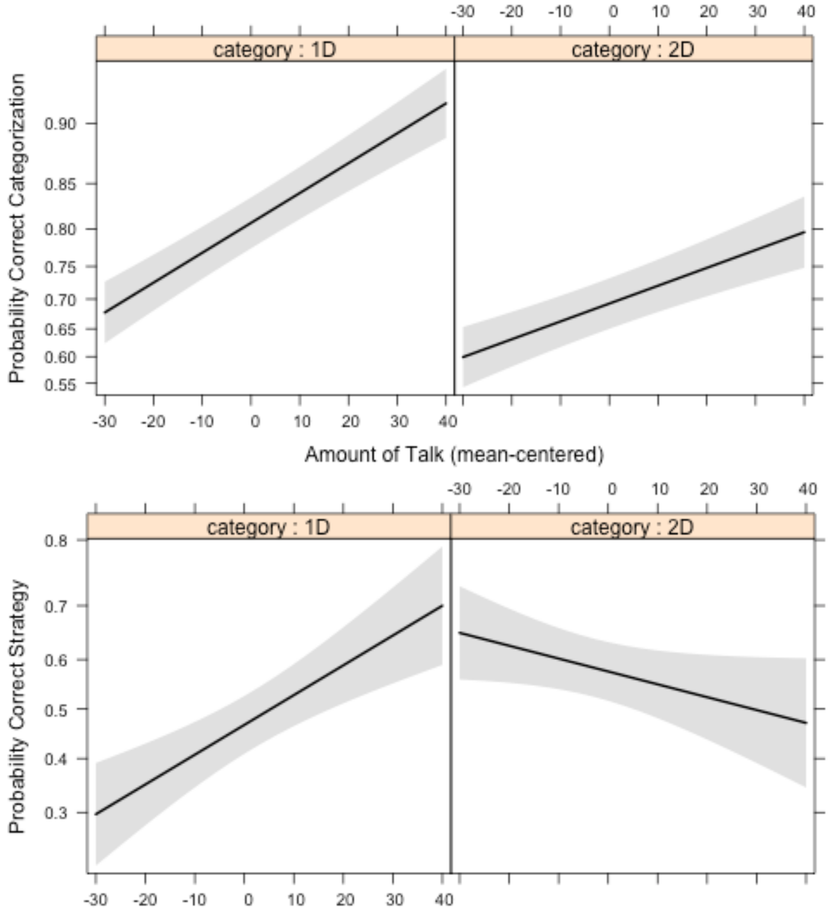

Figure 5-1. Estimated probabilities of correct categorization (top) and strategy use (bottom) by amount of talk and category. Shaded regions represent ± SE (pointwise).

Planning talk. Did dyads that engaged in more planning communication learn more? To examine this question I investigated dyads' overall verbal planning (i.e., not plans and strategies separately). I estimated binomial GLMM models on children's categorization accuracy and correct strategy use with Planning Talk, Age Group (*younger, older*), Category, and their interactions as fixed effects, and by-dyad random intercepts. Planning Talk did not interact with Age Group, however both models yielded Planning Talk by Category interactions (see Figure 5-

2). Planning Talk demonstrated a stronger positive relationship with category accuracy in the 1D trial than the 2D trial, LRT $X^2(1) = 11.95$, $p < 001$. Similarly, planning talk demonstrated a positive relationship with correct strategy use in the 1D trial, but not the 2D trial, LRT $X^2(1) = 6.01$, $p = .014$. Thus as expected, planning talk was a positive predictor of children's learning, but primarily for the 1D trial.

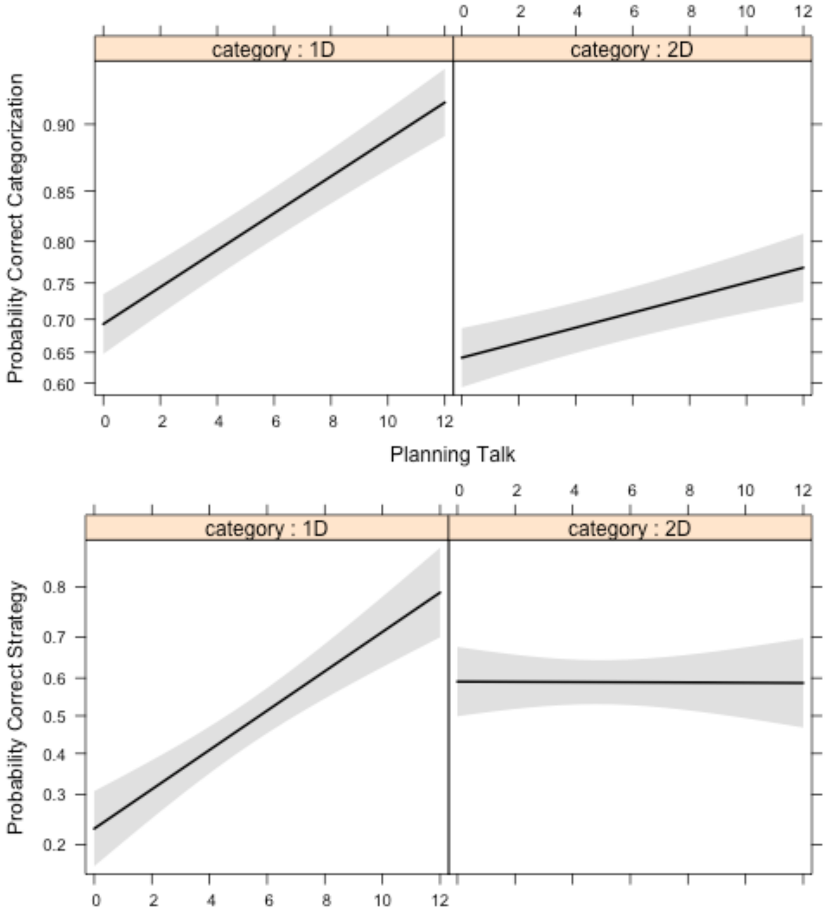

Figure 5-2. Estimated probabilities of correct categorization (top) and strategy use (bottom) by planning talk and category. Shaded regions represent ± SE (pointwise).

Evidence talk. Did discussion of evidence predict dyad learning? To examine this question I investigated dyads' overall evidence talk (i.e., not predictions, hypotheses, and evaluations separately). I estimated binomial GLMM models on children's categorization accuracy and correct strategy use with Evidence Talk (i.e., total number of conversational turns), Age Group (*younger, older*), Category, and their interactions as fixed effects, and by-dyad random intercepts. For the correct categorization model, there was a three-way Evidence Talk by Age Group by Category interaction, LRT $X^2(1) = 19.92$, $p < .001$ (see Figure 5-3). In the 1D trial, talk about evidence demonstrated a stronger positive relationship with categorization accuracy in older dyads (logit $\beta = .38$, $SE = .07$, $p < .001$), than younger dyads (logit $\beta = .16$, $SE = .04$, $p < .001$). In the 2D trial, talk about evidence did no predict older dyads' category accuracy (logit $\beta = .004$, $SE = .03$, $p = .88$), however younger dyads' evidence talk did predict categorization accuracy (logit $\beta = .11$, $SE = .04$, $p = .003$). The age difference in the 2D trial seems to be largely driven by poor learning of younger dyads that didn't talk about evidence.

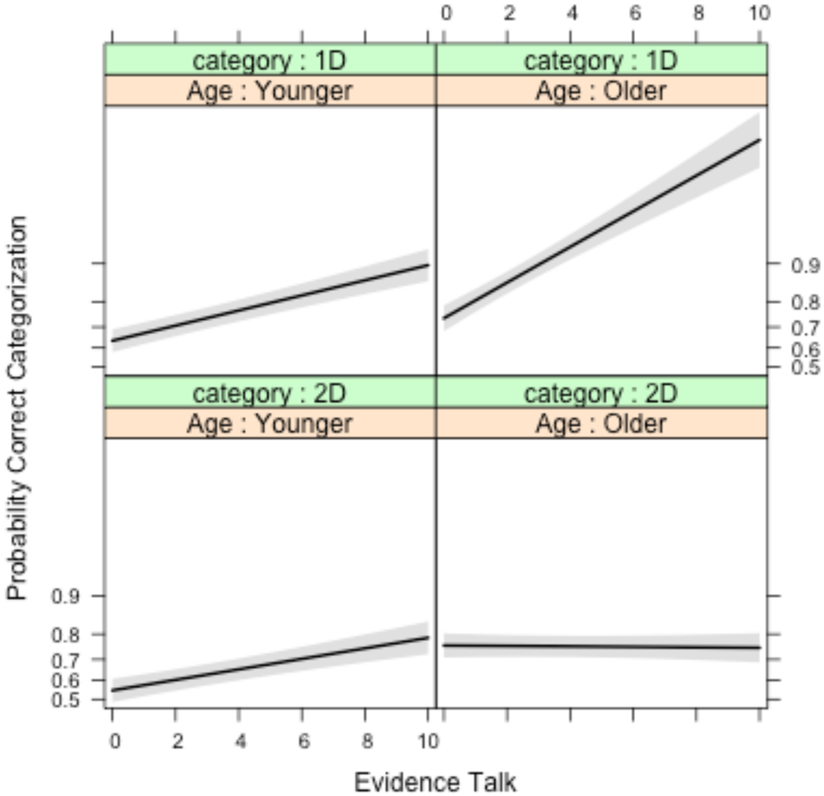

Figure 5-3. Estimated probabilities of correct categorization by evidence talk, age group, and category. Shaded regions represent ± SE (pointwise).

For the correct strategy use model, Evidence Talk did not interact with Age Group, however there was an Evidence Talk by Category interaction, LRT $X^2(1) = 4.94$, $p = .026$ (see

Figure 5-4). Evidence talk demonstrated a positive relationship with correct strategy use in the 1D trial (logit $\beta = .41$, $SE = .11$, $p < .001$), but not the 2D trial (logit $\beta = -.06$, $SE = .08$, $p = .46$), Thus as expected, discussing evidence and hypotheses was a positive predictor of children's learning, but only for the 1D trial.

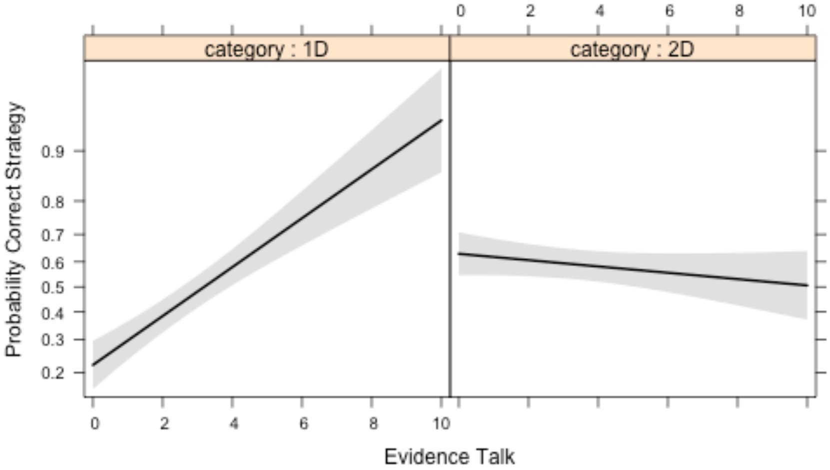

Figure 5-4. Estimated probabilities of correct strategy use by evidence talk, age group, and category. Shaded regions represent ± SE (pointwise)

Unique hypotheses. Generation of alternative hypotheses is a critical component of evidence talk. Did dyads that generated a larger number of explicit verbal hypotheses demonstrate better learning? I estimated binomial GLMM models on children's categorization accuracy and correct strategy use with Alternative Hypotheses (i.e., total number of unique verbal hypotheses turns), Age Group (*younger, older*), Category, and their interactions as fixed effects, and by-dyad random intercepts. Alternative Hypotheses did not interact with Age Group, however both models yielded Alternative Hypotheses by Category interactions (see Figure 5-5). Generation of verbal alternative hypotheses demonstrated a positive relationship with category accuracy in the 1D trial, but not the 2D trial, LRT $X^2(1) = 28.44$, $p < .001$. Similarly, verbal alternative hypotheses showed a positive relationship with correct strategy use in the 1D trial, but not the 2D trial, LRT $X^2(1) = 8.62$, $p = .003$. Thus as expected, generation of alternative hypotheses was a positive predictor of children's learning, but only for the 1D trial.

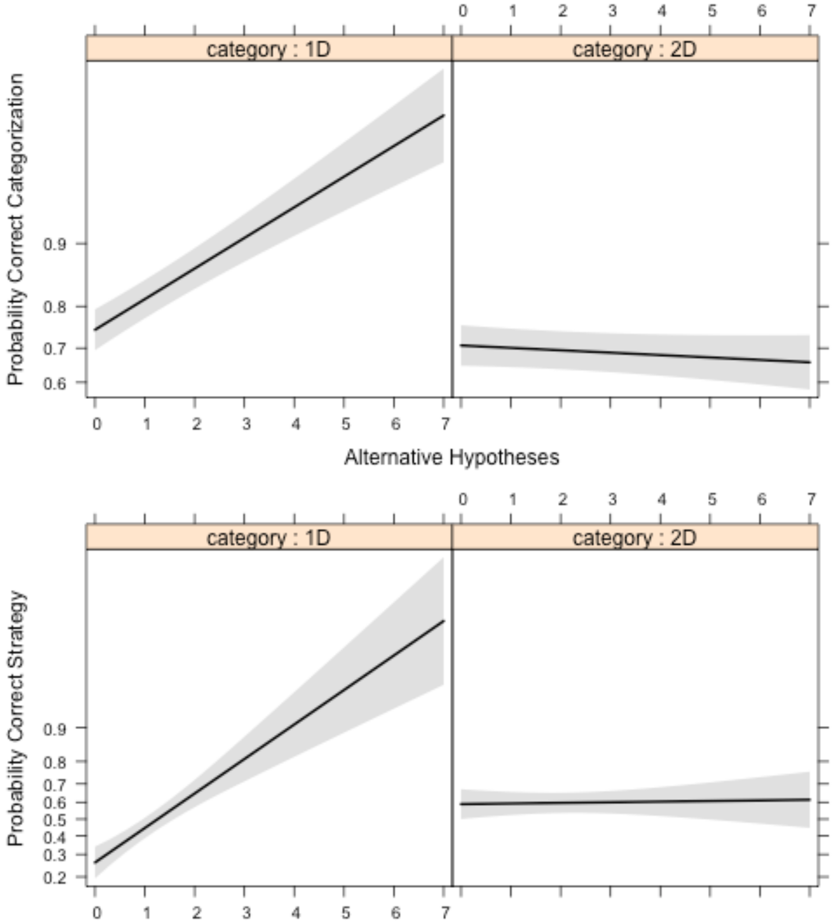

Figure 5-5. Estimated probabilities of correct categorization (top) and strategy use (bottom) by number of alternative hypotheses and category. Shaded regions represent ± SE (pointwise).

Summary. In general, the components of talk examined above were positively related to dyad learning as expected in the 1D category trial. However, talk, like children's sampling behaviors, was far less predictive of children's learning in the 2D trial. In addition, while talk explained additional variance in the above models, Age Group always remained a robust predictor of learning (*ps* < .05).

Relations Between Talk and Information Sampling

Table 5-3 presents correlations between planning talk, evidence talk, sample distance, and sample complexity. Because children's sampling and talk behaviors vary by category, I considered the 1D and 2D trials separately. I expected planning talk to be negatively correlated with sample complexity and evidence talk to be negatively correlated with sample distance. In the 1D trial, neither relation was apparent. In the 2D trial, planning talk demonstrated a marginal positive relation with sample complexity ($r = .19$, $p = .09$) and evidence talk demonstrated a negative relation with sample distance ($r = -0.27$, $p = .016$).In addition, planning talk and evidence talk were positively associated for both the 1D ($r =.55$, $p < .001$) and 2D trials ($r =.54$, $p < .001$). This is not unexpected, as dyads that generate one type of productive talk should be more likely to produce other types of productive talk.

Table 5-3
Spearman Correlations Between Talk and Sampling Behaviors

1D Category	1	2	3	4.
1. Planning Talk	-			
2. Evidence Talk	.55***	-		
3. Sample Distance	.00	.12	-	
4. Sample Complexity	.01	-.18	-.26*	-
2D Category	1	2	3	4
1. Planning Talk	-			
2. Evidence Talk	.54***	-		
3. Sample Distance	-.07	-.27*	-	
4. Sample Complexity	.19•	-.01	-.26*	-

Note: •$p < .10$, *$p < .05$, **$p < .01$, ***$p < .001$

Discussion

The primary goals of this chapter were to assess 1) if dyad talk was associated with learning and 2) whether there were age group differences in dyad talk that could help explain age group differences in collaborative learning.

Dyad Talk and Learning

In the present task, dyads that produced more talk demonstrated better learning. This was the case for the total amount of talk dyads produced, the number of selections that were accompanied by verbal plans, the number of selections that were accompanied by verbal interpretations of evidence, and the number of unique hypotheses that were generated. These data are consistent with a large body of research suggesting dyad talk is a powerful mechanism for collaborative learning (Teasley, 1995; Azmitia & Montgomery, 1993; Howe, 2009).

It is important to note that the relations between talk and learning held for the 1D trial, not the 2D trial. Like children's information sampling behaviors, it seems children's verbal communication was poorly matched to learning a diagonal 2D boundary. Prior research suggests such boundaries are often unexpected and difficult verbalize (Ashby & Maddox, 2005; Markant & Gureckis, 2014). Thus, children's talk might only be productive to the extent that it is consistent with a correct and easily verbalized category.

A potential mechanism for the development of collaborative learning is the development of peer communication. Thus, I tested for differences in younger and older dyads' production of specific types of talk. Younger and older dyads produced a similar overall amount of talk, verbal interpretation of evidence, and number of unique hypotheses. However, older dyads engaged in more verbal planning than younger dyads. This result, along with the positive relation between planning and learning, highlights a pathway by which older dyads are more effective learners than younger dyads.

The present results do demonstrate 5 and 6-year-olds spontaneously produced and benefitted from planning and evidence talk, despite prior work suggesting children of these ages might fail to do so (Azmitia, 1988; Tomasello et al., 1993; Flynn, 2010). Of course, planning and evidence talk can vary in quality, not just quantity. The stronger relation between evidence talk and learning in older dyads possibly suggests they generated “better” evidence talk than younger dyads. Future work should examine the content of children’s predictions, hypotheses, and evidence evaluations to better understand this finding.

I expected lack of verbal planning to be related to more complex sampling. However, dyads’ planning talk and sample complexity were not negatively correlated, and in fact showed a marginal positive relation in the 2D trial. This inconsistency raises a number of questions. First, it may be the case that sample complexity should be conceptualized as something different than unplanned and poorly coordinated sampling. Second, it highlights the specificity of the present coding of children’s talk. I have not yet distinguished between which partner is talking, whether suggested plans are “good” or “bad”, or whether plans are actually followed. Furthermore, I have not yet considered the impact of children’s frequent non-verbal communication in the task (e.g., gesture, points, gaze following). Future research could explore children’s talk at a finer grain, with consideration of the role of non-verbal communication, and could more fully examine the dynamics between children’s information sampling and coordinating talk (see Chapter 6).

Summary

Does talk help explain the collaborative learning effects observed in this research? The present findings suggest dyads’ talk is related to learning (at least in the 1D trial) and that increased use of verbal planning explained additional variance in dyad learning beyond the effect of age. This research found many dyads of both age groups spontaneously coordinated their actions and understandings in the task by producing plans and interpretations of evidence. Those dyads that engaged in more of these efforts demonstrated the greatest learning.

Chapter VI:

General Discussion

This dissertation investigated the development of children's collaborative learning skills to better understand when and how collaboration is effective for learning. An assumption in the extant literature is that effective peer collaborative learning emerges around 7 years of age (Flynn, 2010; Tomasello et al., 1993). However, recent research demonstrates preschool-aged children successfully coordinate collaborative problem solving (e.g., Fletcher et al, 2012; Warneken et al., 2014). To connect these lines of research, this research examined 5 to 8-year-olds' collaborative and individual learning. By using an active category learning task, I was able to examine (1) whether collaboration improved learning in younger (5 and 6-year-old) and older (7 and 8-year-old) children and (2) how skills for coordinating shared representations, gathering information, and communicating develop and contribute to effective collaborative learning.

Summary of Empirical and Theoretical Contributions

The results regarding learning in this study were very clear. Collaboration improved older children's learning over individual learning. However, collaboration did not improve, and sometimes hindered, younger children's learning. Importantly, younger and older individuals demonstrated equivalent learning in the task. The largest learning differences were consistently found between older and younger collaborators. These findings are in line with prior research on early school-aged children's collaborative learning (Young et al., under review) and theoretical analyses of the development of collaborative learning (Tomasello et al., 1993; Tomasello et al., 2005). Furthermore, the findings demonstrate children's abilities to profit from collaboration develop considerably in the early school years.

Development of Collaborative Learning Skills

What accounts for the shift in younger and older children's collaborative learning? I investigated whether developmental changes in children's abilities to coordinate shared representations (Chapter 3), information gathering actions (Chapter 4), and verbal communication (Chapter 5) underlied the age difference in collaborative learning. Notably, I found all three coordination behaviors underwent development and contributed to successful collaborative learning. First, I found younger dyads' learning was undermined by asymmetric role representations. That is, younger children assigned to different roles demonstrated systematically different learning, suggesting they failed to represent their own and their partner's actions during the task from an agent-neutral "bird's-eye view". Older dyads in the task demonstrated far less role asymmetry. Second, I found older and younger dyads' information sampling differed in terms of structure and complexity. When learning a unidimensional boundary, older dyads generated the least complex (i.e., most structured) samples and younger dyads generated the most complex samples, in comparison to individuals. Furthermore, sample complexity was negatively associated with successful learning. That is, children who generated overly complex and uncoordinated sampling patterns tended to learn less. Third, I found older dyads produced more talk about action plans than younger dyads. Importantly, planning talk was positively associated with learning.

Together these results suggest younger dyads had considerable difficulty in coordinating joint action toward a shared learning goal, relative to older dyads. This extends previous research on children's collaborative learning under circumstances of minimal coordination demands (Kushnir et al., 2009; Young et al., under review) and suggests granting young children control over their own collaborative activity is not sufficient to promote learning. Unpacking the relations between the developmental differences observed in this study, and children's collaborative skills more generally, is a complex undertaking of considerable theoretical importance (Flynn, 2010). Figure 6-1 summarizes the relations between 1D and 2D category learning and dyad behaviors found in the current research.

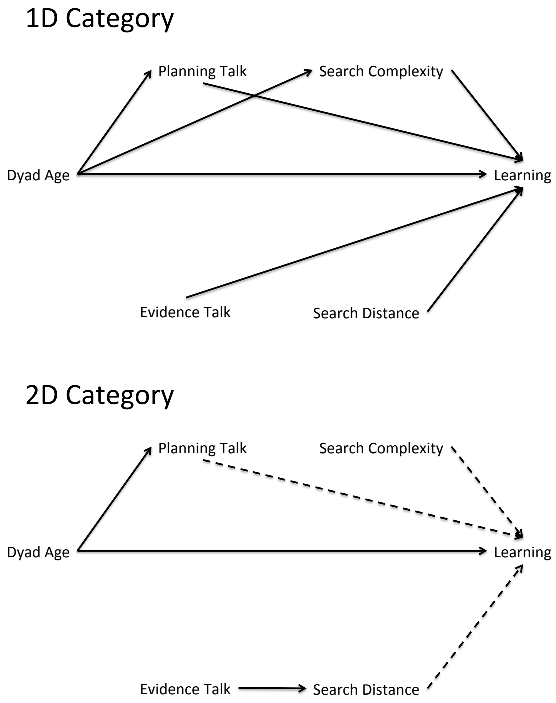

Figure 6-1. Summary of observed relations in dyads' category learning. Dashed lines represent relations that held for a single outcome measure (i.e., categorization accuracy

Unexpectedly, the present research found children's talk was not associated with their information sampling behaviors (at least on the 1D trial). Similarly, the relations between information sampling and learning were consistent across irrelevant and relevant role partners (Note, role assignment is not represented in the figure). These findings suggest the developmental differences observed in the study may stem from different underlying processes.

An important next step in this research is to statistically model the relations between children's coordination behaviors and learning more directly. The present analytical approach found role assignment, information sampling (sample distance and complexity), and talk (planning and evidence) all explained additional variance beyond the effect of age, but never enough shared variance to supplant the effect of age. Thus, while each mattered when modeled separately, none of the behaviors functioned as a primary mechanism of the age difference in collaborative learning. However, it is possible that these factors contribute in a mutually supportive and multidimensional way. Future analyses should formally test for latent relations between children's coordination behaviors and mediation effects on learning (e.g., by using multi-group SEM).

Mechanisms of Collaborative Learning

In the current research, collaborating 7 and 8-year-olds demonstrated greater learning than individuals. Consistent with a large body of work on collaborative learning, I found dyad communication played a large role in successful learning. Collaborators in the present task coordinated their learning by verbally signaling intentions (i.e., plans and strategies) and current interpretations/expectations about the evidence. Dyads that produced more talk about action plans and evidence learned more about the underlying task structure. Prior research suggests it is not just the amount of talk that dyads produce, but also the nature of dyad talk (compared to individual talk) that contributes to learning (e.g., Howe, 2009; Schwarz et al., 2000). Although the present task did not require individuals to talk or some dyads to not talk (e.g., Teasley, 1995), one could compare dyads and individuals that spontaneously talked to those that did not in the current data.

A second mechanism of collaborative learning explored in this dissertation was children's information gathering behaviors, as collaborators may generate different data than individuals. Prior research has explored this possibility in school-aged and adolescent students' scientific reasoning and found collaborators often fail to generate more informative experiments than individuals (e.g. Azmitia & Montgomery, 1993; Teasley, 1995). In line with this research, I found older dyads and individuals produced similarly informative samples, according to a distance to boundary metric (e.g. Markant & Gureckis, 2014). A methodological contribution of the present dissertation was the application of a measure of pattern complexity to children's samples. According to the metric of algorithmic complexity (Kempe et al., 2015), older dyads produced more structured (i.e., less random) samples on the 1D trial than older individuals. This raises the question of how and why dyads generated less complex samples. One plausible hypothesis is that dyad coordination via verbal plans and hypotheses led to more structured sampling. However, the present data suggest dyad talk was independent of sample complexity.

Another possibility and direction for further study is that the division of labor in the present task provided dyads a structural advantage. That is, biases in partners' individual actions and beliefs may combine to generate joint information sampling that is structurally different than individual sampling. I explored this possibility in dyads' sampling of the 2D boundary, predicting they would be less likely to generate a sampling pattern biased by a single unidimensional hypothesis (as is the case with individual adults) because each partner might act according to their own hypothesis-driven sampling bias. This prediction did not hold; that is, dyads did not demonstrate differential information sampling in the 2D trial. However, a similar interactive process could be responsible for older dyads' sampling in the 1D trial. In the future, research should examine the extent to which structural components of collaboration may influence information gathering and further advantage (or disadvantage) collaborative learning. Potentially fruitful possibilities include examining children's sampling under different collaboration policies (e.g., no assignment to dimensions), after manipulating of prior beliefs (e.g., Markant, 2016), and with alternative category structures.

Supporting Collaborative Learning

The developmental progression of collaborative learning observed in this research and other studies (e.g., Riggs & Young, 2016; Young et al., under review) raised the possibility of introducing an intervention to support younger children's collaborative learning. Drawing upon examples of temporal and action constraints facilitating preschool-aged children's learning within problem-solving tasks (e.g., Fletcher et al., 2012; Warneken et al., 2014), the turn-taking manipulation employed in the constrained dyad condition was meant to facilitate younger children's coordination of actions and formation of a "bird's-eye view" representation. The manipulation was partially successful, to the extent that younger constrained dyads demonstrated less role asymmetry than their non-constrained counterparts. However, the manipulation ultimately failed to improve learning in younger children.

Why was the manipulation unsuccessful? Interestingly, older children in the constrained dyad condition on some measures demonstrated poorer performance than their non-constrained counterparts. For example, they generated more complex samples and produced less coordinating talk. It seems older children did not use the turn-taking periods as opportunities to attend and adjust to a partner's actions. Rather, the turn-taking manipulation focused children on coordinating turn-taking instead of action plans and understandings. This possibility can be further examined by coding constrained dyads' verbal transcripts and joint attention behaviors.

Children's Active Learning

This dissertation used an active category-learning task because it was particularly well suited to examine primary research questions about children's collaborative learning and coordination. As a consequence, the current research also contributes a number of novel findings to the nascent literature on children's active category learning. First, as mentioned above, characterizing children's sampling patterns in terms of complexity predicted active category learning. Second, unlike adults, children in the present task demonstrated a positive relationship between sample distance and learning. These results highlight the need for more research on what measures capture better or worse sampling for particular tasks, and possibly, age groups. Additionally, the present research did not include comparisons between children's active learning and passive learning. To better understand the consequences of the sampling decisions individuals and dyads made in the present study, further research should yoke passive learning to their selection data (e.g., Castro et al., 2008; Markant & Gureckis, 2014; Sim et al., 2015).

Children's active learning in the present task varied by category structure and outcome measure. Using categorization accuracy as a measure of learning, children's information sampling behaviors were predictive of learning a 1D category, but not a 2D category. In contrast, children's information sampling behaviors predicted correct strategy use (i.e., a rule consistent decision boundary) for both the 1D and 2D categories. This highlights a distinction that has not yet been made in the literature: information sampling behaviors that support learning a general rule do not necessarily support learning or memory of a specific boundary (or specific examples).

Finally, 7 and 8-year-olds demonstrated stronger relations than 5 and 6-year-olds between information sampling and categorization accuracy on the 1D trial. Existing research has found an advantage of active selection over passive reception for 7-year-olds (Sim et al., 2015). The present research suggests younger children may be less likely to benefit from active selection.

Limitations

As discussed in Chapters 4 and 5, one limitation of the findings of the present study concerns the operationalization and coding of children’s information sampling behaviors and talk. First, at present there is little consensus on whether sample distance to boundary, algorithmic complexity, or an alternative measure of sampling behavior would best summarize children’s information gathering in the present task. Thus, the conclusions of this study with respect to children’s information sampling are conditional on the validity of the current measures. Second, the present study examined children’s communication at a rather coarse grain size. I did not consider the quality of planning or evidence talk, who spoke what, the dynamics of the verbal communication, or the range of non-verbal communication children produced. Fortunately, future research may use these data to examine children’s communication at a variety of levels.

Another limitation is that the task only explored children’s learning of a preference category. Learning about a categorical preference is a knowledge lean task, both in terms of the prior knowledge children brought to bear on the problem and in terms of the conceptual structure of Dobo’s preferences. Children’ information gathering and explanatory behaviors in this domain likely vary from domains in which children have existing conceptual knowledge and causal hypotheses (e.g., Howe, 2009).

The failure of the constrained dyad manipulation was also a considerable limitation of the present research. Preliminary evidence suggests the manipulation focused attention on turn taking itself, not partners’ actions and the overarching task. Furthermore, the experimenter frequently needed to intervene in order to remind children of whose turn it was. Thus, the manipulation likely failed to serve its intended purpose.

Finally, children's peer interaction and collaboration likely vary widely across a number of contexts. The present study examined dyads composed of friends or frequently playmates. Within those dyads there are a number of sources of variation that were not accounted for. For example, dyads varied in terms of how long they have been friends, the contexts in which they regularly collaborate (e.g., in school or out of school), interaction styles, and the degree to which both children directed the collaboration (i.e., power relations). In addition, the present findings might not extend to dyads composed of mere acquaintances or unfamiliar children, as prior research has shown friends coordinate collaboration differently than non-friends (Azmitia & Montgomery, 1993). Similarly, I would not expect the findings of this dissertation to generalize broadly to children from more traditional cultures. For example, Meijia-Arauz, Rogoff, Dexter, and Najafi (2007) found children of indigenous Mexican heritage in the United States are better able to coordinate complex collaborative activity than European heritage children.

Implications and Future Directions

Collaboration can serve as a means to effective learning. Indeed, 7 and 8-year-olds learned more about Dobo's preferences when they collaborated, while 5 and 6-year-olds did not. What is stake is less about learning Dobo's preferences, and more about being able to engage in productive collaboration with a peer. The findings of this dissertation support the view that children have to learn to be effective collaborative learners (Kuhn, 2015). Although preschool-aged children are able to successfully coordinate collaborative problem solving, it is only with additional experience/practice that early school-aged children are able to successfully coordinate collaboration for learning. In the case of the coordination skills examined in the present research, a number of important developments appear to occur in the first years of formal schooling. This is consistent with research suggesting experiences arising from enculturation into formal schooling are responsible for children's changing abilities to learn with others (Flynn, 2010; Young et al., under review).

Beyond the future directions discussed elsewhere in this dissertation, future work should seek to identify the kinds of experiences that drive the development of collaborative learning skills, both in early school-aged children and during other developmental periods. Furthermore, the effectiveness of interventions that encourage the use and development of collaborative skills should be investigated. This was the original intent of the failed constrained-dyad manipulation employed in the current study. Successful interventions may serve to accelerate the development of collaborative abilities in typically developing children and support the subset of children who will need assistance to productively collaborate.

Conclusion

Collaborative learning is a fundamental component of human educational and professional life. The research reported in this dissertation examined early school-aged children's collaborative learning and coordination abilities in an active category-learning task. The findings demonstrate children's collaborative learning undergoes considerable development in the early school years and specifically highlight the development of several collaborative learning skills, including the coordination of shared representation, information gathering, and communication. More broadly, these data contribute to a better understanding of when and how collaboration may be effective for learning, and address a critical gap in research on the development of collaborative skills.

References

Akaike, H. (1974). A new look at statistical-model identification. *IEEE Transactions on Automatic Control, 19,* 716–723.

Ames, G. & Murray, F.B (1982). When two wrongs make a right: Promoting cognitive change by social conflict. *Developmental Psychology, 18,* 894-897.

Ashby, F. G. & Maddox, W. T. (2005). Human category learning. *Annual Review of Psychology*, *56*, 149-178.

Azmitia, M. (1988). Peer interaction and problem-solving: When are two heads better than one? *Child Development, 59,* 87–96.

Azmitia, M., & Montgomery, R. (1993). Friendship, transactive dialogues, and the development of scientific reasoning. *Social Development, 2*, 202–221.

Baker, M. (2015). Collaboration in collaborative learning, *Interaction Studies*, 15, 451-473.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language, 68*, 255-278.

Bates, D., Kliegl, R., Vasishth, S. and Baayen, R. H. (2015). Parsimonious mixed models. arXiv:1506.04967.

Bates, D., Maechler, M., Bolker, B., and Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*, 1-48.

Bratman, M. E. (1992). Shared cooperative activity. *The Philosophical Review*, 101(2), 327-41.

Brown, A. L., & Campione, J. C. (1990). Communities of learning and thinking, or a context by any other name. *Human Development*, 21, 108–129.

Brownell, C. (2011). Early developments in joint action. *Review of Philosophy & Psychology, 2,* 193 – 211.

Bruner, J. (1961). The act of discovery. *Harvard Educational Review, 31,* 21–32.

Bruner, J., Goodnow, J., & Austin, G. (1956). A study of thinking. New York, NY: John Wiley and Sons.

Bruner, J., Jolly, A., & Sylva, K. (1976). Play: Its role in development and evolution. New York, NY: Basic Books.

Burnham, K. P. & Anderson, D. R. (2002). Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach (2nd ed.), Springer-Verlag,

Butler, L.P., & Walton, G.M., (2013). The opportunity to collaborate increases preschoolers' motivation for challenging tasks. *Journal of Experimental Child Psychology, 116*, 953-961.

Castro, R., Kalish, C., Nowak, R., Qian, R., Rogers, T., & Zhu, X. (2009). Human active learning. In D. Koller, D. Schuurmans, Y. Bengio, & L. Bottou (Eds.), *Advances in neural information processing systems* (Vol. 21, pp. 241–249).

Dede, C. (2010). Comparing frameworks for 21st century skills. In J. Bellanca & R. Brandt (Eds.), *21st century skills: Rethinking how students learn* (pp. 51–76). Bloomington IN: Solution Tree Press.

Dillenbourg, P. (1999). What do you mean by ‘collaborative learning?’ In P. Dillenbourg (Ed.), *Collaborative-learning: Cognitive and Computational Approaches* (pp.1–19). Oxford: Elsevier

Dillenbourg, P., Baker, M., Blaye, A. & O'Malley, C. (1996) The evolution of research on collaborative learning. In E. Spada & P. Reiman (Eds) *Learning in Humans and Machine: Towards an interdisciplinary learning science*. (pp. 189-211) Oxford: Elsevier.

Dillenbourg, P. & Traum, D. (2006). Sharing solutions: persistence and grounding in multi-modal collaborative problem solving, *Journal of the Learning Sciences*, *15(1)*, 121-151.

Fawcett, L. M., & Garton, A. F. (2005). The effect of peer collaboration on children's problem-solving ability. *British Journal of Educational Psychology,75,* 157–169.

Fletcher, G.E, Warneken, F., & Tomasello, M. (2012). Differences in cognitive processes underlying the collaborative activities of children and chimpanzees. *Cognitive Development*, 27, 136-153.

Flynn, E. (2010). Underpinning collaborative learning. In B. Sokol, U. Muller, J. Carpendale, A. Young, & G. Iarocci (Eds.), *Self and social regulation: Social interaction and the development of social understanding and executive functions* (pp. 311–335). Oxford: Oxford University Press.

Gauvain, M. (2001). The social context of cognitive development. New York: Guilford.

Gauvain, M., & Rogoff, B. (1989). Collaborative problem solving and children's planning skills. *Developmental Psychology, 25,* 139–151.

Gauvrit, N., Zenil, H., Delahaye, J.-P., & Soler-Toscano, F. (2014). Algorithmic complexity for short binary strings applied to psychology: A primer. *Behavior Research Methods, 46(3),* 732-744.

Gauvrit, N., Zenil, H., & Tegnér, J. (In press). The information-theoretic and algorithmic approach to human, animal and artificial cognition. In G. Dodig-Crnkovic & R. Giovagnoli (Eds.) *Representation and reality : Humans, animals and machines*. Amsterdam : Springer

Griffiths, T.L., & Tenenbaum, J.B. (2004). From algorithmic to subjective randomness. In S. Thrun, L.K. Saul, & B. Schölkopf (Eds.) *Advances in neural information processing systems*, (Vol. 16. pp. 953–960). Cambridge, MA: MIT Press.

Grimm, L.R., Maddox, W.T. (2013). Differential impact of relevant and irrelevant dimension primes on rule-based and information-integration category learning. *Acta Psychologica*. 144, 530-537.

Hamann, K., Warneken, F., & Tomasello, M. (2012). Children's developing commitments to joint goals. *Child Development, 83*, 137-145

Hmelo-Silver, C. E., O'Donnell, A. M., Chan, C. K., & Chinn, C. A. (2013). *International Handbook of Collaborative Learning.* Taylor and Francis.

Howe, C. (2009). Collaborative group work in middle childhood: Joint construction, unresolved contradiction and the growth of knowledge. *Human Development,* 39, 71-94.

Huang-Pollock, C. L., Maddox, W. T., & Karalunas, S. L. (2011). Development of implicit and explicit category learning. Journal of Experimental Child Psychology, 109 (3), 321–335

Huttenlocher, J. (1962). Effects of manipulation of attributes on efficiency of concept formation. *Psychological Reports, 10,* 503–509.

Johnson, D. W., & Johnson, R. T. (2009). An educational psychology success story: Social Interdependence theory and cooperative learning. *Educational Researcher*, 38, 365–379. *Biological Sciences*, 363, 2021–2031.

Kempe, V., Gauvrit, N., & Forsyth, D. (2015). Structure emerges faster during cultural transmission in children than in adults. *Cognition, 136*, 247-254.

Klahr, D., & Nigam, M. (2004). The equivalence of learning paths in early science instruction: Effects of direct instruction and discovery learning. *Psychological Science, 15,* 661–667.

Kolb, D. (1984). *Experiential learning: Experience as the source of learning and development.* Englewood Cliffs, NJ: Prentice Hall.

Köymen, B., Mammen, M., & Tomasello, M. (2016). Preschoolers use common ground in their justificatory reasoning with peers. *Developmental Psychology*, 52(3), 423-429.

Köymen, B., Rosenbaum, L., & Tomasello, M. (2014). Reasoning during joint decision-making by preschool peers. *Cognitive Development, 32,* 74–85.

Kruger, A., & Tomasello, M. (1986). Transactive discussions with peers and adults. *Developmental Psychology*, 22, 681–685.

Kuhn, D. (2015). Thinking together and alone. *Educational Researcher, 44*, 46–53.

Kushnir, T., Wellman, H. M. & Gelman, S. A. (2009). A self-agency bias in children's causal inferences. *Developmental Psychology, 45*, 597-603.

Laughlin, P. R. (2011). *Group problem solving.* Princeton University Press.

Lagnado, D. A., & Sloman, S. (2004). The advantage of timely intervention. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30,* 856–876.

Legare, C.H. & Clegg, J.M.(2015). The development of children's causal explanations. In S.Robson & S. Flannery Quinn (Eds.), The Routledge International Handbook of Young Children's Thinking and Understanding (pp. 65-73). New York, NY:Routledge.

Legare CH, Gelman SA, Wellman HM. (2010). Inconsistency with prior knowledge triggers children’s causal explanatory reasoning. *Child Development*. 81, 929–944

Li, M., & Vitányi, P. (2008). *An introduction to kolmogorov complexity and its applications*. Berlin Heidelberg: Springer.

Maddox, W. T., & Ashby, F. G. (1993). Comparing decision bound and exemplar models of categorization. *Perception & Psychophysics*, 53, 49-70.

Markant, D. (2016). The impact of biased hypothesis generation on self-directed learning. In Papafragou, A., Grodner, D., Mirman, D., and Trueswell, J. (Eds.) *Proceedings of the 38th Annual Conference of the Cognitive Science Society*. Austin, TX: Cognitive Science Society.

Markant, D. and Gureckis, T.M. (2014). Is it better to select or receive? Learning via active and passive hypothesis testing. *Journal of Experimental Psychology: General 143*(1), 94--122.

Mejıa-Arauz, R., Rogoff, B., Dexter, A., & Najafi,B.(2007). Cultural variation in children’s social organization. *Child Development*, 78, 1001–101

Milward, S. J., Kita, S., & Apperly, I. A. (2014). The development of co-representation effects in a joint task: Do children represent a co-actor? *Cognition, 132*, 269-279.

Montessori, M. (1964). *The Montessori method*. New York, NY: Schocken.

National Research Council. (1999). *How people learn*. Washington, DC: National Academies Press.

Nosofsky, R. (1989). Further tests of an exemplar-similarity approach to relating identification and categorization. *Perception & Psychophysics, 45,* 279–290.

Okada, T., & Simon, H. A. (1997). Collaboration discovery in a scientific domain. *Cognitive Science*, 21(2), 109–146.

Qu, L. (2011). Two is better than one, but mine is better than ours: Preschoolers' executive function during co-play. *Journal of Experimental Child Psychology, 108*, 549-566

Pacherie, E. (2012). The Phenomenology of Joint Action: Self-Agency vs. Joint-Agency. In Axel Seemann (ed.), *Joint Attention: New Developments*, Cambridge MA: MIT Press, pp. 343-389.

Papert, S. (1980). *Mindstorms: Children, computers, and powerful ideas*. New York, NY: Basic Books.

Perlmutter, M., Behrend, S. D. Kuo, F., & Muller, A. (1989). Social influences of children's Problem solving. *Developmental Psychology*, 25, 744-754.

Phelps, E., & Damon, W. (1989). Problems solving with equals: Peer collaboration as a context for learning mathematics and spatial concepts. *Journal of Educational Psychology, 81*, 639-645

Piaget, J. (1926). *The language and thought of the child.* New York: Harcourt, Brace.

Piaget, J. (1930). *The child's conception of physical causality*. New York, NY: Harcourt, Brace

Radziszewska, B., & Rogoff, B. (1991). Children's guided participation in planning imaginary errands with skilled adult or peer partners. *Developmental Psychology, 27,* 381–389.

Rakoczy, H., Gräfenhain, M., Clüver, A., Schulze Dalhoff, A., & Sternkopf, A. (2014). Young children's agent-neutral representations of action roles. *Journal of Experimental Child Psychology, 128*, 201-209

Rekers, Y., Haun, D. B. M., & Tomasello, M. (2011).Children, but not chimpanzees, prefer to collaborate. *Current Biology*, 21, 1756–1758.

Riggs, A. E. & Young, A. G. (2016). Developmental changes in children's normative reasoning across learning contexts and collaborative roles. *Developmental Psychology*, 52, 1236-1246.

Rogoff. B. 1998. Cognition as a collaborative process. In Kuhn, D., and R. S. Siegler, eds., *Handbook of Child Psychology, Volume 2: Cognition, Perception, and Language*, 679-744. New York: Wiley.

Rogoff, B. (2003). *The cultural nature of human development.* New York: Oxford University Press.

Roschelle, J. & Teasley, S.D. (1995). Construction of shared knowledge in collaborative problem solving. In C.O'Malley (Ed.), *Computer-supported collaborative learning*. New York, NY: Springer-Verlag.

Saby, J. N., Bouquet, C. A., & Marshall, P. J. (2014). Young children co-represent a partner's task: evidence for a joint Simon effect in five-year-olds. *Cognitive Development, 32*, 38-45.

Sampson, V., & Clark, D. (2009). The effect of collaboration on the outcomes of argumentation. *Science Education*, *93*, 448–484.

Schwartz, D.L. (1995). The emergence of abstract dyad representations in dyad problem solving*. The Journal of the Learning Sciences*, 4 (3), pp. 321-354.

Schwarz, B., Neuman, Y., & Biezunger, S. (2000). Two wrongs may make a right if they argue together! *Cognition and Instruction*, *18*, 461–494.

Shteynberg, G. (2015). Shared attention. *Perspectives on Psychological Science, 1*0, 579-590.

Sim, Z. L., Tanner, M., Alpert, N. Y., & Xu, F. (2015). Children learn better when they select their own data. In R. Dale, C. Jennings, P. Maglio, T. Matlock, D. Noelle, A. Warlaumont, & J. Yoshimi (Eds.), *Proceedings of the 37th annual conference of the Cognitive Science Society* . Austin, TX: Cognitive Science Society.

Singmann, H., Bolker, B., & Westfall, J. (2015). Afex: analysis of factorial experiments. R package version 0.13–145.

Sobel, D. M., & Kirkham, N. Z. (2012). The influence of social information on children’s statistical and causal inferences. In J. B. Benson (Series Ed.) & F. Xu & T. Kushnir (Vol. Eds.), Rational constructivism in cognitive development (pp. 321–350). Waltham, MA: Academic Press.

Sobel, D., & Kushnir, T. (2006). The importance of decision making in causal learning from interventions. *Memory & Cognition, 34,* 411–419.

Soler-Toscano, F., Zenil, H., Delahaye, J.-P., Gauvrit, N. (2013). Correspondence and independence of numerical evaluations of algorithmic-theoretic information measures. *Computability (2)*, 125-140.

Sommerville, J. A., & Hammond, A. J. (2007). Treating another’s actions as one’s own: Children’s memory of and learning from joint activity. *Developmental Psychology*, 43, 1003–1018.

Steyvers, M., Tenenbaum, J., Wagenmakers, E., & Blum, B. (2003). Inferring causal networks from observations and interventions. *Cognitive Science, 27,* 453–489.

Teasley, S. D. (1995). The role of talk in children's peer collaborations. *Developmental Psychology*, 31 (2), 207-220.

Tomasello, M. (2014). *A Natural History of Human Thinking.* Harvard University Press.

Tomasello, M.(2016). Cultural learning redux. *Child Development,* 87, 643-653

Tomasello, M., Carpenter, M., Call, J., Behne, T., & Moll, H. (2005). Understanding and sharing intentions: The origins of cultural cognition. *Behavioral and Brain Sciences*, 28(05), 675-691.

Tomasello, M. & Hamann, K. (2012). Collaboration in young children. *Quarterly Journal of Experimental Psychology*, 65, 1-12

Tomasello, M., Kruger, A. C., & Ratner, H. H. (1993). Cultural learning. *Behavioral and Brain Science, 16*, 495-552

Trilling, B., & Fadel, C. (2009). *21st century skills: Learning for life in our times*. San Francisco, CA: Jossey-Bass.

Tudge, J., & Winteroff, P. (1993). Can young children benefit from collaborative problem solving? Tracing the effects of partner competence and feedback. *Social Development, 2,* 242–259.

Tuomela, R. (2007). *The philosophy of sociality*. New York, NY: Oxford University Press.

Vygotsky, L.S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.

Warneken, F., Steinwender, J., Hamann, K., & Tomasello, M. (2014). Young children's planning in a collaborative problem-solving task. *Cognitive Development, 31,* 48-58.

Young, A.G., Alibali, M.W., & Kalish, C.W. (2012). Causal learning and disagreement: Others' hypotheses affect children's evaluations of evidence. *Developmental Psychology, 48(5),* 1242-1253.

Young, A.G., Alibali, M.W., & Kalish, C.W. (under review). Causal learning from joint activity: A developmental shift in early school-aged children.

Zenil, H., Soler-Toscano, F., Delahaye, J. -P., & Gauvrit, N. (2012). Two- dimensional Kolmogorov complexity and validation of the coding theorem method by compressibility, arXiv 1212:6745