Summer 2012

The role of word learning in the development of dimensional attention

Lynn Krieg Perry
University of Iowa

THE ROLE OF WORD LEARNING IN THE DEVELOPMENT OF DIMENSIONAL ATTENTION

by

Lynn Krieg Perry

An Abstract

Of a thesis submitted in partial fulfillment
of the requirements for the Doctor of
Philosophy degree in Psychology
in the Graduate College of
The University of Iowa

July 2012

Thesis Supervisor: Associate Professor Larissa K. Samuelson

ABSTRACT

Previous work shows that young children focus on holistic (or overall) similarity and older children focus on dimensional similarity (selectively attending to one property to the exclusion of others). Research on early word learning, however, suggests that process of learning new words trains attention towards category-relevant dimensions via regularities in the linguistic and physical environment. Thus, over development, children learn to attend to specific dimensions when making nominal category judgments—they selectively attend to shape, for example, when learning names for solid objects. In four experiments, I asked a question fundamental to our understanding of dimensional attention: does word learning scaffold attention to dimensional similarity in more general contexts. The results of Experiment 1 showed that children who are holistic classifiers are slower than dimensional classifiers to learn categories of objects that vary along both a category-relevant dimension (e.g. size) and a category-irrelevant dimension (e.g. brightness). However, the results of Experiment 2 showed that when children were presented with incidental labels during category learning, holistic classifiers learn the categories as quickly as dimensional classifiers. In a follow-up similarity classification task, children who had been holistic classifiers showed an increase in dimensional attention only if they had been in the label experiment. In Experiments 3 and 4, I examined category learning with and without a label in children who preferred to selectively attend to one dimension of similarity (e.g. brightness) regardless of whether this means selecting dimensional or holistic matches in a classification task. The results of these experiments provide a more complete picture of the continuous developmental trajectory of increasing selective and flexible dimensional attention. By showing how labels support dimensional attention, these results clarify the processes involved in development of similarity perception and potentially unify our understanding of attentional processes in word learning with those in a broader context.

Abstract Approved: ____________________________________
Thesis Supervisor

____
Title and Department

____
Date

THE ROLE OF WORD LEARNING IN THE DEVELOPMENT OF DIMENSIONAL ATTENTION

by

Lynn Krieg Perry

A thesis submitted in partial fulfillment
of the requirements for the Doctor of
Philosophy degree in Psychology
in the Graduate College of
The University of Iowa

July 2012

Thesis Supervisor: Associate Professor Larissa K. Samuelson

Graduate College
The University of Iowa
Iowa City, Iowa

CERTIFICATE OF APPROVAL

PH.D. THESIS

This is to certify that the Ph.D. thesis of

Lynn Krieg Perry

has been approved by the Examining Committee
for the thesis requirement for the Doctor of Philosophy
degree in Psychology at the July 2012 graduation.

Thesis Committee: ______________________________
Larissa K. Samuelson, Thesis Supervisor

____
Karla K. McGregor

____
Bob McMurray

____
Jodie M. Plumert

____
John P. Spencer

____
Teresa A. Treat

To the Language Development student whose name I've long since forgotten who came to my office hours and said, "I don't know what this paper 'Lupyan, Rakison, & McClelland' is," to which I responded, "I don't know either, let me Google that for you." Your lack of attention in class has forever changed the course of my life.

To illustrate, how does one learn to distinguish claret from burgundy? … the adhesion of each wine with its own *name* becomes more and more inveterate, and at last each flavor suggests instantly and certainly its own name and nothing else. The names differ far more than the flavors, and help to stretch these latter farther apart. Some such process as this must go on in all our experience. Beef and mutton, strawberries and raspberries, odor of rose and odor of violet, contract different adhesions which reinforce the differences already felt in the terms.

William James
Principles of Psychology

ACKNOWLEDGMENTS

First, I would like to thank my parents, Kris Krieg Perry and Rick Perry, and my brother, Ray Perry, for their support over the years. My family has been immensely helpful in encouraging me to follow my interests. For example, as a young child, I loved to categorize and sort Halloween candy according to different dimensions of similarity—size, wrapper color, brand, etc.—and then remark upon all the different possible sorts and re-sorts I could make. Instead of finding this incredibly strange, my family indulged my eccentricity, allowing me to eventually pursue these interests in a scientific setting.

I would also like to thank my advisor Larissa Samuelson for her guidance and mentorship these last six years. She has helped me through the ups—always having been there for celebratory dance parties—and the downs—quoting the wise Professor Remus Lupin, "Chocolate helps," and breaking me off a piece of the good stuff the first time I totally botched a data analysis. I have learned countless lessons about research, teaching, mentorship, time management, and being a good colleague, both directly and indirectly from her during my time working at Iowa. One of the most important lessons Larissa has worked to teach me is how to be a good writer. From the onset of my graduate career, she worked closely with me to scaffold my writing. As time passed, she gradually removed the scaffolding bit by bit and really pushed me to develop my writing independently and rise to the occasion. This was, and still is, a difficult challenge, but her extreme patience, empathy, and humor help me improve with each draft. I am grateful beyond words for the opportunities I have had from working with Larissa. She has been and will continue to be a role model for science and life.

I would like to thank Susan Wagner Cook for introducing me to the world of mixed models. Her time and mentorship have allowed me to take my research to exciting new places. Her encouragement—in statistics and beyond—has allowed me to pursue a career path that I would not have even entertained as being possible otherwise.

I would like to thank John Spencer for his support and mentorship. I am appreciative for his time and patience in teaching me about computational modeling. Working with John on DNF modeling has greatly influenced my thinking on cognition and development even in non-modeling contexts, and has forced me to ground my thinking and consider "how would this work in a model?" He has also been such an inspiration to me in thinking about how to simultaneously pursue deep scientific questions and have an influence on public policy. I hope to one day be even half as successful at improving humanity across multiple timescales.

I would like to thank Bob McMurray for challenging me to think deeply about developmental mechanisms and for not only encouraging my contrariness, but also for challenging me and being contrary back. I truly appreciate the opportunity that his teaching gave me to try out and talk though different ideas to see how they felt without feeling prematurely tied to any one in particular.

I would also like to thank the other members of my Dissertation Committee. I would like to thank Jodie Plumert for always making me consider "what's the big picture?" and "what does this mean for development?" when thinking about any research question. I would like to thank Karla McGregor for introducing me to Vygotsky's writings and the concept of scaffolding, both of which have been hugely influential on my development as a developmentalist. I would like to thank Teresa Treat for her career guidance and encouragement and believing me in a way that helped me believe in myself.

I would also like to thank the members of the Language Discussion Group and Delta Center, especially Amanda Owen Van Horne, Julie Gros-Louis, and Bruce Tomblin for their encouragement and positivity. I will always look back fondly on our summer book clubs, Bold Assertion Nights, Delta Center breakfasts, and the stimulating intellectual discussions we had.

I would like to thank my friends in the Developmental Science program for their support over the years. In particular, Sammy Perone, Vanessa Simmering, and Aaron

Buss have all been instrumental in the development and execution of this project. I would especially like to thank Aaron for his patience and programming help, without which I could never have gotten these experiments up and running as quickly as I did. I would also like to thank all the journal club friends—especially Keith Apfelbaum, Amber Dewey, Marcus Galle, Caitie Hilliard, Gavin Jenkins, Effie Kapnoula, Sarah Kucker, Valerie Mendez, Cheyenne Munson, Elizabeth O'Neal, Stephanie Packard, Fabian Soto, Teresa Stephens, Joe Toscano, Zhen Wu, Litsa Xeimariou, and Libo Zhao for sharing many delightful discussions and delicious beers over the years—both of which were critical to the success of this project.

I would like to thank Jessica Horst, who has been the best academic big sister I could ever have hoped for. She has been there for me for the last six years to listen and advise as needed even from across the pond. Her foresight and organizational skills made my transition to and progress through graduate school immensely easier than it would have been otherwise. I am thankful for her amazing ability to sympathize with and clarify any problem—usually by making an analogy to hypothetical scenarios involving famous developmental psychologists. This rhetorical technique has proven to be incredibly helpful in a surprisingly large variety of situations, and I am quite grateful for it.

I would like to thank Michelle Quinn, who has been an incredible source of inspiration and encouragement. She has always been there with a kind word or poem to keep me going. Especially during the final semester of my graduate work she has been there to remind me that living a meaningful, passionate life full of friends and the quest for knowledge is what matters most. I am immensely appreciative of the advice and wisdom I've received from my friend who "exists in eternity, not in time."

I would also like to thank Katie Haggerty for being there for me as a friend ever since we first met at the Iowa interview weekend. I am so glad our paths crossed and that we were able to share in so many wonderful (mis)adventures over the years. I am influenced and inspired daily by her belief in me and in the idea that things happen for a reason and that something great is going to eventually come from every (apparent) setback.

I would like to thank Becky Huber, Joyce Paul, Dawn Miller, Shawn Gelo, and Keith Miller for all their help and patience over the years. Without their much needed assistance, none of my work would been realized. I would also like to thank Ryan Schiffer and all the students who have helped me collect and code data over the past six years. I would not be where I am without their tireless dedication of all my former research assistants. Thank you to: Elizabeth Aljets, Laura Aljets, Eric Anderson, Kristen Badsing, Alison Bahnsen, Allison Bean, Jenny Burdinie, Kara Burns, Janie Cobb, Halsey Darling, Christina Davison, Emily Dasquisto, Carly Flannigan, Allison French, Tracy Glink, Breanne Higgins, Sieon Kim, Danielle Korthals, Allison Kushnir, Lisa Malloy, Amanda Mather, Ian Messenger, Gabriela Mikleau, Stephanie Moffitt, Rachael Mooney, Laura Reali, Colleen Russo, Stephen Ryan, Suzanne Scheib, Tara Schroder, Jordan Simpson, Jodi Smith, Kassandra Stritmatter. And an extra special thanks to Emily, Breanne, and Laura Reali who helped me in the final months when things got really crazy. And an extra extra special thanks to Jenny, Jodi, and Stephen Moliter who help me to remember what the point of it all is. I hope they too continue to pursue their scientific passions and that our paths cross again someday.

I would also like to thank all the participants and parents who made my work possible. I'd especially like to thank Alek and Katya Spencer for selflessly volunteering to participate in my studies over the years. I've been extremely fortunate that my research questions have changed over time in such a way that one or both of them was always the right age to participate in my current project.

I would like to thank Gary Lupyan for writing a paper that has inspired much of my thinking over the past four years and led to the initial conception of this project. I am also extremely grateful to him for taking a chance on me and believing I could do things that I previously would have been too scared to even attempt.

Finally, I would like to thank Linda Smith and Steve Hockema for their encouragement and guidance from the very beginning of my development as a scientist. Their faith in me has helped to get where I am today, and I am forever grateful for their support over the years. Their over-the-top praise ("we've decided—you're going to be famous") has helped me persevere through small setbacks, and I hope to eventually live up to their expectations.

ABSTRACT

Previous work shows that young children focus on holistic (or overall) similarity and older children focus on dimensional similarity (selectively attending to one property to the exclusion of others). Research on early word learning, however, suggests that process of learning new words trains attention towards category-relevant dimensions via regularities in the linguistic and physical environment. Thus, over development, children learn to attend to specific dimensions when making nominal category judgments—they selectively attend to shape, for example, when learning names for solid objects. In four experiments, I asked a question fundamental to our understanding of dimensional attention: does word learning scaffold attention to dimensional similarity in more general contexts. The results of Experiment 1 showed that children who are holistic classifiers are slower than dimensional classifiers to learn categories of objects that vary along both a category-relevant dimension (e.g. size) and a category-irrelevant dimension (e.g. brightness). However, the results of Experiment 2 showed that when children were presented with incidental labels during category learning, holistic classifiers learn the categories as quickly as dimensional classifiers. In a follow-up similarity classification task, children who had been holistic classifiers showed an increase in dimensional attention only if they had been in the label experiment. In Experiments 3 and 4, I examined category learning with and without a label in children who preferred to selectively attend to one dimension of similarity (e.g. brightness) regardless of whether this means selecting dimensional or holistic matches in a classification task. The results of these experiments provide a more complete picture

of the continuous developmental trajectory of increasing selective and flexible dimensional attention. By showing how labels support dimensional attention, these results clarify the processes involved in development of similarity perception and potentially unify our understanding of attentional processes in word learning with those in a broader context.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I: THE DEVELOPMENT OF DIMENSIONAL ATTENTION

Over the course of development, we become increasingly skilled at attending to one thing to the exclusion of others (see Hanania & Smith, 2010 for review). For example, adults can easily focus on the color of a lime, rather than its exact shape or size, in order to distinguish it from a lemon. This ability, known as selective attention, is one that has long been recognized as critical to our cognition. In fact, William James, in his seminal book Principles of Psychology, wrote, "the art of being wise is the art of knowing what to overlook," (1950). Evidence that older children and adults are generally much better than preschool age children at selectively attending to one dimension or property to the exclusion of others comes from a variety of domains including rule use (e.g. Frye, Zelazo, & Palfai, 1995), discrimination learning (e.g. Kendler & Kendler, 1962), and similarity classification (e.g. Smith & Kemler, 1978). However, we still do not know the process by which this change occurs. The goal of this project is to explore the processes that drive changes in selective attention with respect to their effects on one particular domain: similarity perception. For example, how do children become able to judge a lime and a grasshopper as being the same with respect to color even though these objects are different with respect to other dimensions?

Similarity perception is critical to cognition. It has even been proposed to be the basis of categorization (Goldstone, 1994a, Medin, Goldstone, & Gentner, 1993). Thus, similarity has implications for a variety of domains, from basic cognitive processes such as object recognition (Edelman, 1999) and problem solving (Simon 1978), to more complex behaviors such as race perception (Greenberg & Macgregor-Hannah, 2010), to the vast developmental differences present in autism (Gastgeb et al., 2006). James had much to say on this topic as well. He believed similarity to be such a significant driving force underlying cognitive processes that he described it as “the very keel and backbone of our thinking” (1950).

The holistic-to-dimensional shift in similarity classification

In the years since James, much has been done towards understanding how adults perceive similarity while considerably less work has been done on its development. What we do know comes from descriptions of shifts in classification behavior from one age to another. Of particular relevance to the current study is the holistic-to-dimensional shift, or the tendency for young children to focus on holistic similarity and older children and adults to focus on dimensional similarity relationships (Smith & Kemler, 1977). Imagine you are presented with an orange, a yellowish-orange ball, and a yellowish-orange toy car. If you are an adult, you would be more likely to group the ball with the car because they match exactly along one dimension (i.e. they are identical in color). A young child, however, would group the orange with the ball because they are similar along multiple dimensions (i.e. shape and color)—they are holistically similar. These changes in similarity perception occur during the early school-age years of childhood, such that younger children (< 8-years-old) tend to be holistic classifiers and older children dimensional classifiers.

Evidence for such a shift comes from research using free classification tasks. As an example, a triad classification task is pictured in Figure 1. As can be seen in the figure, two stimuli (A and B) match on one dimension (e.g. size) but vary greatly along another dimension (e.g. brightness). The third stimulus (C) is highly similar to the first (A) along both dimensions, but not identical to it (A) on either. If a participant were using holistic similarity, he or she would classify A and C together. If a participant were using dimensional similarity, he or she would classify A and B together. Using this task Smith and Kemler (1977) found that 5-6 year-olds made mostly AC matches and 10-11 year-olds made mostly AB matches. However, other research on the holistic-to-dimensional shift shows even 5-6 year-olds attend to dimensional similarity relationships in some task contexts (Smith & Kemler, 1978; Smith, 1983). Additionally, analyses of individual children's patterns of responding have revealed a common intermediate strategy between holistic and dimensional classification. A child following this "preferrer" strategy would selectively attend to one dimension, but not the other, on all trials, regardless of whether that meant selecting the holistic match or the dimensional match on a given trial (see Hanania & Smith, 2010). Thus, over development children go from an initial tendency to attend to overall similarity, to preferentially attending to one dimension of similarity, to being able to flexibly shift attention between dimensions depending on context. Together, these results suggest that the development of similarity classification is not a discrete change so much as a continuous shift towards increasingly selective and flexible attention.

This more continuous picture fits with research on the development of attention switching showing that much younger children are able to attend to one dimension to the exclusion of others in some tasks. For example, in the dimensional change card sort (DCCS) task, 3-4 year-old children are asked to sort cards according to a rule, such as cards with pictures of blue shapes go in a pile and those with red shapes go in another pile (regardless of object shape). The central question in this task is whether, after attending to one dimension for a number of trials, children will be able to flexibly switch to attending to values along the other dimension. The canonical result is that on "post-switch" trials, when the rule is switched so that now cards with one shape go in one pile and cards with another shape go in different pile (regardless of color), 3-year-old children tend to perseveratively sort by the first rule, while 4-year-old children are able to switch to the new rule (Frye et al., 1995). What processes enable 4-year-old children to show "mature" dimensional attention in this task while 5-6 year-old children cannot in the triad classification task?

One possibility is that the DCCS task makes fewer demands on children's dimensional attention than the triad classification task. For example, in the DCCS children have to attend to specific values (e.g. blue and red) rather than a range of values along a dimension. Likewise, in deciding where to sort cards in the DCCS task, children have to choose between two visible targets that each match the sort card exactly on one dimension. In the triad classification task, on the other hand, overall (holistic) similarity is always pitted against dimensional similarity. Thus, a child who has even a small amount of trouble attending to one dimension at a time will be biased to choose the holistic match. For this reason, these two tasks are historically thought to measure different aspects of attentional development; the DCCS task is typically thought to measure the development of attention switching rather than selective attention.

Nevertheless, Hanania & Smith (2010) recently argued that the development of selective attention also increases flexible attention switching. The crux of their argument comes from (what Hanania and Smith refer to as circumstantial) evidence that the developmental changes involved in children's performance in the DCCS task mirror the changes involved in their performance in the triad task. In both tasks, children first become better at being able to attend selectively to one dimension. However, data from both tasks also show "sticky attention" to that one dimension—either pereseveratively sorting by the first rule in the DCCS task, or preferentially attending to only one dimension of similarity in triad classification. Thus, in both these cases, it is clear that the child with sticky attention lacks the flexibility to switch between attending to different dimensions. Considered together, the data from attention switching and selective attention tasks reveal a general trend over development from nonselective attention, to relatively selective but not flexible attention, to selective and flexible. Hanania and Smith therefore hypothesize that flexible attention is a necessary party of mature selective attention.

For this reason, then, understanding the developmental changes that allow children to succeed in flexible attention switching should shed light on the changes that allow them to succeed in selective attention more generally. An example of such research comes from a study using the DCCS task that suggests a relationship between words and attention. Kirkham, Cruess, & Diamond (2003) found that asking children to label the relevant dimension as they sorted cards on the post-switch trials of the DCCS resulted in less perseveration and more correct sorting by the new rule. One possible suggestion from this experiment is that when children name the specific value along the relevant dimension (e.g. "blue" when sorting by color), it forces them to attend to that value Once children's attention is drawn to the relevant dimension they are able to correctly sort the card. Thus, the process of labeling can support the weak selective attention abilities of young children. This supporting process, similar to what socio-cultural theorists have described as "scaffolding," occurs during the immediate timescale of online attention and decision making. However, modeling work has shown how this scaffolding can support the emergence of dimensional attention over longer timescales of learning and development.

Over the same period in which dimensional attention abilities are developing, children are also learning many dimension words (e.g. color, size, etc.) and the names of values along those dimensions (e.g. blue, small, etc.). Modeling work by Smith and colleagues (1997) suggests that learning these dimension and value words is what builds representations of these dimensions. Specifically, Smith, Gasser and Sandhofer (1997) taught a neural network that demonstrated holistic perception to label attributes of individual objects and to compare attributes between objects. Labeling and comparing attributes led the network to selectively isolate both dimensions that corresponded to sensory dimensions pre-specified in its environment and novel dimensions that did not correspond to these sensory dimensions. This led the network to make similarity comparisons and thus selectively attend to these dimensions. The results of this model suggest that word learning plays a critical role in the change from holistic to dimensional similarity classification. This, along with the results of Kirkham et al., 2003, suggests that naming specific features and dimensions helps children and neural networks attend to those dimensions both over developmental and immediate timescales. The modeling work shows how learning the names of dimensions can help a system to learn to attend to dimensions at all, while the work using the DCCS tasks show how using already-learned dimension names can support flexible, selective attention to already-learned dimensions in a weak attentional system. Other work, however, suggests that word learning may more generally play a role in the perception of similarity.

Around the same time children are undergoing these changes in dimensional attention, there are also changes in the way words influence their perception of similarity. For example, work by Katz (1963) demonstrated that when two objects are named with the same label, 7- and 9-year-old children were more likely to perceive them as the same than if they were named with different labels. This effect was more pronounced in 7 than 9-year-olds suggesting younger children's similarity perception is more affected by label cues. Similarly, work by Landau & Shipley (2001) demonstrates that when two objects are given the same label, children and adults will generalize that label to all intermediate morphs between the objects. If, however, the two objects are given two different names or no names at all, participants divide the intermediates into two distinct categories. The authors argue that this provides evidence for labels "boosting the equivalence" of objects. This work by Landau and Shipley (2001) and Katz (1963) advanced our understanding of the relationship between word learning and similarity. However we cannot tell if, for example, children judged objects with a common label as being the same because they selectively attended to the dimension in common, or because they know objects within a common category (or associated with a common name as in the Landau & Shipley study) should be treated equivalently.

It is clear, however, that there is a tight link between words, categories and similarity: words refer to categories, categories are organized around some form of similarity, and words can highlight the nature of that similarity. Thus, in order to understand the development of any one of these processes, we must consider the interactions between them. In particular, how do category learning and word learning increase dimensional attention? The goal of this dissertation is to understand the developmental interactions between word learning and categorization that lead to selective attention to dimensional similarity. In the rest of this chapter, I review recent findings from the developmental and adult cognitive literatures that point to relationships between word learning, categorization, and attention to dimensional similarity. This work helps to clarify the mechanisms driving the holistic to dimensional shift in similarity classification. I first review work on the relationship between dimensional attention and category learning. Next I describe recent work on the relationship between word learning and category learning. Then, I discuss work on the relationship between word learning and dimensional attention. Finally, I propose an experimental account of how these processes fit together over development.

The relationship between dimensional attention and category learning

A rich body of research suggests that learning a category changes our perceptions of the similarity of its members. Adults' ability to selectively attend allows them to form novel categories of objects that require attention to small changes along one relevant dimension (e.g. size) but not to changes along an irrelevant one (e.g. brightness) (e.g. Goldstone, 1994b). Learning categories organized in this way then changes adults' ability to make discriminations along relevant and irrelevant dimensions (Goldstone, 1994b). In particular, Goldstone found that adults show enhanced discrimination for between-category exemplars along relevant dimensions and diminished discrimination for within-category exemplars along irrelevant dimensions when these dimensions are perceptually separable (i.e. dimensions for which adults can selectively attend to one and ignore the other, such as size and brightness). When these dimensions are perceptually integral (i.e. dimensions for which adults cannot typically selectively attend to one and ignore the other, such as hue and brightness), adults still showed enhanced discrimination for between-category exemplars along the relevant dimension but no diminished discrimination along the irrelevant dimension.

That adults could learn categories that forced them to attend to one dimension and ignore another even when those dimensions are not separable, suggests that learning a category increases attention to separate dimensions. Perhaps, then, even young children with holistic similarity perception could learn to separate dimensions that are "integral" to them in a task such as this. In fact, work by Minda, Desroches, and Church, (2008) demonstrated that children as young as 3-years-of-age can successfully learn a category that is organized by a single-dimensional rule. For example, children learned to assign all blue objects to one category and all orange objects to another, regardless of size and shape. However, the children in Minda et al.'s study were not tested on novel category exemplars after learning (nor were adults in Goldstone's study). Thus, we do not know whether children had actually learned to attend to the relevant dimension. They could also have learned to remember exact exemplars or the relevant attributes rather than dimensions (i.e. orange and blue rather than color). Thus, while learning to make categorical distinctions seems to enhance dimensional attention, it may be that one must have the ability to selectively attend to a single dimension to even learn such categorical distinctions.

In fact, a review of earlier work on discrimination learning suggests that children can learn a single dimension category (as in Minda et al.'s 2008 study) without abstracting a dimensional basis for the distinction (Kendler, Kendler & Wells, 1960; Kender & Kendler, 1962; and see Hanania & Smith, 2010 for review). This line of research showed that both preschool and school age children can learn a category distinction, such as white triangles and white circles go in one category and black triangles and black circles go in another, quite well. However, the two groups of children did not learn to categorize in the same way. Evidence for the different methods of learning came from children's ability to learn a new categorization rule after they had mastered the first. In the second phase of the experiment, the rule was switched so that children either had to categorize by 1) the same dimension as before but in the opposite way so that, for example, the black objects now go in the first category and the white objects now go in the second category (i.e. an intra-dimensional shift) or 2) by the other dimension so that triangles now go in one category and circles in another (extra-dimensional shift). Older children were quicker to learn a new rule after an intra-dimensional shift rather than after an extra-dimensional shift suggesting they learned to pay attention to the color dimension and ignore the shape dimension. Younger children did not show this advantage and instead seemed it make these category distinctions by memorizing whole stimuli. Memorizing which whole objects go into which categories makes it easier to switch rules regardless of whether the switch is intra- or extra-dimensions. Thus, while even young children are often skilled category learners, the selective attention skills of older children may be required to attend to category-relevant dimensions of similarity.

Developmental changes in category learning are related to changes in children's selective attention abilities. Selective attention is necessary for learning some kinds of categories, but less so for others. In particular, selective attention is especially important for learning categories organized by similarity along a single dimension (Lupyan, Mirman, Hamilton, & Thompson-Schill 2012; Sloutsky, 2010; Ashby & Maddox, 2011). These categories, which I will call "dimensional," (alternatively called "rule-based" by Ashby, "selection-based" by Sloutsky, or "low-dimensional" by Lupyan), require the learner to selectively attend to one category-relevant dimension and abstract across category-irrelevant dimensions. In contrast, what I will call “holistic” categories (alternatively called “compression-based” by Sloutsky or “high-dimensional” by Lupyan), have a large number of covarying features and require the learner to attend to overall similarity across multiple dimensions rather than selectively attending to one dimension. Developmental research has shown that children learn holistic categories as well as adults do (Kloos & Sloutsky, 2008; Minda et al., 2008), but that they require explicit verbal instructions about the relevant feature to learn dimensional categories (Kloos & Sloutsky, 2008). In fact, Ashby and colleagues have argued that the optimal strategy for learning dimensional categories is to verbalize a rule (e.g. green objects go here, blue objects there). However, for holistic categories, learners need to integrate information across dimensions, and verbalization is not a useful strategy (Ashby & Maddox, 2011; and see Sloutsky, 2010). In other words, using a verbalizable rule that labels a relevant dimension or feature value can help learners categorize only in cases where selective attention is needed. While adults can learn dimensional categories even without externally provided verbal rules, they may engage language to support this categorization. In fact, this idea is supported by studies that show verbal interference (Lupyan, 2009) and language impairments such as aphasia (see Vignolo, 1999 for review) impair adults’ dimensional categorization and sorting (see R. Cohen, Kelter, & Woll, 1980; R. Cohen, Woll, Walter, & Ehrenstein, 1981; and Lupyan & Mirman, under review, for similar arguments).

Children, however, have relatively immature selective attention abilities and thus require externally provided verbal rules to learn dimensional categories. This developmental change supports the proposal that there is a tight link between dimensional categorization, attention and language. Thus, if labeling relevant dimensions or features increases selective attention, then changes in children’s knowledge and use of labels for dimensions should be related to changes in selective attention to those dimensions. We know that selective attention is required for dimensional categorization and that language has been shown to support selective attention. However, it still unclear what the nature of the relationship is and whether language is what *drives* developmental changes in selective attention.

The relationship between labels and category learning

Recently, research on adult's category learning clarified the nature of the relationship between language and categorization (Lupyan, Rakison & McClelland, 2007). Lupyan and colleagues demonstrated that when learning categories, adult participants given incidental labels learned more quickly and were more accurate at test than those not given labels. Participants were presented with novel objects that varied along two dimensions and asked to decide to which of two novel categories they belonged. Participants in the label condition were presented with a novel category label after feedback on each category learning trial. The other group only received feedback and no label. Importantly, the labels presented to participants in the label condition were completely redundant to the category—i.e. they provided no new information above and beyond the feedback. Nevertheless, participants in the label condition learned the categories more quickly than participants in the no-label condition. Furthermore, other redundant cues, such as an associated spatial location, did not affect participant's learning. The authors suggested that labels facilitate learning because they are (historically) better predictors of category membership than other cues.

Other researchers have referred to this facilitative effect as linguistic bootstrapping (Yoshida and Smith, 2005). Yoshida and Smith (2005) argued that redundant associations between linguistic cues strengthen the associations between other perceptual cues and category structure. These authors taught Japanese–speaking children categories of solid things organized by similarity in shape and nonsolid things organized by similarity in material, either with or without redundant syntactic cues (e.g. always using the classifier "hitotsu no" when naming solid objects and the quanitifier "sukoshi no" when naming nonsolid substance. Children trained with these syntactic cues were more accurate in a novel noun generalization transfer task than children trained without syntactic cues. Potentially, this was because these redundant syntactic cues facilitated attention to the associated category-relevant dimensions: to shape for solid things and to material for nonsolid things. Importantly, Yoshida and Smith argued that the facilitative effect of labels is fundamentally developmental. Labels do not begin with a privileged status in learning, but rather begin to take on this role over the course of development. That is, over the course of word learning, the frequent redundancy between labels and other cues such as solidity, syntax, and category organization helps children use labels to facilitate their learning. The more experience children have learning regularities in these overlapping cues, the better they can attend to category-relevant information in word learning.

The role of word learning on the development of dimensional attention

This work by Yoshida and Smith ties into a larger body of work on the development of dimensional attention in word learning. That work suggests that the process of learning words does scaffold dimensional attention, at least in its application to further word learning. During word learning, labels guide children's attention to category-relevant dimensions. In particular, children acquire certain word-learning biases that help them attend to relevant dimensions of novel stimuli when learning new words in a specific context. For example, children acquire a shape bias, or the tendency to generalize names of novel objects by similarity in shape (Landau, Smith, & Jones, 1988). Research has suggested that this bias emerges from the regularities present in the linguistic environment. Specifically, most words that children learn early name categories of solid objects organized by similarity in shape, e.g. ball, cup, (Samuelson & Smith, 1999). As children learn more of these labels, their attention is trained such that they automatically attend to a specific feature an object—the shape—when learning names for novel solid objects (Smith et al., 2002). As they learn more words that name categories organized in other ways, for example names for nonsolids in categories organized by similarity in material substance or adjectives that name properties of objects, they acquire other biases and learn to flexibly attend to context-appropriate dimensions (Jones & Smith, 1993).

Thus, Smith and colleagues (2002) describe word learning as "on-the-job training for attention," and have shown that teaching young children names of categories organized by similarity in shape leads them to precociously attend to shape when learning new words (see also, Perry, Samuelson, Malloy, & Schiffer, 2010; Samuelson, 2002). Thus, over the course of development, learning words helps direct children's attention to dimensional similarity in future word-learning contexts. The critical unanswered question, however, is whether learning words also helps direct children's attention to dimensional similarity in non-linguistic contexts. I propose that word learning provides on-the-job training for attention more globally, meaning changes we see in word-learning biases are connected to changes we see in similarity perception—labels scaffold dimensional attention in similarity perception.

Specific proposal

It has already been established that category learning influences adults' perceptual discrimination along category-relevant dimensions (Goldstone, 1994b). It has not yet been established, however, whether category learning can influence children's perceptual discrimination in the same way. In the early school years, children are undergoing large changes in their ability to attend to one dimension to the exclusion of others. These changes should affect not only the way in which they learn categories, but the influence category learning has on their perception. Additionally, what drives changes in children's emerging dimensional attention in similarity classification tasks has not yet been established. Research on novel noun generalization has demonstrated that word learning trains children's selective attention in the service of learning new words. Thus, I propose that word learning scaffolds the emergence of dimensional attention over development, but also within the context of learning an individual category. It is my prediction, then, that while children who attend to holistic similarity will have difficulty learning categories that require selective attention to one dimension to the exclusion of others, the presence of a novel label can scaffold this attention. In 4 experiments, I asked: 1) if there are developmental differences in dimensional category learning that relate to dimensional attention in similarity classification, 2) whether labels can support dimensional attention and facilitate categorization, and 3) if there are developmental differences in the way category learning influences perceptual discrimination.

In Experiment 1, I examined whether the influence of category learning on perception of category-relevant dimensions rests on a 5-to-8-year-olds' ability to demonstrate attention to dimensional similarity. I predicted that children who classify holistically in a triad task would have more difficulty learning categories defined by similarity along a single dimension (i.e. dimensional categories). Further, I predicted these children would not show changes in discrimination along either the dimension that was criterial for the category they learned or an irrelevant dimension. On the other hand, I predicted that children of the same age who could classify dimensionally should be able to learn dimensional categories and would show changes in discrimination similar to the pattern demonstrated by adults in earlier studies (i.e. Goldstone, 1994b).

If labels facilitate categorization by increasing selective attention to relevant dimensions, then they may also scaffold dimensional attention in children who are not able to attend dimensionally on their own. Thus, in Experiment 2, I examined whether the presence of an incidental label would enable even holistic classifiers to learn dimensional categories and lead them to subsequently show changes in discrimination and classification. Furthermore, I predicted that the advantage in categorization these children would demonstrate relative to their counterparts in Experiment 1 would relate to increased changes in discrimination to relevant and irrelevant dimensions.

If the holistic to dimensional shift is a continuous increase in selective and flexible attention, then examining children who use an intermediate classification strategy, such as the preferrer strategy, should greatly inform our understanding of selective attention across the developmental spectrum. Thus, in Experiment 3, I examined the influence of category learning on preferrers' attention to category-relevant dimensions. In Experiment 4, I examined whether the presence of a label during category learning would affect preferrers' subsequent discrimination and classification abilities. Because little is currently known about what it means to prefer a dimension when judging similarity, I have no specific predictions regarding these children's performance. Nevertheless, examining their behavior in these experiments will play an important role in expanding our understanding of the continuity in the development of selective attention.

Thesis organization

Experiments 1-4 used nearly identical tasks that differed only in the presence of labels in Experiment 2 and 4. These tasks were designed to compare category learning and subsequent attentional changes in children with different classification biases when categories are labeled versus not. Because all four experiments used almost identical procedures, I describe the general methods in Chapter 2, with specific details of the stimuli and modifications to test specific hypotheses provided in subsequent chapters.

The questions proposed in this dissertation required me to measure children's ability to categorize stimuli that differ along size and brightness dimensions and subsequent changes in discrimination. However, there was no basis in the current literature for understanding what stimulus values children might be able to differentiate along these dimensions. Thus, Chapter 3 presents a preliminary calibration study that examined 5-to-8-year-old children's perceptual discrimination of metric changes in size and brightness. I then used this information to design the stimuli used in the subsequent experiments.

Chapter 4 presents the first experiment. In this experiment, children were sorted into holistic versus dimensional classifier groups based on their performance in a triad classification task. I then compared these groups' abilities to learn dimensional categories containing stimuli that differ along both relevant (e.g. brightness) and irrelevant (e.g. size) dimensions. I also examine children's ability to perceptually discriminate stimuli that differ along either the relevant or irrelevant dimension.

Chapter 5 presents the second experiment. In this experiment, children were again sorted into groups based on triad classification performance. I then compared the groups' abilities to learn the same categories presented in Experiment 1 in the presence of an incidental label. Subsequently I tested children's discrimination abilities with the same stimuli used in Experiment 1.

Chapter 6 presents a direct comparison of the results of the holistic classifiers in Experiments 1 and 2. This direct comparison allows further examination of the respective contributions of category learning and labels on increases to holistic classifier's selective attention to dimensional similarity.

Chapter 7 presents the results of the third and fourth experiments. These experiments were analogous to Experiments 1 and 2, with the exception that all the participants were classified as "preferrers". That is, these children preferentially selected stimuli that most closely matched the exemplar on one dimension (e.g. size or brightness) on every trial of the triad classification task, regardless of whether the stimulus was a holistic or dimensional match. I compare brightness preferrers' and size preferrers' abilities to learn labeled and unlabeled categories and their subsequent abilities to make discriminations along size and brightness. In addition, category learning, discrimination and subsequent dimensional attention in the post-test triad classification task of children in the preferrer groups was directly compared across the two experiments.

Chapter 8 presents a general discussion of the results of Experiments 1-4. I also discuss the implications of these results and future directions for research on the role of labels and word learning in the development of dimensional attention.

CHAPTER II: GENERAL METHODS

Participants

Participants for all experiments were monolingual English-speaking children between 5- and 8-years-of-age. Previous research (e.g. Smith & Kemler, 1977) has shown that most children at the lower end of this range are not yet dimensional classifiers but that many at the upper end can attend dimensionally. Thus I examined children within this large age range to be able to compare children from a variety of developmental levels with varying attentional abilities. Informed consent was obtained from children's parents or guardians prior to the session. Children were recruited via birth records and received a small toy for participation after each session.

Stimuli and apparatus

The stimuli were a set of squares that vary metrically along size and brightness, similar to those used in Goldstone's 1994 study (see Figure 2). They can be divided into categories such that either size or brightness is the critical dimension. The particular values used were determined by a preliminary calibration study. All stimuli were presented on a pc with a 114.3cm touch screen monitor using Eprime 2.0 software. Children made responses by touching the computer screen.

Discrimination task

The discrimination task measured children's ability to distinguish between close values on test dimensions (see Figure 3). Stimuli were squares that vary in size and brightness. Children sat in front of a touch-screen computer and were presented with a target and two test stimuli. The three stimuli were arranged in a triangle such that the target was at the bottom of the screen and the test stimuli were on top. The children had to indicate which of the test stimuli matched the target by touching it. All three stimuli were present until the children made their response. The target matched one of the test stimuli on every trial. The left/right position of the matching test stimulus was counterbalanced across trials.

Past experiments examining the effects of categorization on perceptual discrimination have used different tasks, some of which have later been shown to introduce varying degrees of categorical bias because of the extra memory demands they create (Pisoni & Tash, 1974; Hanley & Roberson, 2011). The method used in the present experiments, known as the ABX method, does show a bias in the speech domain, but preliminary work suggest a lack of bias with the stimuli used here; likely due to temporal differences in the presentation of auditory and visual stimuli. A third task, known as the AXBX task, where stimuli are presented in two pairs and participants have to decide which pair contained identical stimuli, has been shown to reduce categorical bias in both speech and color domains. However, this method has never been used with children and seems potentially difficult for children to comprehend. Thus the ABX method is the most child-friendly procedure that reduces memory bias.

Triad task

The triad task measures holistic versus dimensional attention in similarity classification (see Figure 4). Children were presented with three stimuli: an exemplar, a holistic match, and a dimensional match, and asked to pick which test stimulus went with the exemplar (see Smith & Kemler, 1977). Children indicated their choice by touching the stimulus. The size and brightness of the square stimuli were varied so that the dimensional match was the same size or brightness value as the exemplar but two steps away on the other dimension, and the holistic match was one step away on both dimensions. On half the trials the dimensional match was the same size as the exemplar, and on the other half of the trials it was the same brightness value. By comparing children’s accuracy on both dimensions I was able to examine how categorization and word learning influence attention to relevant and irrelevant dimensions.

Categorization task

In the categorization task children were presented with a square, told it is a rock, and asked to sort it into one of two categories by selecting where it comes from (ocean or jungle). Children indicated their response by touching the appropriate place on the screen (see Figure 5). Geographical assignment of categories (ocean or jungle) and left/right assignment of response pictures were counterbalanced across children. During the learning phase, children received feedback regarding accuracy of each decision. Correct categorization was followed by a bell sound and visual presentation of the stimulus and the correct location (ocean or jungle) directly beneath the exemplar. The experimenter then told the child “Yeah, good job, that did come from the [ocean/jungle]!” Incorrect categorization was followed by a buzzer sound and visual presentation of the stimulus and the correct location. The experimenter then told the child, “Uh oh, that one actually came from the [ocean/jungle]!”

Learning blocks were made up of 8 trials, 4 trials for each category. Children received a sticker after completion of each block, regardless of performance, to keep them engaged with the task. The learning criterion was accurate categorization on 7 out of 8 trials per block, 2 blocks in a row. If a child did not reach criterion after 30 blocks, the learning phase of the categorization task ended.

The learning phase was followed by a test phase. Stimuli used at test included the 4 exemplars from each category that were presented during category learning and 6 novel exemplars from each category (see Figure 6). The test phase consisted of 4 blocks of 20 trials. At test, no feedback was given after each trial. Children again received a sticker, regardless of performance, after each block.

Approach

Assignment of participants to experiments

Participants were 129 monolingual-English speaking 5-8-year-old children. Six children did not complete the entire experimental procedure (5 quit, 1 equipment error) and were excluded from analyses. This left 123 children in the final group (70 females; *M:* 6 years, 10 months; range: 5 years, 0 months to 8 years 11 months). Participants were assigned to one of the four experiments based on their performance in the triad pretest. The first 16 children who were holistic classifiers and the first 16 who were dimensional classifiers (based on pretest performance) were assigned to Experiment 1 (No Label), while the second 16 of each classifier type were assigned to Experiment 2 (Label). The first 16 children who were size preferrers and the first 16 who were brightness preferrers were assigned to Experiment 3 (No Label), while the remaining 16 children who were size preferrers and 11 children who were brightness preferrers and were assigned to Experiment 4 (Label). In each experiment, half of the children within a classifier group were randomly assigned to the brightness category learning group and the other half to the size category learning group. See Table 1 for participant information for each group in each experiment.

Analyses

I used mixed linear regression models to analyze the speed of category learning, and mixed logistic regression for all other tasks that used forced choice measures (e.g. dimensional versus holistic match, ocean versus jungle, and left stimulus versus right stimulus). I took this approach because recent arguments suggest ANOVA's on categorical outcome variables such as mine are inappropriate (see Jaeger, 2008). Additionally, these models enable control for any potential random effects due to individual differences between children or specific stimuli. This is especially important in analyzing the data of young children who are notoriously noisy and variable. I removed collinearity from the models by sum-coding data and scaling continuous variables. I began with a completely specified random effects structure including random slopes for all variables included in a given model. Using model comparison I systematically removed uninformative random effects to find an appropriate model (c.f. http://hlplab.wordpress.com/2009/05/14/random-effect-structure/). All final models included random intercepts for subject and items, unless otherwise specified. The models of performance in each task were identical in each experiment, and thus I describe their structure here.

In each experiment, to assess differences in the number of blocks it took children to reach criterion for category learning, I created a mixed linear regression model of the interaction between classifier type (e.g. holistic versus dimensional) and category structure (brightness versus size). This model included random subject effects (but no random item effects as I was examining overall time to criterion rather than accuracy on each trial).

All other models in each experiment were done separately for each classifier type and category learning type (e.g. only holistic classifiers who learned brightness categories) so that I could explore qualitative patterns in each group's performance separately. These mixed logistic regression models each compared a given group's likelihood of a given binomial response (e.g. correct versus incorrect) on various trial types (e.g. between, within-relevant, or within-irrelevant discriminations). Some of these models of more in-depth measures of trial type accuracy included an interaction between two factors related to trial type. In particular, the general structure of each of these models was:

Dependent Variable ~ Trial Type x optional second trial type factor + random subject and items effects

Graphical representation of data

Results are depicted graphically with bar graphs. This will give the reader an idea of what the data look like. However, there is a limitation in this depiction in that bar graphs of, for example, proportion correct on a given measure, depict mean data collapsed across participants. Thus, these graphs do not capture the fact that the analysis took into account random subject and item effects and every point of data. For this reason, I cannot depict a measure of variance such as standard error on the graphs. To do so would be misleading because, again, such measures capture average variance whereas the models accounted for the variance of individual participants across trials. Nevertheless, to help the reader integrate statistical significance information from the models with the depictions of the data, asterisks between bars denote significant differences between children's performance on various measures.

CHAPTER III: PRELIMINARY CALIBRATION STUDIES

Calibration Study A

The purpose of Calibration Study A was to determine size of metric brightness and size changes to use be used in the main experiments.

Method

Participants

Thirty-nine 5-8-year-old children (24 females, average age: 6 years, 9 months; range: 4 years, 11 months – 8 years, 10 months) participated.

Procedure

Children completed the discrimination task with both the size and the brightness stimulus sets. In the discrimination task, I used a staircasing procedure similar to that used in prior studies with children (Simmering & Spencer, 2008). If a child made two accurate discriminations on two trials in a row, the similarity distance between the stimuli was decreased by one step. If the child made an inaccurate discrimination, then the distance increased by one change unit. The number of stepwise changes by which the discrimination stimuli differed was adjusted until a child oscillated between two distances that were two steps apart. Each change in size was equal to an increase of .07 cm in both width and height (as in Goldstone, 1994b). Each change in brightness was equal to an increase of 2 units in each red, blue, and green color space. Threshold information was collected from each child for each dimension. I found the 75th percentile for discrimination between stimuli along both the size and brightness dimensions. These numbers were 6 steps (or an increase of 12 in RGB) for brightness and 16 steps (or an increase of 1.12 cm in both width and height) for size.

Calibration Study B

The purpose of Calibration Study B was to 1) ensure that I knew about any *a priori* differences in perception along these dimensions related to children's performance in a similarity classification task and 2) to get a sense of how prevalent each of the four relevant similarity classification types (i.e. holistic classifiers, dimensional classifiers, brightness preferrers, and size preferrers) were within my chosen age range.

Method

Participants

Twenty-six 5-8-year-old children (12 females, average age*:* 6 years, 5 months, range: 5 years, 0 months to 8 years, 2 months) participated.

Triad Stimuli

I created triad stimuli in the same manner as Smith & Kemler 1977 (see Figure 1) such that one step in similarity space was equal to the 75th percentile of perceptual discrimination found in Calibration Study 1.

Procedure

Children completed a triad pretest and the discrimination test using the same staircasing procedure as in Calibration Study A. Performance on the triad pretest was used to assign children into one of four groups according to dimensional attention abilities: holistic, dimensional classifier, brightness preferrer, or size preferrer. Dimensional classifiers had more than 13 dimensional matches on the triad pretest (out of 24 trials). Holistic classifiers made fewer than 12 dimensional matches with no indication of a preference for either dimension. The remaining children fell into preferrer category. These children demonstrated dimensional attention to one, but not the other, dimension. Brightness preferrers made dimensional matches on at least 11 of the 12 brightness match trials and no more than 1 dimensional match on the size match trials. Size preferrers made dimensional matches on at least 11 of the 12 size match trials and no more than 1 dimensional match on the brightness match trials.

Of the twenty-six children who completed the triad pretest, 4 were dimensional classifiers, 6 were brightness preferrers, 8 were size preferrers, and 8 were holistic classifiers. There were no significant differences between classification groups or between age groups in discrimination thresholds. Thus I used the 75th percentile of all children for each dimension as the final step size between category stimuli.

CHAPTER IV: EXPERIMENT 1

The purpose of Experiment 1 was to examine the influence of category learning on dimensional attention. Thus, I compared how children's varying abilities to attend to dimensional similarity in classification relates to their abilities to learn dimensional categories and discriminate stimuli along relevant and irrelevant dimensions.

Experiment 1

Method

Participants

Participants were 33 5- to 8-year-old children. One child did not complete the entire experiment and was excluded from these analyses. Thus there were 32 children in the final group.

Procedure

The children first completed a triad pretest. Children were then divided into classifier groups based on performance in this task. There were 16 holistic classifiers and 16 dimensional classifiers. See Table 1 for participant information.

Categorization

Children next completed the category-learning task and category test. For category learning, half of the children in each classification group were trained with categories organized by similarity in size (size learners), half with categories organized by similarity in brightness (brightness learners). Category spacing was based on discrimination thresholds from the calibration study: each exemplar was one threshold unit away from the next stimulus in brightness and size similarity. Trained exemplars for the brightness learners are shown in Figure 6a and trained exemplars for the size learners are shown in Figure 6b. As can be seen in the Figure, novel stimuli probed at test were sampled from both in- and outside of the learned category space. This allowed me to examine whether children were attending to the category-relevant dimension or to specific values along that dimension.

Discrimination

Next, children completed the discrimination task. The procedure was identical to the preliminary calibration study with the following exceptions: 1) stimuli were selected from the category space used in the category-learning task and 2) stimuli did not change in brightness or size during the task based on children's performance. Instead, discrimination abilities were probed at .5, 1, and 1.5 times the previously-defined perceptual threshold for each dimension (from the calibration study). Discrimination was tested both within and across category boundaries (see Figure 6). All pairs were presented four times—each stimulus within a pair was presented twice as target and matching test item, and twice as foil—for a total of 96 trials. Left/right presentation of match/foil was counterbalanced across trial types for each pair. Both size-learners and brightness-learners were presented with the same pairs, such that any given trial forced children to make a discrimination along a dimension that was only relevant for one group's learned category. For example, a stimulus pair that differed only in size would test discrimination along the relevant dimension for the size-learners and the irrelevant dimension for brightness-learners. This allowed an examination of changes in children's ability to make both discriminations along category-relevant and irrelevant dimensions and discriminations both within and between categories.

<u>Strongly versus weakly learned areas:</u> Locations of trained exemplars were chosen to allow for some areas of the category space to be more strongly-learned (e.g. between stimuli 2-4 and 2-5 in figure 6a) and other areas to be more weakly-learned (e.g. between stimuli 7-4 and 7-5 in figure 6a). This provided an assessment (with the discrimination task) of whether children who learned the categories did so by learning

about individual exemplars or regions of the category space, or if they learned to attend to the relevant dimension. If children were more accurate at making discriminations between stimuli in strongly-learned areas of the space than between stimuli in weakly-learned areas of the space, this would suggest that they did not learn to attend a whole dimension (e.g. size) but instead learned to associate regions of the category space with the category response.

Blocking: The order of presentation of discrimination trials was blocked so that children alternated between a block of 12 discriminations along the dimension that was relevant for their category task (i.e. single type relevant block), a block of 12 discriminations along the irrelevant dimension (i.e. single type irrelevant block), and a block of 24 trials alternating between discriminations on relevant and irrelevant dimensions (i.e. mixed block). This blocked structure provided an additional measure of changes in dimensional attention related to category learning. If children show a discrimination advantage in the block of discriminations on the relevant dimension relative to the block of discriminations on the irrelevant dimension, it would suggest that learning a categorical distinction based on that dimension increased attention to that dimension. If children also show a discrimination advantage in a relevant block over the relevant-dimension discriminations in a mixed block, it would suggest that once children increased attention to a given dimension, there was a cost to switch attention to another dimension. If, however, children show no advantage for one type of block over another, it would suggest that their attention towards each dimension had not been affected, and that they were attending to overall similarity rather than dimensional similarity to discriminate between stimuli. Alternatively, although less likely, a lack of difference in accuracy between blocks could mean that children are actually very skilled, flexible switchers.

Posttest triad task

The posttest triad task was identical to the pretest version.

Predictions

Categorization

It was expected that children who classified dimensionally in the pretest would be able to learn both the size and brightness categories and generalize to novel category exemplars. However, children who classified holistically in the pretest should have more difficulty. While they should still be able to learn to categorize trained exemplars, they should have more trouble with novel exemplars. Correct categorization of novel exemplars should require children to have learned something about category organization rather than something about specific stimuli. As can be seen in Figure 6, novel stimuli were drawn from both inside and outside of the category space. Novel stimuli from inside the category were overall quite similar to learned exemplars, but they were also quite similar to stimuli on either side of the category boundary. Thus, these stimuli could be harder to categorize correctly for children who had relatively weak selective attention to the relevant dimension and tended to attend more to overall similarity.

On the other hand, novel exemplars from outside the category space were, by definition, quite different from the stimuli at the category boundary. These stimuli were only similar to the trained exemplars along one dimension – the dimension critical for the category. Thus, correct categorization of novel exemplars from outside the category space should require children to attend to the category-relevant dimension rather than to overall similarity or specific learned values along the relevant dimension. However, because these stimuli were, by necessity, farther from the category boundary, it is also possible that much less attention to relevant dimensions is required to correctly categorize novel stimuli from outside the category space.

Discrimination

I predicted that children who were dimensional classifiers on the triad pretest would show worse within-category discrimination along the irrelevant dimension and enhanced between-category discrimination on the relevant dimension than their counterparts. They should also show a switch cost for trials in the mixed block. However, because Goldstone had previously found that adults who learned brightness categories, but not size categories, showed enhanced discrimination along the relevant dimension and poorer discrimination along the irrelevant dimension, I also predicted that the advantage for relevant discriminations over irrelevant discriminations might only be found for dimensional classifiers who learned brightness categories. Because it was predicted that holistic classifiers would learn dimensional categories not by selectively attending to category-relevant dimensions, but instead by attending to the overall similarity between the stimuli, they should not demonstrate differences in attention to relevant versus irrelevant dimension. Within the context of the discrimination task, this means that children who were holistic classifiers on our triad pretest, should not show differences in accuracy for relevant versus irrelevant discriminations. Additionally, these children should show no differences in performance for blocked versus mixed trials in the discrimination task.

Posttest triad

I predicted that individual classification patterns would not generally differ from those at pretest. Children who were dimensional classifiers on the triad pretest should still classify dimensionally as learning dimensional categories should not decrease their ability to selectively attend to dimensional similarity. Children who were holistic classifiers on the triad pretest should still classify holistically. This is because it was expected that holistic classifiers will not be able to selectively attend to category-relevant dimensions, the category learning should not change their attention to dimensional similarity in the posttest triad task.

Results and discussion

My primary questions in this experiment concerned how classifier type (dimensional versus holistic) influenced dimensional category learning and changed subsequent discrimination abilities on category-relevant versus category-irrelevant dimensions and subsequent similarity classification abilities. Thus to address these questions, I first examined the results of category learning, overall results of the discrimination test, and the results of the posttest triad task. Then, in order to more deeply examine the extent to which children were selectively attending to dimensional similarity in each of these tasks, I next examined the results of category test and the more in-depth measures of discrimination, including accuracy in strongly- versus weakly-learned areas, accuracy across different threshold sizes, and accuracy in the context of mixed versus blocked trials.

Category learning

If selective attention to dimensional similarity is necessary for learning categories organized by a single dimension, then dimensional classifiers should be faster to learn categories than holistic classifiers. To examine this I measured the number of blocks to criterion for each child in the category-learning task. These data are presented in Figure 7. As can be seen in the figure, dimensional classifiers took fewer blocks to reach criterion than holistic classifiers, and both groups took fewer blocks to reach criterion if they learned categories organized by brightness rather than categories organized by size. A mixed linear regression model of the interaction between classifier type (holistic versus dimensional) and category structure (brightness versus size organization) on the number of blocks it took children to reach criterion, revealed a significant effect of classifier type such that dimensional classifiers were faster to reach criterion than holistic classifiers,

t=21.88, p<.0001.[1] This model also showed a significant effect of category structure such that children were faster to reach criterion when learning categories organized by brightness, t =-23.00, p<.0001. This advantage for learning categories organized by similarity in brightness replicates Goldstone's (1994b) finding with adult participants. There was also a significant interaction between classifier type and category structure such that holistic classifiers showed less of an advantage for learning brightness categories than dimensional classifiers, t=2.66, p<.01.

Thus, the data appear to support my prediction; children who were able to selectively attend to a single dimension were able to learn a novel category distinction based on attention to one dimension faster than children who were holistic attenders. However, because the dimensional classifiers were generally older than the holistic attenders, it is possible that age or another developmental factor could be the basis for these results. To test this possibility, I created additional models to assess whether age, gender, or SES (as measured by maternal education), rather than classification group, might have driven the difference in learning. Model comparison revealed that classifier type was necessary to account for the findings, whereas age, sex, and SES were not. A model with both classifier type and age was significantly better than one with only age, $X^2(1)$ =12659, p<.0001, but no different from one with only classifier type. A model with both classifier type and sex was significantly better than one with only sex, $X^2(1)$ =10665, p<.0001, but no different from one with only classifier type. Finally, a model with both classifier type and SES was significantly better than one with only SES, $X^2(1)$ =14095, p<.0001, but no different from one with only classifier type.

In addition, to check that the results were not due to the particular criterion I used to assign children to classifier groups, I also ran a model using a continuous measure of

[1] Because of the difficulty in determining degrees of freedom in mixed linear models, we conducted MCMC sampling to find p-values (see Baayen et al., 2008).

children's dimensional responses on the triad task. This model used the interaction between the proportion of dimensional responses children made on brightness match trials and the proportion of dimensional responses they made on size match trials to predict the number of blocks it took children to reach criterion. Importantly, comparing the log likelihood values of the two models revealed that the likelihood of the categorical model accurately capturing the data was higher than the likelihood of the continuous model accurately capturing the data. Thus, grouping children in terms of their performance in the triad classification task is the most appropriate way to capture category learning performance.

Overall, then, the results of category learning fit my predictions. Holistic classifiers were slower to learn dimensional categories than dimensional classifiers. In fact, 4 out of 16 children in the holistic group failed to reach criterion at all and thus completed the maximum number of blocks (30), while only 1 out of 16 children in the dimensional group failed to reach criterion. A critical question, however, is whether the holistic classifiers who managed to learn the categories did so in different ways than the dimensional classifiers. For example, did the holistic classifiers learn the categories by attending to the overall similarity between specific stimuli, while the dimensional classifiers increased attention to the category relevant dimension and decreased attention to the category irrelevant dimension? To answer this question, I next examined children's performance in the categorization test, discrimination task and posttest triad task. These tasks were designed to measure any changes in selective attention to category-relevant and irrelevant dimensions. Thus differences in the classifier groups' performance in these tasks should clarify whether the different groups of children who learned the dimensional categories did so using the same or different mechanisms. These analyses included only data from those children who reached criterion on category leaning. I report results of each classification group separately for two reasons. First, overall, all the participants did very well on the generalization and discrimination tests; performance was above 50%

across the board. Second, I am more interested in whether dimensional and holistic classifiers both show enhanced between-category discrimination, for example, than whether dimensional classifiers are more accurate than holistic classifiers. Thus, assessing the results of the two groups separately will allow me to find differences in their qualitative patterns of results. In addition, because there were differences in the learning of categories organized by similarity in brightness and size that correspond to differences previously described in the literature (e.g. Goldstone 1994b), I examined performance of children who learned each category structure separately.

Discrimination

The discrimination task was designed examine whether or not children of each classifier type had learned to selectively attend to the category-relevant dimension and ignore the category irrelevant dimension. To the extent that a child had learned to attend to the category-relevant dimension, she should show increased accuracy for discriminations along the relevant dimension—particularly for between-category discriminations—compared to those along the irrelevant dimension. Dimensional classifiers should already be able to selectively attend to category-relevant dimensions; therefore I predicted they would show increased accuracy for relevant discriminations and decreased accuracy for irrelevant discriminations. If holistic classifiers show this pattern, then it would suggest that their selective attention was supported by the category learning task. Goldstone found that adult participants demonstrated enhanced accuracy for discrimination across the category-relevant dimensions and decreased accuracy for discrimination across the category-irrelevant dimension if they had learned brightness categories but if they had learned size categories. Thus it is expected that the dimensional classifiers should also show this pattern for brightness but not size.

To examine changes to children's selective attention to category-relevant and irrelevant dimensions, I measured children's accuracy for each type of discrimination: between-category, within-category along the relevant dimension and within-category along the irrelevant dimension. Overall, children were quite accurate in discriminating stimuli. In particular, the dimensional classifiers who learned brightness categories, M=.82, $t(7)$=15.71, p<.0001, the dimensional classifiers who learned size categories, M=.81, $t(6)$=6.27, p<.001, the holistic classifiers who learned brightness categories, M=.74, $t(6)$=5.96, p<.001, and the holistic classifiers who learned size categories, M=.73, $t(4)$=5.04, p<.01, were all significantly better than chance (.50) at discriminating stimuli in this task. As can be seen in Figure 8a, dimensional classifiers who learned brightness categories were more accurate at between-category discriminations than within-category discriminations. This was supported by a mixed logistic regression model of the effect of discrimination comparison type (between, within relevant, within irrelevant) on discrimination accuracy which showed that dimensional classifiers in the brightness learning group had significantly increased discrimination accuracy when making between category discriminations compared to within category discriminations on either the irrelevant, z= -3.91, p<.0001 or relevant dimension, z=-3.84, p<.001. However, these children were also significantly more accurate at making within-category discriminations along the relevant than the irrelevant dimensions, z=-3.70, p<.001, demonstrating that, overall, they were more accurate at discriminating across the relevant than the irrelevant dimension. However, as can also be seen in the Figure, dimensional classifiers who learned size categories were no more accurate at making between-category discriminations than within-category discriminations on either the irrelevant, z=.25, *NS,* or the relevant dimension, z=-.98, *NS.* These children were not any more accurate at making one kind of within-category discrimination over another, z=.70, *NS.*

If holistic classifiers were unable to selectively attend to category relevant dimensions, they should not show the same discrimination abilities as dimensional classifiers. However, as can be seen in Figure 8b, holistic classifiers who learned brightness categories performed like dimensional classifiers, although with slightly lower overall accuracy. Holistic classifiers were more accurate at between-category discriminations than within-category discriminations. This was confirmed by a mixed logistic regression model showing that holistic classifiers who learned brightness categories were more accurate at between category discriminations than within category discriminations along both the irrelevant dimension, z=-2.33, p<.05, and the relevant dimension, z=-1.93, p<.05. However, these children were significantly more accurate at making within-category discriminations along the relevant than the irrelevant dimensions, z=-2.11, p<.05, demonstrating that, overall, they were more accurate at discriminating across the relevant than the irrelevant dimension. As can also be seen in Figure 8b, holistic classifiers who learned size categories were no more accurate at making between-category discriminations than within-category discriminations along either the relevant, z=-.68, *NS*, or irrelevant dimensions, z=-.39. These children were not any more accurate at making one kind of within-category discrimination over another, z=.01, *NS*.

Thus, both types of classifiers showed evidence of an effect of category learning on their perceptual discrimination abilities—but only if they learned a dimensional category for which brightness was the critical dimension. Importantly, this effect was such that overall, the brightness learners were more accurate at making discriminations along the relevant dimension suggesting that they had learned something about dimensions rather than just categories.

Posttest triad

I next examined the results of the posttest triad task. The primary question of interest is whether there was a change in the number of dimensional matches children choose from pre- to posttest. The extent of such a change indicates the extent to which learning dimensional categories—which require the learner to selectively attend to one dimension to the exclusion of another—increased children's selective dimensional attention in similarity classification. To examine this, I measured the number of dimensional responses children made in both pre- and posttest triad tasks. It was predicted that the dimensional classifiers would not show an increase in dimensional responding from pre to post test because they were already attending dimensionally. As can be seen in Figure 9a, the dimensional classifiers who learned brightness categories did not increase the number of dimensional choices they made. This was supported by the results of a mixed logistic regression model that showed that dimensional classifiers who learned brightness categories were no more likely to choose dimensional matches during the posttest compared to the pretest, z=-.81 *NS.* Similarly, as can also be seen in the Figure, dimensional classifiers who learned size categories were no more likely to choose dimensional matches during the post test, z=-.15, *NS*.

The holistic classifiers, on the other hand, were also not predicted to show an increase in dimensional responding but for a different reason—because they were not expected to be attending dimensionally in the category-learning task. As can be seen in Figure 9b, the holistic classifiers who learned brightness categories were no more likely to choose dimensional matches during the posttest compared to the pretest. This is supported by the results of a mixed linear regression model showing that holistic classifiers who learned brightness categories showed no increase in dimensional responding, z=.78, *NS.* As can be seen in the right panel of Figure 9b, however, those who learned size categories *were* marginally more likely to select dimensional matches during the post test than the pre test, z=1.94, p<.10. This result was unexpected but suggests that perhaps these children did learn to selectively attend to size during category learning. As can be seen in Figure 10b, holistic classifiers initially showed a slight preference in the pretest for selecting dimensional matches on brightness trials. The brightness categories were much easier for both holistic and dimensional classifiers to learn, suggesting that brightness is an easier dimension to selectively attend to than size. This could mean that even if the holistic classifiers who learned brightness categories did learn to attend to brightness during category learning, they would not show an increase in dimensional responding—because they already had a slight preference for attending to brightness. Those holistic classifiers who learned size categories, on the other hand, could have shown an increase in dimensional responding because they learned to attend to size in category learning, which would lead to a change in their pattern of dimensional responses in the posttest triad.

An important secondary question, then, is whether the particular category type learned influenced children's choice of dimensional matches on triad task trials that probed the critical dimension for the category they had learned. I next examined the extent to which learning to selectively attend to a given dimension in categorization changed the specific dimension to which children selectively attended in similarity classification. Even though they did not show an overall increase in dimensional responding, I expected dimensional classifiers to select dimensional matches on the trials corresponding to the dimension relevant for the category they learned because they should have been primed to attend to that dimension after having done so during category learning. As can be seen in Figure 10a, dimensional classifiers who learned brightness categories were more likely to select dimensional matches on brightness trials than they were on the pretest. This was supported by the results of a mixed logistic regression model of the interaction between trial type (size or brightness) and test time (pre- or posttest triad) on the likelihood of selecting a dimensional match. The model revealed that dimensional classifiers who learned brightness categories were significantly more likely to choose dimensional matches on brightness trials than size trials, $z=-2.68$, $p<.01$, and there was a significant interaction such that they were even more likely to do so on the posttest trials, $z=-2.01$, $p<.05$. As can also be seen Figure 10a, dimensional classifiers who learned size categories, on the other hand, were more likely to choose dimensional matches on posttest size trials than they were during the pretest. This was supported by the results of a mixed logistic regression model that showed dimensional classifiers who learned size categories were significantly more likely to choose dimensional matches on size trials, $z=-2.85$, $p<.01$, and that this interacted with test time such that they were more likely to do so on the posttest, $z=1.92$, $p<.10$. Together, these results suggest that dimensional classifiers learned to attend more to the dimension that was relevant to the categories they learned.

It was predicted that the holistic classifiers would not choose dimensional matches more on trials corresponding to the dimensional structure of the category they learned. Holistic classifiers had more difficulty learning the categories relative to dimensional classifiers—presumably because they had greater difficulty attending selectively to dimensional similarity. Therefore, despite the fact that holistic classifiers (who learned brightness categories) did show some evidence of selective attention in the discrimination test, I did not predict that this would be enough of an increase to carry over to the triad task and result in changes to their classification. As can be seen in Figure 10b, the holistic classifiers who learned brightness categories were more likely to choose dimensional matches on brightness trials during both pre- and posttest. A mixed logistic regression model of the interaction between trial type and test time showed that holistic classifiers who learned brightness categories were more likely to select dimensional matches on brightness trials, z=-6.04, p<.0001, however, there was no interaction, z=-1.23, *NS*. This suggests that during both pre- and posttest triad holistic classifiers were more likely to select dimensional matches on brightness trials. Furthermore, as can be seen in the right panel of Figure 10b, holistic classifiers who learned *size* categories were also more likely to select dimensional matches on brightness trials. This was supported by a mixed logistic regression model showing that these children were significantly more likely to select dimensional matches on brightness trials, z=-6.42, p<.0001, and a significant interaction such that holistic classifiers who learned size categories were even more likely to select dimensional matches on posttest brightness trials than pretest ones z=-3.75, p<.001. This main effect suggests that these children did not learn to attend selectively to the dimension that had been relevant for categorization (size), and suggests instead that these children generally preferred to attend to similarity in brightness regardless of the structure of the category they learned. More than that, though, the interaction suggests that somehow the category learning task increased their selective attention to the irrelevant dimension—brightness.

Together, the results of the triad task show that dimensional classifiers' similarity classification was influenced by the categories they had learned earlier in the experiment. While dimensional classifiers did not show an increase in the overall proportion of dimensional responses they made from pre- to posttest, they did show a change in the particular dimension to which they paid the most attention. Thus, they clearly utilized selective attention to learn their categories and their later attention was influenced by the specifics of those categories. Holistic classifiers' similarity classification, on the other hand, was not directly influenced by the categories they had learned earlier in the experiment. It is still unclear, then, whether or not they were utilizing selective attention to learn their categories. Therefore, the next series of analyses take a more in-depth look at the extent to which children were selectively attending to dimensional similarity during the experiment.

<u>In-depth analysis of dimensional attention</u>

The results of category learning task supported my prediction that holistic classifiers should have more difficulty than dimensional classifiers learning dimensional categories. However, the results of the discrimination task showed that both dimensional and holistic classifiers who learned brightness categories demonstrated a pattern of discrimination accuracy indicative of attention to category-relevant dimensions. The results of the posttest triad, on the other hand, suggest that the dimensional classifiers increased attention to category relevant dimensions while the holistic classifiers did not. Thus, the extent of each group's selective attention is not clear. If the holistic-to-dimensional shift is actually more of a continuous increase in selective flexible attention over development, rather than a discrete shift, then it is entirely possible that even holistic classifiers can—with the help of learned category distinctions—selectively attend to one dimension and ignore another. This could mean that the holistic classifiers showed enough of an increase in dimensional attention that it could influence their accuracy in the discrimination task, but not enough that it could influence their overall preference for holistic similarity relationships in the triad task.

To clarify this issue I next examined children's accuracy in categorization and discrimination with respect to stimulus type. The category test addressed whether or not children are more accurate at categorizing familiar than novel stimuli, and whether they are more accurate at categorizing novel stimuli from within the category space than outside the category space. These two questions will clarify the extent to which children learned to pay attention to the category-relevant dimension or to specific stimuli. The discrimination test addressed whether or not children's discrimination accuracy was influenced by 1) where in the learned category space the comparison stimuli were sampled, 2) the size of difference between stimuli, and 3) whether the discrimination occurred within a mixed or blocked set of trials. Examining these issues together will clarify the process by which children learned the categories: whether by selectively attending to a category-relevant dimension or by attending to the overall similarity of the stimuli

Category test

Overall, children in both groups who reached the learning criterion were very accurate in the categorization test. The dimensional classifiers who learned brightness categories, $M=.87$, $t(7)=41.61$, $p<.0001$, the dimensional classifiers who learned size categories, $M=.77$, $t(6)=4.98$, $p<.01$, the holistic classifiers who learned brightness categories, $M=.83$, $t(6)=6.33$, $p<.001$, and the holistic classifiers who learned size categories, $M=.70$, $t(4)=3.01$, $p<.05$, were all significantly better than chance at categorizing the test stimuli. If children learned these categories by selectively attending to the relevant dimension and ignoring the irrelevant dimension, then they should be equally accurate at categorizing familiar and novel exemplars at test. Additionally, they should be equally accurate at categorizing novel exemplars both inside and outside of the learned category space. If, however, children learned these categories by memorizing individual stimuli or learning about specific local regions of the stimulus space (rather than learning about a *dimension*), they should be more accurate at categorizing familiar exemplars relative to novel exemplars and might be worse at categorizing novel stimuli from inside of the learned category space than those outside. I predicted that the dimensional classifiers should show evidence of attending to dimensions rather than specific stimuli. The holistic classifiers, however, should show evidence of attending to specific stimuli rather than dimensions.

To address this I first examined children's accuracy categorizing familiar (presented during category learning) versus novel stimuli. As can be seen in Figure 11a, the data suggest, as predicted, that dimensional classifiers (right set of bars) who learned brightness categories categorized familiar and novel stimuli equally accurately, z=.61, *NS.* Likewise, dimensional classifiers who learned size categories categorized familiar and novel stimuli with equal accuracy as well, z =.54, *NS.* These findings could be taken to suggest that the dimensional classifiers learned something about the category's rule rather than something about specific stimuli.

The holistic classifiers, on the other hand, were predicted to show differences in accuracy for novel and familiar stimuli. As can be seen in Figure 11b, however, the holistic classifiers who learned brightness categories were equally good at categorizing familiar and novel stimuli, z=1.26, *NS.* Likewise, the holistic classifiers who learned size categories were also were equally accurate at categorizing both types of stimuli, z =-.17, *NS.* These results thus suggest that the holistic classifiers also learned something about the category-relevant dimension rather than something about specific stimuli. However, because the null result was found, the possibility still exists that children in both groups learned something about specific regions of the stimulus space, rather than the specific dimension relevant for categorization.

To examine this possibility, I analyzed accuracy of children's categorization of novel stimuli with respect to whether the stimuli were drawn from inside or outside the category space. It was predicted that the dimensional classifiers would either be equally good at categorizing the novel stimuli that fell outside of the trained stimulus space, or potentially even better at categorizing those outside the space as they were maximally dissimilar from the incorrect category. As can be seen in Figure 12a, these children were in fact better at categorizing stimuli from outside the category space. A mixed logistic regression model of the effect of novel trial type on accuracy revealed that dimensional classifiers who learned brightness categories were significantly more accurate at categorizing novel stimuli from outside the trained category space than those inside the trained space z=-2.49, p<.05. In contrast, the dimensional classifiers who learned size categories were not any more accurate at categorizing novel stimuli from outside the category space than inside the category space, z=-.75, *NS*. This suggests that at least the brightness learners were attending to the dimension relevant to category organization.

It was expected that the holistic classifiers, on the other hand, should be worse at categorizing the novel stimuli that fell outside the trained stimulus space relative to those from inside the space because the outside stimuli were very different from the stimuli presented during learning. As can be seen in the left panel of Figure 12b, however, the holistic classifiers who learned brightness categories were equally accurate at categorizing novel stimuli from outside the trained space compared to those from inside the space, z=-1.27, *NS*. As can be seen on the right of Figure 12b, on the other hand, the holistic classifiers who learned size categories, were more accurate at categorizing novel stimuli from outside the category space than inside, z=-2.78, p<.01. This result suggests that the holistic classifiers, at least those who learned size categories, might have learned something about the category relevant dimension.

Results of the categorization test are mixed overall. The dimensional classifiers who learned brightness categories, but not those who learned size categories, appear to have learned something about the specific dimension of category organization. However, for the holistic categorizers, it was the ones that learned size categories, and not the ones who learned brightness categories who appear to have learned something about the specific dimension of category organization—as evidenced by greater categorization accuracy for novel stimuli from outside the learned category space. Assessing the results of the discrimination test with respect to accuracy for discriminations in strongly- versus weakly-learned areas should help clarify the extent to which these two groups were attending to dimensions rather than specific stimuli and regions of the category space.

Strongly versus weakly learned areas of the category space

If differences found in children's accuracy for relevant and irrelevant discriminations were due to children having learned to selectively attend to a whole dimension, then they should be equally good at making discriminations in both strongly and weakly learned areas of the category space. If however, children had learned to attend to overall similarity between exemplars within a category, then they should be better at making discriminations between stimuli from strongly learned rather than weakly learned areas of the category space. To examine this, I analyzed accuracy for between and within-category discriminations in both strongly and weakly learned areas of the category space. A mixed logistic regression model of the interaction between stimulus location (strongly versus weakly learned space) and type (between, within relevant, within irrelevant) showed that dimensional classifiers were equally accurate on discriminations from both strongly and weakly learned areas of the stimulus space, regardless of whether they had learned brightness categories, z=1.25, *NS,* or size categories, z=.38, *NS.* Similarly, a mixed logistic regression model revealed that holistic classifiers were also equally accurate on strongly-learned and weakly-learned discriminations regardless of whether they had learned brightness categories, z=-.21, *NS*, or size categories, z=-.26, *NS.* The results of these models demonstrated that there was no difference between children's accuracy in discriminating stimuli from strongly- versus weakly-learned areas of the space. This provides some evidence that both groups of children who learned the trained categories may not just have learned about specific stimuli or regions of the space, and perhaps instead learned to selectively attend to category-relevant dimensions.

Discrimination step size

Next I examined whether learning to attend to brightness or size for a categorization task changed children's discrimination abilities. Recall that discrimination was tested at .5, 1, and 1.5 times the chosen discrimination threshold. A mixed logistic regression model of the interaction between the step size (.5, 1, or 1.5) and discrimination type (between, within relevant, within irrelevant) showed that dimensional classifiers who learned brightness categories were significantly more accurate at making discriminations when the stimuli were more different, z= 4.42, p<.0001, as were those who learned size categories, z=5.10, p<.0001. Holistic classifiers who learned brightness categories also showed an effect of discrimination step size, z=2.18, p<.05 as did those who learned size categories, z=2.83, p<.01. These results suggest that, overall, children were better at making discriminations as the difference between stimuli increased. The question of interest for analysis was, after learning the categories, are children more accurate at making relevant discriminations even across smaller distances. The answer is clearly no. Accuracy in each discrimination type simply increased with step size.

Blocking

Finally, I examined children's discrimination accuracy with respect to the block of trials where the discrimination occurred. If children really were attending more to one dimension than another, then they should show switch costs when they have to make discriminations along the other dimension. This would lead to decreased accuracy for discriminations along the relevant dimension that occurred within the context of mixed blocks compared to those that occurred within single-type blocks, for example. If however, children are attending to the overall similarity of specific stimuli when they make discriminations, they should not show an advantage for one type of block over another because they do not have to reallocate their attention from trial to trial. To examine this, I measured children's accuracy for between and within-category discriminations in single-type (relevant-only or irrelevant-only) and mixed blocks. As can be seen in Figure 13a, dimensional classifiers who learned brightness categories were slightly more accurate at making relevant between- (blue bars) and within-category (green bars) as well as irrelevant (orange bars) discriminations in the context of single-type blocks than during mixed blocks. This was supported by a mixed logistic regression model of the effects of block on discrimination accuracy showing that when a trial occurred within a block of single-type trials, as compared to a mixed block, dimensional classifiers who learned brightness categories were marginally more accurate at making between category discriminations, z=1.81, p<.10, marginally more accurate at making within category discriminations along the relevant dimension, z=1.86, p<.10, and marginally more accurate at making within category discriminations along the irrelevant dimension, z=1.94, p<.10. Similarly, as can be seen in the figure, dimensional classifiers who learned size categories were significantly more accurate at making between-category discriminations in single-type blocks than in mixed blocks, z=2.05, p<.05. However, those who learned size categories showed no advantage for blocked trials over mixed when making within category discriminations along either the relevant, z=.38, *NS,* or irrelevant dimension, z=.96, *NS.* Overall, these results suggest that the dimensional classifiers were possibly selectively attending to dimensions, although not to the same extent in the two different category-learning conditions.

It was predicted that holistic classifiers would attend to overall similarity when making discriminations and thus should not show an advantage for trials in single-type blocks. In fact, as can be seen in Figure 13b, those who learned brightness categories showed no advantage of block type on between category discriminations, z=.67, *NS*, or within category discriminations along either the relevant, z=-.56, *NS*, or irrelevant dimensions, z=.14, *NS*. Similarly, as can be seen in the figure, those who learned size categories showed no advantage for single-type blocks on between category discriminations, z=.64, *NS,* or within category discriminations along either the relevant, z=.42, *NS,* or irrelevant dimensions, z=.06, *NS.* The finding that holistic classifiers did not show any switch costs in discrimination and could indicate that they had not, in fact, learned to selectively attend to the relevant dimension in categorization. This could suggest that the categorical perception-like effects they appear to have demonstrated in the categorization test could have arisen from their attention to overall similarity of stimuli.

Overall, the results of the discrimination test demonstrate that children's ability to make discriminations along dimensions was affected by category learning—that is if they learned brightness categories. In general, these children demonstrated increased accuracy for relevant versus irrelevant discriminations. However, the results of the analysis of blocking indicate that the two classifier types may have learned the categories in different ways. Specifically, the dimensional classifiers' performance suggests that they were selectively attending to category relevant dimensions while the holistic classifiers' performance suggests that they were attending to overall similarity and thus did not show a cost for making discriminations within the context of a mixed block of trials.

Conclusions

Learning to categorize stimuli along a given dimension increases attention to that dimension and leads to changes in discrimination. For this to happen, however, the learner has to be able to attend to the relevant dimension in the first place. The results of Experiment 1 generally support this idea by demonstrating that holistic classifiers were slower to learn categories organized by a single dimension of similarity. Furthermore, analysis of the blocking structure for the discrimination test suggest that of those children who did learn the categories, only the dimensional classifiers demonstrated increased attention to category relevant dimensions and decreased attention to category irrelevant dimensions. The blocking measure suggests that the holistic classifiers may have learned the category distinction by attending to the overall similarity of individual stimuli within the categories. The fact that the holistic classifiers were equally accurate at making discriminations in single-type and mixed blocks, lends some support to my idea that they are engaging in less selective attention than their dimensional counterparts, making it more difficult for them to learn these category distinctions.

However, the results of the categorization test and the null results of the discrimination test with respect to stimulus location are more mixed and relatively difficult to interpret. These tests do not clearly reveal whether children were selectively attending to dimensional similarity. It is possible they learned the categories by some other means or that these tests were not strong enough measures of attention.

Overall, the results of Experiment 1 generally suggest that children who classify holistically have difficulty learning to attend to one dimension in the category-learning task, and instead might be learning about individual stimuli or regions of the category space, as evidenced by their slow category learning and their lack of a switch cost for discrimination. However, work on early word-learning shows that even very young children can do this in the context of novel noun generalization (Smith et al., 2002). Furthermore, we know that labels can facilitate category learning even in adults (Lupyan et al., 2007). If labels play a scaffolding role in the development of dimensional attention more generally, then we should also see a facilitative effect of labels on category learning in children who are holistic classifiers. In Experiment 2, I examined category learning in the context of redundant labels to assess subsequent changes in attention.

CHAPTER V: EXPERIMENT 2

The results of Experiment 1 demonstrated that, as predicted, children who attend to holistic similarity relationships have greater difficulty learning dimensional categories than children who attend to dimensional similarity relationships. If, over the course of development, word learning supports the emergence of dimensional attention in similarity classification, then the presence of incidental labels should support even holistic classifiers' dimensional attention during category learning. Thus, the purpose of Experiment 2 is to examine whether redundant category labels can increase children's dimensional attention and category-learning accuracy.

Experiment 2

Method

Participants

35 5-8-year-olds participated. Three children did not complete the entire experiment (2 for quitting and 1 for equipment error) and thus were excluded from the analyses. Thus there were 32 children in the final group.

Procedure

General methods were identical to those of Experiment 1. Children were divided into classifier groups based on performance in the triad pretest. There were 16 holistic classifiers and 16 dimensional classifiers. See Table 1 for participant information. However, during category learning, a novel category label was presented immediately after accuracy feedback. The two labels “grecious" and “leebish” were recorded by a female speaker and were selected because they had been used in previous categorization work (e.g. Lupyan et al., 2007). Thus, on each trial, after the child responded they heard either a bell or a buzzer to give them feedback about their accuracy. This was followed by the novel label associated with the stimulus's category (either grecious or leebish). After the presentation of the label, the experimenter said (as in Experiment 1), "Yeah, good job, that did come from the [ocean/jungle]!" if the child had categorized correctly, or "Uh oh, that one actually came from the [ocean/jungle]!" if the child had categorized incorrectly. Assignment of labels to categories was counterbalanced across subjects. Results of each task were analyzed in the same manner as in Experiment 1: first I examine differences in speed of category learning between the classifier types of each category learning group. I then assessed accuracy in the discrimination and posttest triad tasks. Finally, I analyzed the results of categorization and discrimination tests to gain a more in-depth understanding of children's dimensional attention. As in Experiment 1, all analyses (other than that for category learning) were conducted separately for each classifier type and category learning group, and included only children who successfully learned the categories. I directly compare performance of holistic classifiers with and without labels in the next chapter.

Results and discussion

Category learning

If labels increase selective attention to category-relevant dimensions, then holistic and dimensional classifiers should be equally fast to learn dimensional categories in this Experiment. To examine this I measured the number of blocks to criterion for each child in the category-learning task. As can be seen in Figure 14, the holistic classifiers learned the categories as quickly as the dimensional classifiers, and both groups learned brightness categories as quickly as they did size categories. This was supported by a mixed linear regression model of the interaction between classifier type (holistic versus dimensional) and category structure (brightness versus size organization) on the number of blocks it took children to reach criterion. This analysis revealed that both groups of children learned the categories equally quickly as there was no effect of classifier type, t=.04, *NS,* and no effect of category structure, t =-.30, *NS*.

Thus, the data appear to support my prediction; labels facilitate the learning of dimensional categories, such that holistic classifiers were now as quick to reach criterion as dimensional classifiers. In fact, only 1 out of 16 children in the holistic group failed to reach criterion and thus completed the maximum number of blocks (30), and none of the 16 children in the dimensional group failed to reach criterion. A critical question, however, whether the presence of labels facilitated category learning for the holistic classifiers because it increased their selective attention to category-relevant dimensions. To answer this question, I next examined children's performance in the categorization test, discrimination task and posttest triad task.

Discrimination

The discrimination task was designed to further address the question of whether or not children of each classifier type had learned to selectively attend to the category relevant dimension and ignore the category irrelevant dimension. Overall, children were quite accurate in discriminating stimuli. In particular, the dimensional classifiers who learned brightness categories, M=.83, $t(6)$=16.40, p<.0001, the dimensional classifiers who learned size categories, M=.80, $t(6)$=10.90, p<.0001, the holistic classifiers who learned brightness categories, M=.76, $t(6)$=7.42, p<.001, and the holistic classifiers who learned size categories, M=.77, $t(7)$=6.27, p<.001, were all significantly better than chance at discriminating stimuli in this task.

It was predicted that, as in Experiment 1, dimensional classifiers who learned brightness categories would show increased between category discrimination and decreased within category discrimination. As can be seen in Figure 15a, dimensional classifiers who learned brightness categories were more accurate at between-category discriminations than within-category discriminations. This was supported by a mixed logistic regression model of the effect of discrimination comparison type (between, within relevant, within irrelevant) on discrimination accuracy; dimensional classifiers in the brightness learning group were significantly more accurate when making between category discriminations compared to within category discriminations for both irrelevant, $z = -4.50$, $p < .0001$ and relevant dimensions, $z = -3.93$, $p < .0001$. These children were also significantly more accurate at making within-category discriminations along the relevant than the irrelevant dimensions, $z = -4.50$, $p < .0001$, demonstrating that, overall, they were more accurate at discriminating across the relevant than the irrelevant dimension. As can be seen in the figure dimensional classifiers who learned size categories were no more accurate when making between category discriminations than when making within category discriminations on either the irrelevant, $z = 1.49$, *NS,* or the relevant dimension, $z = -.42$, *NS.* These children were no more accurate at making within-category discriminations along either dimension, $z = 1.49$, *NS.*

It was predicted that holistic classifiers who learned brightness categories would show an advantage for between-category discriminations relative to within-category discriminations, as in Experiment 1. As can be seen in Figure 15b, holistic classifiers who learned brightness categories were, in fact, more accurate at between-category discriminations than within-category discriminations. This was supported by a mixed logistic regression model showing that holistic classifiers who learned brightness categories were more accurate when making between category discriminations than within category discriminations along both the irrelevant, $z = -5.29$, $p < .0001$, and the relevant dimensions, $z = -4.66$, $p < .0001$. These children were also significantly more accurate at making within-category discriminations along the relevant than the irrelevant dimensions, $z = -5.29$, $p < .0001$, demonstrating that, overall, they were more accurate at discriminating across the relevant than the irrelevant dimension. However, as can be seen in the figure, holistic classifiers in the size learning group did not show any differences when making between-category discriminations relative to within-category discrimination along the irrelevant dimension, $z = 1.28$, *NS*, or along the relevant dimension, z=-.43. These children were no more accurate at making within-category discriminations along either dimension, z=1.49, *NS.*

Thus, just as in Experiment 1, both types of classifiers showed some evidence of an effect of category learning on their perceptual discrimination abilities—but only if that dimension was brightness.

Posttest triad

Next, I examined the results of the posttest triad task. It was predicted that, as in Experiment 1, the dimensional classifiers would not increase in their amount of dimensional responding from pre- to posttest. However, as can be seen in Figure 16a, the dimensional classifiers who learned brightness categories did not demonstrate an increase in dimensional responding. This was supported by the results of a mixed logistic regression model that showed that dimensional classifiers who learned brightness categories were no more likely to choose dimensional matches during the post test than they did on the pretest, z=-.65 *NS.* However, as can be seen in the figure, dimensional classifiers who learned size categories were more likely to choose dimensional matches during the post test, z=2.43, p<.05.

If labels drive attention to dimensional similarity, then the facilitative effects of labels in the category-learning task should have been a result of increased attention to dimensional similarity. Thus, it was predicted that the holistic classifiers would show increases in dimensional responding from pre to posttest. As can be seen in Figure 16b, the holistic classifiers who learned brightness categories were more likely to choose dimensional matches during the posttest. This is supported by the results of a mixed linear regression model showing that holistic classifiers who learned brightness categories showed an increase in dimensional responding, z=2.43, p<.05. Similarly, as can be can also be seen in the figure, those who learned size categories were more likely to select dimensional matches during the post test than the pre test, z=3.81, p<.001. These results support the prediction that the presence of a label during category learning increases dimensional attention in similarity classification.

A secondary question is whether the particular category type learned influenced children's choice of dimensional matches on triad task trials that probed the dimension they had trained on during category learning. I next examined the extent to which learning to selectively attend to a given dimension in categorization changes the specific dimension to which children selectively attend in similarity classification. It was expected that dimensional classifiers should be more likely to choose dimensional matches on posttest trials corresponding to the dimension they learned to categorize by relative to how often they did so on the pretest. However, as can be seen in Figure 17a, the dimensional classifiers who learned brightness categories were more likely to select dimensional matches on size trials during *both* pre- and posttest. This was supported by the results of a mixed logistic regression model of the interaction between trial type (size or brightness) and test time (pre- or post- category learning) on the likelihood of selecting a dimensional match. In particular the model showed that dimensional classifiers who learned brightness categories were significantly more likely to choose dimensional matches on size trials than brightness trials, z=2.98, p<.01. There was no significant interaction between trial type and test time, z=-.22, *NS*. As can also be seen in the figure, dimensional classifiers who learned size categories, on the other hand, were no more likely to choose dimensional matches on either size or brightness trials during either pre- or posttest. This was supported by the results of a mixed logistic regression model that revealed no effect of trial type, z=.01, *NS*, and no interaction between trial type and test time, z=.01, *NS*. Together, these results suggest that dimensional classifiers did not select the dimension corresponding to their learned category more often during posttest.

It was predicted that the holistic classifiers would choose dimensional matches more on trials corresponding to the dimensional structure of the category they learned. My proposal is that labels support selective attention to dimensional similarity. Thus, holistic classifiers should be able to selectively attend to category-relevant dimensions in a labeling context. This increase in dimensional attention should then influence children's responding in the triad task. As can be seen in Figure 17b, the holistic classifiers who learned brightness categories were more likely to choose dimensional matches on brightness trials. However, while a mixed logistic regression model of the interaction between trial type and test time showed that holistic classifiers who learned brightness categories were more likely to select dimensional matches on brightness trials, z=-9.03, p<.0001, there was no interaction, z=-1.23, *NS*. This suggests that during both the pre- and posttest triad tasks holistic classifiers were more likely to select dimensional matches on brightness trials. Furthermore, as can be seen in the figure, holistic classifiers who learned size categories were also more likely to select dimensional matches on brightness trials. This was supported by the results of a mixed logistic regression model showing that these children were significantly more likely to select dimensional matches on brightness trials, z=-2.24 p<.03, and a significantly interaction such that holistic classifiers who learned size categories were even more likely to select dimensional matches on size trials on posttest trials than pretest trials, z=3.14 p<.01.

Together, the results of the triad task show that holistic, but not dimensional classifiers' similarity classification was influenced by the categories they had learned earlier in the experiment. This differs from the results of Experiment 1, which demonstrated that only the dimensional classifiers' classification was influenced by the categories they learned. Clearly, the presence of redundant labels during category learning supported strong enough selective attention to category-relevant dimensions to carry over to similarity task even in holistic classifiers. It is not clear why, however, this attention to category-relevant dimensions did not carry over to the similarity task for dimensional classifiers in this experiment given that it did in Experiment 1.

In-depth analysis of dimensional attention

The results of category learning supported my prediction that the presence of labels would help even holistic classifiers learn dimensional categories more quickly. However, the critical question is whether this happened because holistic classifiers increased selective attention to category-relevant dimensions. The results of the posttest triad task support the prediction that the increased speed of category learning was related to an increase in dimensional attention. The results of the discrimination test support my hypothesis as well, in that both dimensional and holistic classifiers who learned brightness categories showed enhanced accuracy for category-relevant discriminations relative to category-irrelevant discriminations. However, this was the same finding from Experiment 1, where, even though the holistic classifiers were slow to learn dimensional categories and did not show relevant changes in the triad classification task, there was evidence of their dimensional attention in the discrimination task. Thus, as in Experiment 1, I conducted a more in-depth examination of children's accuracy in categorization and discrimination to assess the extent to which children were selectively attending to category-relevant dimensions.

Category test

As in Experiment 1, children in both groups who reached learning criterion were very accurate in the categorization test. In particular, the dimensional classifiers who learned brightness categories, $M=.84$, $t(6)=11.11$, $p<.0001$, the dimensional classifiers who learned size categories, $M=.86$, $t(6)=15.16$, $p<.0001$, the holistic classifiers who learned brightness categories, $M=.85$, $t(6)=13.32$, $p<.0001$, and the holistic classifiers who learned size categories, $M=.69$, $t(7)=4.30$, $p<.01$, were all significantly better than chance at categorizing stimuli in this task. Because neither group showed a difference in accuracy for categorizing familiar versus novel stimuli in Experiment 1, I predicted again neither group would show a difference.

As can be seen in Figure 18a, the data suggest, as predicted, that dimensional classifiers who learned brightness categories are equally accurate at categorizing familiar and novel stimuli. This was supported by a mixed logistic regression model of the effects of trial type on categorization accuracy showed that dimensional classifiers who learned brightness categories were no different in their accuracy of categorizing familiar than novel stimuli, z=.63, *NS.* As can be seen in the figure, dimensional classifiers who learned size categories were also equally accurate at categorizing all types of stimuli as well, z=.53, *NS.* This suggests that the dimensional classifiers may have learned something about the category's organization rather than just something about specific stimuli.

As can be seen in Figure 18b, the holistic classifiers who learned brightness categories were also equally good at categorizing familiar and novel stimuli. This is supported by a mixed logistic regression model of the effect of trial type on accuracy at category test showed that holistic classifiers who learned brightness categories were equally accurate at categorizing familiar and novel stimuli., z=.56, *NS.* As can be seen in Figure 18b, the holistic classifiers who learned size categories were also were equally good at categorizing both types of stimuli, z=.49, *NS.* This would suggest that the holistic classifiers also learned something about the category's organization rather than something about specific stimuli. However, because a null result was found, the possibility still exists that children in both groups paid attention to overall similarity rather the specific dimension relevant for categorization.

To examine the extent to which children were attending to the dimension relevant for categorization rather than to specific regions of the category space, I examined accuracy of children's categorization to novel stimuli with respect to whether the stimuli were drawn from inside or outside of the category space. It was predicted that the dimensional classifiers should be equally good or better at categorizing the outside stimuli relative to the inside stimuli. Although, as can be seen in Figure 19a, the dimensional classifiers who learned brightness categories were generally more accurate in categorizing novel stimuli outside the category space than inside the category space, this trend was not significant. This is supported by the results of a mixed logistic regression model of the effect of novel trial type on accuracy z=-1.54, *NS*. Similarly, the dimensional classifiers who learned size categories were also not any more accurate at categorizing novel stimuli from outside the category space than inside the category space, z=-1.28, *NS*. This suggests that at least the brightness learners may not have been attending to specific stimuli or regions of the category space, but instead perhaps to category relevant dimensions.

It was expected that the holistic classifiers would be equally good or better at categorizing the outside stimuli relative to the inside stimuli. As can be seen in Figure 19b, the holistic classifiers who learned brightness categories were equally accurate at categorizing novel stimuli inside and outside the learned category space. This was supported by a mixed logistic model of the effect of novel stimuli type on categorization accuracy, z=-.86, *NS*. Although, as can be seen in Figure 19b, the holistic classifiers who learned size categories were generally more accurate at categorizing novel stimuli from outside the category space than inside a mixed logistic regression model revealed that this difference was not significant, z=-1.11, *NS*.

Results of the categorization test overall show that both groups of children were equally accurate at categorizing all types of stimuli. This differs from the results of Experiment 1 that showed that dimensional classifiers who learned brightness categories and holistic classifiers who learned size categories were more accurate at categorizing novel stimuli from outside the learned category space. It is unclear why this would be. It may be that children in both experiments were overall so accurate in categorization that it is not possible to see differences in their attention by comparing accuracy for each trial type. It could be that generalizing to the novel stimuli in this task does not require as much selective attention as hypothesized and instead children were able to achieve a high degree of accuracy by learning about specific regions of the category space. Assessing the results of the discrimination test with respect to strongly- versus weakly-learned areas of the stimulus space should help clarify the extent to which these two groups were attending to dimensions rather than specific stimuli and regions of the category space.

Strongly versus weakly learned areas of the category

If children were more accurate at making discriminations from strongly-learned than weakly-learned areas of the category space, this would suggest that they had learned about specific regions of the space rather than category-relevant dimensions of similarity. To examine this, I measured accuracy for between and within-category discriminations in both strongly and weakly learned areas of the category space. A mixed logistic regression model of the interaction between stimuli location (strongly- versus weakly-learned space) and type (between, within-relevant, within-irrelevant) showed that dimensional classifiers were equally accurate at making discriminations in both strongly and weakly learned areas of the stimulus space, regardless of whether they had learned brightness categories, z=.33, *NS,* or size categories, z=-.56, *NS.* Similarly, a mixed logistic regression model revealed that holistic classifiers were also equally accurate regardless of whether they had learned brightness categories, z=.96, *NS*, or size categories, z=.16, *NS.* As in Experiment 1, the results of these models suggest that children might not have just learned about specific regions of the category space, but instead might have learned to selectively attend to category-relevant dimensions.

Discrimination step size

Next I examined discrimination accuracy by asking if learning these categories affect the children's ability to make discriminations along relevant dimension at even smaller metric differences than those learned. It was predicted that as in Experiment 1, both groups of children would be better at making discriminations across larger step sizes, but that this would not interact with discrimination type. A mixed logistic regression model of the interaction between the step size (.5, 1, or 1.5 times the discrimination JND) and discrimination type (between, within-relevant, within-irrelevant) shows that dimensional classifiers who learned brightness categories were significantly more accurate at making discriminations when the stimuli were more different, $z= 3.94$, $p<.0001$, as were those who learned size categories, $z=3.39$ $p<.001$. However, holistic classifiers who learned brightness categories showed no effect of discrimination step size, $z=.36$, *NS,* but did show an effect of trial type such that they were significantly more accurate at making between-category discriminations than within-category discriminations across both the relevant, $z=3.27$, $p<.01$, and irrelevant dimensions, $z=3.03$, $p<.01$. This model also revealed a significant interaction between trial type and step size, such that holistic classifiers who learned brightness categories were significantly less affected by step size when making between-category discriminations, z= -2.03, $p<.05$. On the other hand, those holistic classifiers who learned size categories did show a significant effect of step size on discrimination accuracy, $z=4.38$, $p<.0001$. There were no significant interactions. These results suggest that, with the exception of those holistic classifiers who learned brightness categories, children were better at making discriminations as the difference between stimuli increased. Overall, this replicates the results of the corresponding analysis from Experiment 1. However, it is critical to note the interesting results of the holistic classifiers who learned brightness categories. First, that they did not show an overall effect of step size on discrimination accuracy suggests that the holistic classifiers who learned brightness categories were equally accurate at making discriminations even at the smallest step size. Second, that these children showed an interaction such that they were least likely to show an influence of step size on accuracy for between-category discriminations shows that these children were selectively attending to the category-relevant dimension and this increased the accuracy with which they could make these discriminations.

Blocking

Finally, I examined children's discrimination accuracy with respect to the specific context of the block of trials where the discrimination occurred. It was predicted that both groups of children would show a cost for making discriminations in mixed blocks relative to single-type blocks. In Experiment 1, only the dimensional classifiers showed this switch cost. If labels are supporting holistic classifiers' selective attention to category-relevant dimensions, then they too should show a cost of switching between trial types. However, as can be seen in Figure 20a, dimensional classifiers who learned brightness categories were equally accurate at making discriminations in the context of single-type blocks than during mixed blocks. This was supported by a mixed logistic regression model of the effects of block on the discrimination accuracy showing that dimensional classifiers who learned brightness categories were equally accurate at making between category discriminations, z=.26, *NS*, equally accurate at making within category discriminations along the relevant dimension, z=-.89, *NS*, and equally accurate at making within category discriminations along the irrelevant dimension, z=.10, *NS* when they occurred within a single-type block than in a mixed block. Similarly, as can also be seen in the figure, dimensional classifiers who learned size categories were equally accurate at making between-category discriminations in single-type blocks and mixed blocks, z=.71, *NS,* equally accurate at making within-category discriminations along the relevant dimension in single-type and mixed blocks, z=-.69, *NS,* and equally accurate at making within-category discriminations along the irrelevant dimension in irrelevant and mixed blocks, z=1.22, *NS.* These results were surprising given that the dimensional classifiers in Experiment 1 did show a difference in discrimination accuracy between mixed and single-type blocks of trials.

It was predicted that holistic classifiers would show increased accuracy on trials occurring in single-type blocks relative to mixed blocks. However, as can be seen in Figure 20b, those who learned brightness categories showed no advantage of block type on between category discriminations, z=.89, *NS*, or within category discriminations along either the relevant, z=1.20, *NS*, or irrelevant dimensions, z=.85, *NS*. Similarly, as can also be seen in the figure, those who learned size categories showed no advantage for single-type blocks on between category discriminations, z=1.36, *NS,* or within category discriminations along either the relevant, z=-1.54, *NS,* or irrelevant dimensions, z=-.70, *NS.* These results, like those of the dimensional classifiers, are surprising given that the dimensional classifiers in Experiment 1 did show a cost for making discriminations in mixed blocks. It was expected that if children had learned to selectively attend to category-relevant dimensions then they should show a cost for making discriminations in mixed blocks. However, in both experiments children are overall very accurate in categorization and discrimination. It is possible that these measures are not sensitive enough to capture differences related to selective attention if children are performing at or near ceiling for all trial types.

Conclusions

I proposed that over development, word learning scaffolds selective attention to dimensional similarity, not just in the context of learning new words, but also more generally. If my proposal is correct, then we should see a role for labels in scaffolding children's dimensional attention in non-linguistic tasks, such as the triad classification task. The primary results of Experiment 2 generally support this proposal by demonstrating that, unlike in Experiment 1, holistic and dimensional classifiers are equally quick to learn dimensional categories. Furthermore, the results of the posttest triad task demonstrate that, also unlike in Experiment 1, both holistic and dimensional classifiers show an increase in selective attention to the dimension of similarity relevant to the category they had learned. Interestingly, however, the results of the discrimination task do not differ from those of Experiment 1—namely that both holistic and dimensional classifiers who learned brightness categories showed enhanced accuracy for relevant discriminations over irrelevant discriminations. This could suggest that dimensional category learning itself can increase selective attention to dimensional similarity.

Thus, in children who are holistic classifiers, and have relatively weak selective attention abilities, it takes a long time to learn dimensional categories. Once the categories are learned, however, these children show slight increases in dimensional attention—as evidenced by their performance in the discrimination task. However, this increase was not strong enough to push them into very large increases in dimensional attention in the posttest triad task. Because labels appear to support even weak selective attention, however, when holistic classifiers learn dimensional categories in the context of redundant labels, they can learn the categories quite quickly. This increase in their selective attention then cascades forward to both their discrimination abilities and their classification biases. So, as Lupyan suggested in his 2008 study of effects of category grouping on visual processing, “categories matter, and named categories matter more.”

However, it is important to note that the results of the secondary in-depth analyses of categorization and discrimination are mixed. Overall, these tests did not reveal clear evidence of children having changed attention to relevant versus irrelevant dimensions. It is unclear whether these tests are not sensitive enough to reveal a change in attention, or if children did not have such a change. To further examine the respective contributions of category learning and labels on increases to holistic classifier’s selective attention to dimensional similarity, in the next chapter, I directly compare the performance in each task of holistic classifiers in Experiments 1 and 2. This comparison should clarify the nature of these mixed results.

CHAPTER VI: DIRECT COMPARISON OF HOLISTIC CLASSIFIERS IN EXPERIMENTS 1 AND 2

The results of Experiments 1 and 2 demonstrated that holistic classifiers are slower to learn dimensional categories than dimensional classifiers, but the presence of incidental labels eliminates this difference. Additionally these experiments demonstrate that the effects labels have on dimensional attention in category learning cascade forward to a similarity classification task. Together these results generally support the hypothesis that labels scaffold young children's selective attention to dimensional similarity. However, the results of the discrimination task also demonstrate that regardless of the presence or absence of labels, category learning itself might also support selective attention. Holistic classifiers who learned brightness categories in both experiments demonstrated enhanced accuracy for category-relevant discriminations relative to category-irrelevant discriminations. Thus, both category learning and category learning in the context of labels appear to support some degree of selective attention (for those children who learned brightness categories). The purpose of this chapter is to directly compare the performance of holistic classifiers in the two experiments. Doing so will help clarify exactly how labels affected children's performance in these tasks above and beyond category learning.

Comparison Predictions

Children's performance in each task was compared across the two experiments for holistic classifiers who learned brightness categories and for those who learned size categories. The same basic mixed logistic regression model structure was used to make each comparison: the interaction between experiment (label versus no label) and task manipulation (e.g. between, within-relevant, or within-irrelevant discrimination type). This allowed me to examine how the presence of a label might interact with the effects of category learning.

There were two qualitative differences in holistic classifiers' performance in Experiments 1 and 2 that have already been described. First, holistic classifiers in Experiment 2 (label) learned dimensional categories as quickly as dimensional classifiers did, while those in Experiment 1 were significantly slower than dimensional classifiers. Thus, it was expected that the comparison analysis between the two holistic classifier groups would show that those in Experiment 2 were significantly faster to learn the categories than those in Experiment 1. Second, holistic classifiers in Experiment 2 showed increases in dimensional responding from pre- to posttest that related to the structure of the categories they learned, while those in Experiment 1 did not. Thus it was expected that the comparison analysis would show a significantly larger increase in dimensional responding from pre- to posttest for the holistic classifiers in Experiment 2 relative to those in Experiment 1.

There were no other qualitative differences between holistic children's performance in the two experiments. Thus, it was anticipated that any differences revealed in the comparison analyses would be related to overall accuracy rather than differences in the overall patterns of performance. For example, it is possible that children's overall highly accurate performance in the discrimination task could mask differences in discrimination accuracy for strongly- versus weakly-learned areas of the category space. A direct statistical comparison of the two experimental groups, however, might reveal such fine-grained effects, and thus help clarify the mixed results of the categorization and discrimination tasks.

Comparison Analyses and Discussion

Category learning

If labels facilitate category learning, even in young children who still generally attend to holistic similarity relationships, then the holistic classifiers in Experiment 2 should be significantly faster to learn categories than the holistic classifiers in Experiment

1. In fact, a mixed linear regression model of the interaction between experiment (label versus no label) and category structure (size versus brightness) on the number of blocks to reach learning criterion revealed that holistic classifiers were significantly quicker to reach criterion in the label experiment than the no label experiment, t=-26.57, p<.0001. There was an overall effect of category structure, such that children were faster to learn brightness categories, t=-23.66, p<.0001, however, there was also a significant interaction such that children in the label experiment showed less of an advantage for learning brightness categories, t=7.47, p<.0001. Thus, holistic classifiers who learned categories in the context of redundant labels were not only equally fast to learn those categories as their dimensional counterparts (as we saw in the results of Experiment 2), but they were significantly faster than their holistic counterparts in the no label experiment. Further, these children learned brightness and size categories at almost the same rate. This result provides direct evidence that labels facilitate dimensional category learning in children who have difficulty selectively attending to dimensional similarity.

Discrimination

If labels facilitate category learning because they increase children's selective attention to category-relevant dimensions, then the children in the label experiment should show the biggest enhancement in accuracy for relevant over irrelevant discriminations. However, we also know from the results of Experiment 1 and 2 that holistic classifiers who learned brightness categories demonstrated the same general pattern of results in the discrimination tasks of both experiments. Nevertheless, it is still possible that the label had an effect on children's discrimination, such as increasing accuracy overall, that would be revealed by a direct comparison between the two experiments. In fact, a mixed logistic regression model of the interaction between experiment (label versus no label) and discrimination type (between, within-relevant, or within-irrelevant) revealed a marginal effect of experiment such that the holistic classifiers who learned brightness categories in the context of an incidental label had a tendency to be more accurate overall in discriminations than those in the no label experiment, z=1.69, p<.10. Additionally, there were significant interactions such that children in the label experiment showed a larger difference in accuracy in between-category discriminations relative to within-category discriminations on the relevant dimension, z=2.25, p<.05, and relative to within-category discriminations on the irrelevant dimension, z=2.28, p<.05. However, the equivalent mixed logistic regression model of holistic classifiers who learned size categories did not reveal significant effects of experiment, z=2.28, *NS*, or any interactions. Together, these results demonstrate, at least for those children who learned brightness categories, that the presence of a label made the effects of category learning on perceptual discrimination somewhat stronger. This is evidence that labels increase selective attention to dimensions above and beyond category learning.

<u>Posttest triad</u>

If word learning is what drives the emergence of selective attention to dimensions in similarity classification over development, then we may also see indices of these changes in attention over the course of an experiment if there are incidental labels presented during category learning. We saw in both Experiments 1 and 2 that holistic classifiers showed some increases in dimensional responding. However, only those children in Experiment 2 showed any evidence of changes in their dimensional responding related to the specifics of their category learning. In particular, children in Experiment 1 who learned size categories only showed a slight increase in attention to brightness—a finding that suggests any effects of category learning did not carry forward to the similarity classification. Therefore I predicted that the holistic classifiers in the label experiment should show the largest increases in dimensional responding from the pre- to posttest triad task regardless of which type of category they learned. This prediction was supported by a mixed logistic regression model of the effect of experiment (label versus no label) on the change in dimensional responses from pre- to posttest triad for holistic classifiers who learned brightness categories. This model revealed that those in the label condition showed a significantly larger increase in dimensional responding from pre- to posttest, $z=3.34$, $p<.001$. Similarly, the equivalent mixed logistic regression model of the effect of experiment on change in dimensional responding for holistic classifiers who learned size categories revealed that those in the label experiment also showed a significantly larger increase in dimensional responding from pre- to posttest, $z=3.55$, $p<.001$. These results demonstrate that the presence of a label not only immediately facilitated category learning, but also led to changes in selective attention that cascaded forward to a similarity classification task.

In-depth analysis of dimensional attention

Together, the results of the comparison analysis support my proposal that labels scaffold dimensional attention: holistic classifiers are faster to learn dimensional categories in the presence of labels; they demonstrate more extreme influences of attention to category-relevant/category-irrelevant dimensions on their perceptual discrimination abilities; and they demonstrate significantly larger increases in dimensional attention in a similarity classification task from pre- to posttest. Next, I examined performance in the categorization and discrimination tasks at the detailed level of stimulus location, threshold size, and blocking to examine whether the presence of a label had an effect on children's accuracy.

Category test

The results of Experiment 1 and 2 revealed no differences in categorization accuracy for familiar and novel stimuli by classification group, category structure or experiment. This was likely because all the children performed very well on the categorization test. Thus, it was expected that there would be no significant differences between the two groups of holistic classifiers. In fact, a mixed logistic regression model of the interaction between experiment (label versus no label) and trial type (familiar or novel) on children's categorization revealed that holistic classifiers who learned brightness categories showed no effect of experiment, z=.29, *NS.* Similarly, the equivalent mixed logistic regression model for holistic classifiers who learned size categories revealed no effect of experiment, z=-.12, *NS.* Clearly all children were equally accurate at categorizing familiar and novel stimuli. It is unclear just from this test, however, if this is because all children had truly learned to attend to category-relevant dimensions or whether this test was not a strong enough measure of dimensional attention. For this reason, I next examined children's accuracy in categorization with respect to type of novel stimulus (inside or outside the category space).

In Experiment 1, the holistic classifiers who learned size categories were significantly more accurate at categorizing novel stimuli from outside the category space than from inside the space. In Experiment 2, however, while holistic classifiers generally demonstrated higher accuracy for categorizing novel stimuli outside the space, they were not statistically different at categorizing stimuli from outside and inside. Thus, directly comparing the performance will help clarify whether the children in the two experiments actually were different with respect to accuracy In fact, the results of mixed logistic regression models demonstrate that there was no difference between the children in the two experiments in accuracy of categorizing novel stimuli. In particular a mixed logistic regression model of the interaction between experiment (label versus no label) and trial type (novel inside or novel outside) showed that holistic classifiers who learned brightness categories were equally accurate in both experiments, z=.43, *NS.* Similarly the equivalent model for holistic classifiers who learned size categories revealed no effect of experiment, z=-.55, *NS.* Together, these comparisons of the holistic classifiers' accuracy in categorization demonstrate that once children learned dimensional categories, they are equally accurate at categorizing different types of stimuli as each other, regardless of whether there had been labels present or not.

Strongly- versus weakly-learned areas of the category

The results of Experiment 1 and 2 revealed that no group of children, regardless of classifier type, learned category structure, or experiment, showed any differences in accuracy for discriminations between stimuli from strongly- versus weakly-learned areas of the category space. It is possible that all children's very high accuracy in discrimination overall prevented me from seeing any difference related to attention to category relevant dimensions rather than specific regions of the category space. Exploring performance in this task by directly comparing holistic classifiers from the two experiments allowed me to assess whether there were any differences not captured by the qualitative results. A mixed logistic regression model of the interaction between experiment (label versus no label), location (strongly- versus weakly-learned), and discrimination type (between, within-relevant, or within-irrelevant) showed a marginal effect of experiment such that children in the label experiment were overall slightly more accurate at making discriminations than those in the no label experiment, z=1.69, p<.10. There were not, however, any significant effects of location, z=.40, *NS.* However the equivalent mixed logistic regression model of discrimination accuracy for holistic classifiers who learned size categories showed no effect of experiment, z=.56, *NS* and no interaction with location, meaning that children in both experiments were no different in making discriminations in both strongly- and weakly-learned areas of the category space. Thus these two models do not tell us anything above and beyond the results of the earlier comparison model of discrimination that showed an effect of experiment on discrimination accuracy. It is unclear whether the lack of an effect is because these two groups of children are both showing selective attention to category-relevant dimensions or because this test is not sensitive enough to capture differences. Assessing results of the other two discrimination measures may help clarify this issue.

Discrimination step size

The results of Experiment 1 and 2 demonstrated that children generally tended to be more accurate at discrimination with bigger differences between the stimuli. It was therefore expected that there would be no difference between the holistic classifiers in the two experiments with respect to threshold size. In fact, a mixed logistic regression model of the interaction between experiment (label versus no label), threshold size (.5, 1, or 1.5x JND), and discrimination type (between, within-relevant, or within-irrelevant) on the discrimination accuracy of holistic classifiers who learned brightness categories did not reveal any effect of condition, z=1.42, *NS,* or any interactions with threshold size. Additionally, the equivalent model for holistic classifiers who learned size categories did not reveal any effect of condition, z=-.02, *NS,* or any interactions with threshold size. These results suggest that children in the two experiments were equally affected by the size of the difference between stimuli. Again, it is not clear from the results of this test whether this means children are demonstrating selective attention or not. Assessing differences between holistic classifiers' performance in the discrimination task with respect to the blocking of trials may help clarify this issue.

Blocking

The results of Experiment 1 and 2 showed no effect of blocking type on holistic classifiers' discrimination accuracy. However, the results of Experiment 1 did show that dimensional classifiers were more accurate when making discriminations in the context of single-type blocks than in mixed blocks. This was taken as evidence that dimensional classifiers in Experiment 1 were selectively attending to category-relevant dimensions and thus showed a cost for making discriminations in mixed blocks where they had to switch between attending to each dimension. It is possible that I did not find such a switch cost for children in Experiment 2—even if they had been selectively attending to the relevant dimension during category learning—because they were so accurate overall at discrimination. Making a direct comparison between holistic classifiers in the two experiments will clarify whether this is the case. I predicted that if there was a difference between the two experiments it would be such that the holistic classifiers in the label condition were more accurate overall and would show a greater cost for making discriminations in mixed blocks than those in the no label condition. However, a mixed logistic regression model of the interaction between experiment (label versus no label), block (mixed versus relevant or irrelevant), and discrimination type (between, within-relevant or within-irrelevant), revealed no significant effect of experiment on the accuracy of holistic classifiers who learned brightness categories doing between-category discriminations, z=1.25, *NS,* within-relevant discriminations, z=.21, *NS* or within-irrelevant discriminations, z=.14, *NS*, nor were there any interactions with block. Similarly, the equivalent model for holistic classifiers who learned size categories revealed no significant effect of experiment on the accuracy of between-category discriminations, z=-.06, *NS,* within-relevant discriminations, z=.98, *NS* or on within-irrelevant discriminations, z=1.35, *NS*, nor were there any interactions with block. These results suggest that the presence of a label did not affect children's likelihood of showing a cost for making discriminations in the context of mixed blocks.

Conclusions

The results of the primary comparison analyses between holistic classifiers in the two experiments add to the qualitative analyses of the previous two chapters by directly examining the extent to which labels support selective attention above and beyond category learning. In particular, the analysis of learning speed revealed that holistic classifiers in the label experiment were significantly faster to learn dimensional categories—either organized by brightness or size. This added to the results of

Experiment 2 by demonstrating that it was not just the case that the presence of labels removed the difference between holistic and dimensional classifiers, but that it led the holistic classifiers to be significantly faster to learn than their counterparts in the no label experiment.

Additionally the results of the comparison of discrimination task accuracy between the two experiments revealed that holistic classifiers in the label experiment were not only significantly more accurate at discrimination overall, but they showed a larger difference in accuracy for the different trial types. In other words, children in the label experiment showed greater enhanced attention to category-relevant dimensions and more decreased attention to category-irrelevant dimensions. This adds to the qualitative assessment of Experiments 1 and 2, both of which demonstrated that holistic classifiers who learned brightness categories showed the "adult" pattern found by Goldstone (1994b). The analysis presented in this chapter shows that the holistic classifiers in the label experiment show a more extreme pattern, adding support to the claim that category learning affects attention to dimensions, but labeled category learning affects it more.

Finally, the comparison of the two holistic groups' changes in dimensional responding in the triad task revealed that those in the label experiment showed significantly larger increases in dimensional responding from pre- to posttest than those in the no label experiment. This offers the strongest evidence that labels scaffold selective attention to dimensional similarity. The only difference between the two experiments was the presence of incidental, redundant, labels during the feedback in category learning. Yet this difference was enough to change children's pattern of responding in a seemingly unrelated similarity classification task. Together, the results of these three primary analyses clearly show that labels can support selective dimensional attention even in children who tend to attend to holistic similarity relationships.

The results of the secondary analyses—which were supposed to provide a more in-depth analysis of the extent to which children demonstrate selective attention to dimensional similarity—however, do not provide such clear support. These secondary analyses reveal no effects of experiment (label versus no label) on children's accuracy in categorizing familiar versus novel stimuli, or novel inside versus novel inside stimuli, or in discrimination based on stimulus location, threshold size, or block type. The null results of these secondary analyses (with respect to experiment) coupled with the similarly unclear results of these analyses in the qualitative assessments of Experiments 1 and 2 suggest that these tests might not be strong enough measures of selective attention to be informative.

Nevertheless, the results of the direct comparison are nevertheless informative. The purpose of Experiments 1 and 2 was to explore the extent to which the presence of incidental labels could support dimensional attention in holistic above and beyond category learning. While the results of these experiments demonstrate that labels do in fact scaffold dimensional attention above and beyond category learning, category learning still supports a degree of dimensional attention even in holistic classifiers. That holistic classifiers can be pushed into selectively attending to dimensional similarity via category learning with or without labels supports the proposal that the change from holistic to dimensional attention is one of continuous developmental change rather than a discrete.

Previous studies have also proposed continuity in the development of selective dimensional attention. For example, researchers have argued that children who preferentially attend to one dimension of similarity (e.g. brightness) regardless of whether this means classifying by dimensional or holistic similarity may be an intermediate part of this developmental continuum (see e.g. Hanania & Smith, 2011). Thus, in the next chapter, I explore the effects of category learning with and without incidental labels on changes to selective dimensional attention in "preferrer" children. Understanding how category learning and labels influence these children's selective attention to dimensions will help clarify how children eventually begin to attend to dimensional similarity without linguistic support.

CHAPTER VII: EXPERIMENTS 3 AND 4

If the holistic to dimensional shift is a continuous increase in selective and flexible attention, then examining children who use an intermediate classification strategy, such as the preferrer strategy, is necessary for our understanding of selective attention across the developmental spectrum. Thus, in Experiment 3, I examined the influence of category learning on preferrers' attention to category-relevant dimensions. In Experiment 4, I examined whether the presence of a label during category learning would affect these children's subsequent discrimination and classification abilities. At the end of this chapter, I conduct a direct comparison between the two experiments for each preferrer group.

Experiment 3

The purpose of this experiment was to examine preferrers' category learning and its effects on their dimensional attention. It is unclear why children should prefer the specific dimension they do when they demonstrate a preference in similarity classification. It is also unclear how stable this preference is and whether it will influence children's learning and discrimination later in the experiment. Some researchers have explained the preferrer strategy by arguing that individual dimensions "become available" to children at different times as their perceptual systems mature, and that the order in which dimensions become available differs between children (Cook & Odom, 1992). If this preference is relatively stable, then preferrers should be able to learn dimensional categories fairly quickly if the category is organized by the dimension they prefer (i.e. brightness preferrers should be faster to learn brightness categories than size categories and size preferrers should be faster to learn size categories than brightness categories).

A second possibility is that the preference is not stable. It could be that children end up classified as a preferrer not because they are really good at focusing on brightness (for example) but instead because they are able to selectively attend to one dimension but cannot yet flexibly switch between dimensions. This type of attention could be described as, "sticky attention" (see Hanania & Smith, 2010). If this is the case, then preferrers may be equally quick to learn either type of category, and would show dramatic difference in their ability to make category-relevant versus category-irrelevant discriminations. Additionally, if it was the case that they have sticky attention, preferrers may switch which dimension they prefer from pre- to posttest triad depending on their learned category structure.

A third possibility is that there is something inherently different about brightness and size dimensions. The results of Goldstone's 1994 study of perceptual learning and the holistic and dimensional classifiers' category learning in Experiment 1 both demonstrate that, in general, children and adults are faster to learn brightness categories than size categories. The results of these studies also demonstrate that both children and adults show changes to their abilities to make discriminations along relevant versus irrelevant dimensions if they learn brightness—but not size—categories. It therefore seems probable that size is a more difficult dimension to attend to than brightness and even when people learn categories that force them to attend to size, they still are really good at attending to brightness. If size is just a more difficult dimension to attend to, this could mean that size preferrers actually have much stronger selective attention abilities than brightness preferrers. Size preferrers might be equally good at learning either type of category, then, because they are good at attending to size and because brightness is an easier dimension for anyone to attend to. Brightness preferrers, on the other hand, would be quick to learn brightness categories, but show difficulty in learning size categories (just as holistic and even dimensional classifiers did in Experiment 1).

By comparing the speed with which children of each preferrer type learn categories of each type, and comparing the qualitative patterns of results in categorization, discrimination, and posttest triad tests, I should be able to clarify not only what it means to “prefer” a dimension, but also how children using this intermediate strategy fit into the continuum from holistic to dimensional attention..

Method

Participants

Participants were 33 5- to 8-year-old children. One child did not complete the entire experiment and was excluded from these analyses. Thus there were 32 children in the final group.

Procedure

The procedure was identical to that of Experiment 1—the no label experiment. The only difference was that all the children in this experiment used a preferrer strategy on the triad classification test. That is, they preferentially selected the stimulus that most closely matched the exemplar on one dimension (size or brightness) regardless of whether that meant selecting the holistic or the dimensional match. Children were divided into classifier groups based on performance in the triad pretest. There were 16 brightness preferrers and 16 size preferrers. See Table 1 for participant information.

Results and Discussion

Category Learning

It was thought that preferrers might differ from each other with respect to the speed with which they learned categories of each type. As can be seen in Figure 21, the size preferrers were equally fast to learn both brightness and size categories while the brightness preferrers were only fast at learning brightness categories. This was supported by a mixed linear regression model of the effect of classifier type (brightness versus size preferrer) on the number of blocks it took children to reach criterion. The model revealed a significant effect of classifier type such that size preferrers were significantly faster to reach criterion than brightness preferrers, t=-50.60, p<.0001, a significant effect of category structure such that children learned brightness categories more quickly than size categories, t=-65.80, p<.0001, and significant interaction such that size preferrers were less likely to show an advantage for learning brightness categories, t=4.98, p<.0001. These results support the idea that there is something different about brightness and size dimensions that makes categorize defined by size more difficult to learn. By this view, if a child is good at attending to size (the more “difficult” dimension) then they are also good at attending selectively to brightness (the “easy” dimension) and can learn either category structure relatively quickly. On the other hand, if a child is good at attending to brightness, they have more difficulty attending to size, and they are therefore much quicker to learn brightness categories than size categories.

Discrimination

Overall, children were quite accurate in discriminating stimuli. In particular, the brightness preferrers who learned brightness categories, M=.82, $t(7)$=9.51, p<.0001, the brightness preferrers who learned size categories, M=.82, $t(6)$=13.96, p<.0001, the size preferrers who learned brightness categories, M=.85, $t(7)$=18.29, p<.0001, and the size preferrers who learned size categories, M=.80, $t(7)$=8.35, p<.0001, were all significantly better than chance at discriminating stimuli in this task.

Children’s accuracy in making relevant versus irrelevant discriminations could differ depending on whether the category they learned was “easy” for them to learn given their classifier type. In particular, it was thought that brightness preferrers who learned brightness categories could show an advantage for category-relevant over category-irrelevant discriminations, but those who learned size categories might not. As can be seen in Figure 22a, brightness preferrers who learned brightness categories were more accurate at making between-category than within-category discriminations. This was supported by the results of a mixed logistic regression model of the effects of discrimination type (between, within-relevant, within-irrelevant) on accuracy. The model showed that brightness preferrers who learned brightness categories were significantly more accurate at making between-category discriminations than within-category discriminations along the irrelevant dimension, z=-4.83, p<.0001, or the relevant-dimension, z=3.83, p<.001. These children were also significantly more accurate at making within-category discriminations along the relevant than the irrelevant dimensions, z=-4.27, p<.0001, demonstrating that, overall, they were more accurate at discriminating across the relevant than the irrelevant dimension. However, as can be seen in the figure, if brightness preferrers learned size categories, they actually demonstrated the reverse pattern. These children were actually more accurate at making discriminations along the irrelevant dimension than the relevant dimension. This was supported by a mixed logistic regression model of the effect of discrimination type on accuracy that revealed that brightness preferrers who learned size categories were significantly more accurate at making discriminations along the irrelevant dimension than either making between-, z=2.31, p<.05, or within-category discriminations, z=3.50, p<.001. These children were actually significantly more accurate at making within-category discriminations along the irrelevant than the relevant dimensions, z=3.50, p<.001. Together, these results demonstrate that brightness preferrers were very accurate at making brightness discriminations regardless of whether brightness was relevant or irrelevant to their category learning. However, if they learned brightness categories, they do appear to be affected by their learning in that they are more accurate at making between-category than within-category discriminations along the relevant dimension.

Because size preferrers were equally quick to learn either category structure, they should show increased accuracy for discriminations along the category-relevant dimensions relative to the category-irrelevant dimension regardless of which type of category they learned. As can be seen in Figure 22b, size preferrers who learned brightness categories were in fact more accurate in making between-category discriminations than within-category discriminations. This is supported by a mixed logistic regression model of the effect of discrimination type (between, within-relevant, within-irrelevant) on accuracy. This model revealed a significant effect of discrimination type such that size preferrers who learned brightness categories were more accurate at making between category discriminations than within-category discriminations along either the relevant dimension, z=4.04, p<.0001, or the irrelevant dimension, z=-4.02, p<.0001. These children were also significantly more accurate at making within-category discriminations along the relevant than the irrelevant dimensions, z=-3.11, p<.01, demonstrating that, overall, they were more accurate at discriminating across the relevant than the irrelevant dimension. However, as can also be seen in the figure, the same was not true for those size preferrers who learned size categories. A mixed logistic regression model of the effect of discrimination type on accuracy revealed that these children were equally accurate at making between-category discriminations as they were making within-category discriminations along the relevant, z=.40, *NS*, and irrelevant dimensions, z=-43, *NS*. These children were no more accurate at making within-category discriminations along either dimension, z=-.15, *NS.* Thus, despite the fact that size preferrers focused on size similarity in the classification task, and were quite quick to learn size categories, they do not show any evidence of increased attention to size in the discrimination task.

Posttest Triad

As was the case in Experiment 1, the primary question of interest with respect to the results of the posttest triad task is was there an increase in dimensional responding from pre- to posttest triad task? It was expected that neither classifier group would show an increase in dimensional responding if they learned the type of category corresponding to the dimension they preferred. On the pretest, each preferrer group was already choosing the maximum number of dimensional responses corresponding to their dimension. Even if category learning increases attention to relevant dimensions, these children will not be able to exceed the number of dimensional matches that they choose on pretest. However, if these children learn the category structure organized by their non-preferred dimension, then they might show an increase in dimensional responding. The question is, though, would category learning actually lead to an increase in dimensional responding. It was thought that brightness preferrers who learned brightness categories would not show an increase. It was also thought that brightness preferrers who learned size categories would not increase, because they had had so much difficulty learning their categories that the learning should not have impacted their subsequent attention. As can be seen in Figure 23a, brightness preferrers who learned brightness categories did not show an increase in dimensional responding from pre- to posttest. This was supported by a mixed logistic regression model of the effect of test time (pretest or posttest) on dimensional responding that revealed no effect of test time, z=1.54, *NS.* Similarly, as can be seen in Figure 23a, brightness preferrers who learned size categories did not show an increase in dimensional responding either. This was supported by a mixed logistic regression model of the effect of test time (pretest or posttest) on dimensional responding that revealed no effect of test time, z=.66, *NS.* Together these results show that category learning did not alter brightness preferrers' dimensional attention in the triad classification task.

Similarly, it was predicted that size preferrers who learned size categories would not show any increase in dimensional responding. As can be seen in Figure 23b, size preferrers who learned size categories did not in fact show any increase. This was supported by a mixed logistic regression model of the effect of test time (pre versus post) on dimensional responding that revealed no effect of test time, z=-.10, *NS*. However, it was predicted that those who learned brightness categories might show an increase in dimensional responding as these children had been able to learn their categories relatively easily. But, as can be seen in the figure, size preferrers who learned brightness categories did not show any change in dimensional responding from pre- to posttest triad. This was supported by a mixed logistic regression model of the effects of test time (pre or post) on dimensional responding that reveal no effect of test time, z=.92, *NS.* Thus, neither brightness nor size preferrers showed an increase in dimensional responding from pre- to posttest. It remains to be seen, however, whether the specific pattern of dimensional choices children made was influenced by the specifics of the categories they learned.

Thus, a secondary question is whether preferrers change the dimension to which they prefer to attend. If category learning can affect the specific dimensions of similarity to which children attend, then preferrers may switch their preferred dimension. This would support the idea of preferrers having "sticky attention." Once they selectively attend to a given dimension, they get stuck and cannot flexibly switch to attending to another dimension. However, if they can learn a dimensional category, this could get them stuck on a dimension that is different than the one they initially preferred. Thus, it was thought that brightness preferrers who learned brightness categories might continue to prefer brightness and size preferrers who learned size categories would continue to prefer size. Brightness preferrers who learned size categories did not show increased accuracy for the relevant dimension over the irrelevant dimension in discrimination, and so it was not expected that they would prefer size in the posttest. Size preferrers who learned brightness categories on the other hand, did show increased accuracy for the relevant dimensional over the irrelevant dimension in discrimination, and so it was expected that they might develop a preference for brightness in the posttest.

As can be seen in Figure 24a, brightness preferrers who learned brightness categories did not change their preference from pre- to posttest. This was supported by a mixed logistic regression model of the interaction between trial type (size or brightness match) and test time (pretest or posttest) that revealed no significant effect of test time, z=1.44, *NS.* This model did reveal a significant effect of trial type such that children were most likely to select dimensional matches on brightness trials, z=-6.04, p<.0001. Thus brightness preferrers remained brightness preferrers. Similarly, as can be seen in the figure, those brightness preferrers who learned size categories showed an increase in their likelihood of selecting dimensional matches on size trials. This was supported by a mixed logistic regression model of the interaction between trial type (size or brightness) and test time. This model revealed a significant effect of time, such that children were more likely to select dimensional matches on the posttest, $z=3.66$, $p<.001$, a significant effect of trial types, such that they were more likely to select the dimensional match on brightness trials, $z=-6.42$, $p<.0001$, and a significant interaction such that children were more likely to choose dimensional matches on size trials during the posttest, $z=-3.75$, $p<.001$. Thus, despite the fact that brightness preferrers had difficulty learning size categories and did not show increased attention to the relevant dimension in discrimination, they do show an increase in attention to similarity in size if they learned categories organized by size.

As can be seen in Figure 24b, size preferrers who learned brightness categories showed a small increase in their likelihood of selecting dimensional matches on size trials from pre- to posttest. This was supported by the results of a mixed logistic regression model that revealed a significant effect of test time such that children were more likely to select dimensional matches on the posttest triad, $z=2.89$, $p<.01$, a significant effect of trial type such that overall children were more likely to select dimensional matches on size trials, $z=10.13$, $p<.0001$, and a significant interaction such that children were more likely to select dimensional matches on brightness trials during the posttest triad, $z=-2.95$, $p<.01$. Thus, size preferrers continued to be size preferrers, but showed a slight increase in their likelihood of attending to brightness. As can also be seen in the figure, size preferrers who learned size categories continued to prefer size as well. This was supported by the results of a mixed logistic regression model that revealed no significant effect of time, $z=.96$, *NS,* a significant effect of trial type, such that children were more likely to select the dimensional match on the size trials, $z=10.32$, $p<.0001$, and no interaction, z=-1.35, *NS.* Thus, size preferrers continued to be size preferrers if they had learned size categories.

In-depth analysis of dimensional attention

The results of Experiment 3 so far suggest that children who prefer classifying by size are somewhat different than those who prefer to classify by brightness. Size preferrers learn both size and brightness categories equally quickly, but only show changes to their discrimination abilities if they learned brightness categories. Brightness preferrers on the other hand learn brightness categories much more quickly than size categories and show increased accuracy for brightness discriminations no matter whether they learned brightness or size categories. In the triad task, both classifier types, however, show an increase in attention to the non-preferred dimension if they learned categories organized by that dimension. This last result would suggest that both preferrer groups are at some intermediate place on the developmental continuum from holisitic to dimensional attention. The results of Experiment 1 demonstrated that holisitic classifiers' posttest responses were not influenced by the specifics of the categories they learned, while the dimensional classifiers were greatly influenced. The current results show that preferrers are also influenced, but not to the extent that they flexibly choose dimensional matches on a majority of trials. Why then, do brightness preferrers not show increased attention to the relevant dimension during discrimination? Next I take a more in-depth look at the results of categorization and discrimination to clarify this issue.

Categorization

As in Experiment 1, children in both groups who reached learning criterion were very accurate in the categorization test. In particular, the brightness preferrers who learned brightness categories, M=.84, $t(7)$=11.19, p<.0001, the brightness preferrers who learned size categories, M=.74, $t(6)$=3.68, p<.01, the size preferrers who learned brightness categories, M=.89, $t(7)$=30.89, p<.0001, and the size preferrers who learned size categories, M=.76, $t(7)$=10.36, p<.0001, were all significantly better than chance at categorizing stimuli in this task. The results of Experiment 1 revealed that neither holistic nor dimensional classifiers demonstrated any difference in their accuracy of categorizing familiar versus novel stimuli. It was predicted that this would hold true for preferrers too. As can be seen in Figure 25a, brightness preferrers who learned brightness categories were equally accurate at categorizing familiar and novel stimuli. This was supported by the results of a mixed logistic regression model of the effect of stimulus type (familiar or novel) on categorization accuracy, z=1.09, *NS.* Similarly, as can be seen in the figure, brightness preferrers who learned size categories were equally accurate at categorizing familiar and novel stimuli. This was supported by the results of a mixed logistic regression model of the effects of stimulus type on categorization, z=.81, *NS.* Regardless of the type of category they learned, brightness preferrers were equally accurate at categorizing familiar and novel stimuli. As can be seen in Figure 25b, size preferrers who learned brightness categories were also equally accurate at categorizing both kinds of stimuli. This was supported by the results of a mixed logistic regression model of the effects of stimulus type on categorization, z=-.26, *NS*. Similarly, as can be seen in the figure, size preferrers who learned size categories also were equally accurate at categorizing both familiar and novel stimuli. This was supported by the results of a mixed logistic regression model of the effects of stimulus type on categorization, z=.75, *NS*. Together these results show that both preferrer groups were equally good at categorizing both types of stimuli. Next, I compared children's accuracy at categorizing novel stimuli inside and outside of the learned category space to examine the extent to which children had learned to attend to a dimension of similarity versus a region of the category space.

It was predicted that the brightness preferrers who learned brightness categories would either be equally accurate at categorizing all novel stimuli or be even more accurate in categorizing novel stimuli from outside the category than from inside the category because they should be attending to the dimension of brightness rather than specific areas of the category space. In fact, as can be seen in Figure 26a, brightness preferrers who learned brightness categories were equally accurate at categorizing both types of novel stimuli. This was supported by a mixed logistic regression model of the effect of stimulus type (novel inside or novel outside) on categorization accuracy. This model revealed no effect of stimulus type, z=-1.46, *NS.* However, while it was predicted that those brightness preferrers who learned size categories might be more accurate at categorizing novel stimuli from inside the category space because they might not have learned to attend to the entire dimension of size, as can be seen in the figure, they were actually more accurate at categorizing stimuli outside the category space. This was supported by the results of a mixed logistic regression model of the effects of stimulus type on categorization accuracy that revealed a significant effect of stimulus type such that brightness preferrers who learned size categories were actually significantly more accurate to categorize novel stimuli from outside the category space, z=-2.95, p<.05. It could be the case that brightness preferrers are worse at making fine grained discriminations between similarly sized objects, and so are more accurate at categorizing stimuli outside the space, which are maximally different from the (more ambiguous) stimuli near the category boundary. Together, these results suggest that brightness preferrers may be attending to relevant dimensions of similarity rather than specific stimuli or regions of the category space.

It was predicted that size preferrers would also show evidence of attending to dimensions, and would be either be equally accurate at categorizing stimuli inside or outside of the space, or even better at categorizing stimuli from outside the space. In fact, as can be seen in Figure 26b, they were actually equally accurate in categorizing novel stimuli. This was supported by a mixed logistic regression model of the effects of stimulus type on categorization accuracy that revealed no effect of stimulus type, z=3-1.01, *NS.* As can also be seen in the figure, however, those size preferrers who learned size categories were actually more accurate at categorizing stimuli outside the category space. This was supported by a mixed logistic regression model of the effects of stimulus type on categorization that revealed a significant effect of stimulus type, z=-2.05, p<.05. Together these results demonstrate that the size preferrers might not have been merely learning about specific regions of the category space.

Strongly versus weakly learned areas of the category

The results of Experiments 1 and 2 demonstrated that no group of children ever showed an advantage for making discriminations in strongly- versus weakly-learned areas of the category space. This could mean that this was not a sensitive measure of dimensional attention. Thus it was not expected that either type of preferrer would show effects of stimulus location on discrimination. In fact, brightness preferrers who learned brightness categories were equally good at making both kinds of discriminations. This was supported by the results of a mixed logistic regression model of the interaction between stimulus location (strongly- versus weakly-learned areas of the space) and discrimination type (between, within-relevant, within-irrelevant) which revealed no significant effect stimulus location z=-.55, *NS,* and no interactions of location and type. Similarly, brightness preferrers who learned size categories were equally good at making both kinds of discriminations. This was supported by the results of a mixed logistic regression model of the interaction between stimulus location and discrimination type that revealed no effect of location, z=-.33, *NS,* and no interactions with location. These results provide some indication that brightness preferrers were attending to dimensions rather than specific regions of the category space.

Similarly, size preferrers who learned brightness categories were equally accurate at making discriminations in both strongly- and weakly-learned areas of the category space. This was supported by the results of a mixed logistic regression model of the interaction between stimulus location (strongly- versus weakly-learned) and discrimination type (between, within-relevant, and within-irrelevant) that revealed no significant effect of stimulus location, z=.15, *NS,* and no interactions with location. Likewise, size preferrers who learned size categories were equally accurate at making discrimination in both locations. This was supported by the results of a mixed logistic regression model of the interaction between stimulus location and discrimination type that revealed no significant effect of location, z=-.24, *NS.* Together these data suggest that size preferrers might not be attending just to specific regions of the learned category space, but instead may attend to the relevant dimension.

Discrimination Threshold

Results of Experiments 1 and 2 revealed that holistic and dimensional classifiers' discrimination accuracy was influenced by the size of the difference between the to-be-discriminated stimuli. It was therefore expected that both types of preferrers would show increased accuracy for larger differences. Brightness preferrers who learned brightness categories were more accurate at making discriminations across larger distances. This was supported by a mixed logistic regression model of the interaction between threshold size (.5, 1, or 1.5x JND) and discrimination type (between, within-relevant, or within-irrelevant) that revealed a significant effect of threshold size such that the bigger the difference between stimuli, the more accurate children were at discriminating between them, z=3.58, p<.001. There was also a significant interaction such that this effect was less strong for within-category discriminations across the irrelevant dimension, z=-2.01, p<.05. This suggests that brightness prefers were relatively poor at making size discriminations regardless of how different the stimuli were from each other. Similarly, brightness preferrers who learned size categories were also more accurate at making discriminations across larger distances. This was supported by a mixed logistic regression model of the interaction between threshold size and discrimination type that revealed a significant effect of threshold size such that the bigger the difference between the stimuli, the more accurate children were at discriminating between them, z=3.47, p<.001 and no significant interactions. This result suggests that brightness preferrers who learned to attend to size simply increase accuracy the more different discrimination stimuli become.

It was thought that size preferrers would show a similar pattern of results. However, size preferrers who learned brightness categories actually were no more accurate at discrimination as threshold size increased. This was supported by a mixed logistic regression model of the interaction between threshold size and discrimination type that revealed no significant effect of threshold size, z=.87, *NS,* and no interactions. Those size preferrers who learned size categories, on the other hand, were more accurate as discrimination as threshold size increase. This was supported by a mixed logistic regression model of the interaction between threshold size and discrimination type that revealed a significant effect of threshold size such that the more different stimuli were, the more accurate children were in discriminating them, z=5.05, p<.0001, and a significant interaction such that children showed less of an effect for within-category discriminations along the irrelevant dimension. This result is similar to that of the brightness preferrers who learned brightness categories in that when a child's preferred dimension matches the dimension relevant for categorization, children did not seem to show typical attentional changes to the irrelevant dimension. It is unclear what exactly this means about their dimensional attention. The results of the blocking analysis might clarify this issue.

Blocking

It was expected that if children did learn to selectively attend to the category-relevant dimension, then they would show a cost for switching between discrimination types in the mixed blocks. Because both preferrer types had shown some indication of selective attention to dimensions in the other tasks, it was predicted that they would show such a cost. As can be seen in Figure 27a, brightness preferrers who learned brightness categories do show some cost: they are more accurate at making between-category discriminations in the context of single-type blocks than mixed blocks (blue bars), more accurate at making within-category discriminations across the irrelevant dimension in the context of single-type blocks than mixed blocks (orange bars), but equally accurate at making within-category discriminations across the relevant dimension in both types of blocks (green bars). This was supported by the results of a mixed logistic regression models of the effects of block type on accuracy that revealed brightness preferrers who learned brightness categories were marginally more accurate at making between-category discriminations in relevant blocks, z=1.88, p<.10, marginally more accurate at making within-category discriminations across the irrelevant dimension in single-type blocks, z=-1.66, p<.10, but equally accurate at making within-category discriminations in both types of block, z=-.14, *NS.* This suggests that overall, these children have learned to selectively attend to dimensions.

As can also be seen in the figure, brightness preferrers who learned size categories showed no switch costs. This was supported by mixed logistic regression models of the effects of block type on discrimination accuracy that revealed no effect of block on accuracy for between-category discriminations, z=56, *NS,* for within-category discriminations along the relevant dimension, z=-.76, *NS,* or for within-category discriminations along the irrelevant dimension, z=-.29, *NS.* These results suggest that brightness preferrers who learned size categories may not have learned to selectively attend to their category-relevant dimension.

Similarly, as can be seen in Figure 27b, size preferrers who learned brightness categories did not show any switch costs either. This was supported by the results of mixed logistic regression models of the effects of block type on discrimination accuracy that revealed no effect of block on accuracy for between-category discrimination, z=-.03, *NS,* or for within-category discriminations along the irrelevant dimension, z=-.86, *NS,* or along the relevant dimension, z=-.16, *NS.* This suggests that size preferrers who learned brightness categories might not have fully learned to selectively attend to their category-relevant dimension.

Finally, as can also be seen in the figure, size preferrers who learned size categories only showed a switch cost for within-category discriminations along the irrelevant dimension (orange bars). This was supported by mixed logistic regression models of the effects of block on discrimination accuracy that revealed no effect of block on accuracy for between-category discrimination, z=1.58, *NS,* or for within-category discriminations along the relevant dimension, z=.86, *NS,* but did show an effect of block for within-category discriminations along the irrelevant dimension, z=-2.12, p<.05. This suggests that even though these size preferrers learned about categories organized by similarity in size, they appear to be selectively attending to brightness such that they a cost for switching between making these discriminations in a mixed block relative to a single-type block. Together the results of the results of the blocking analysis lend some support to the idea that children learned to selective attend to dimensional similarity, but not always to the dimension related to the learned category.

Overall, the results of Experiment 3 suggest that category learning increases preferrers' selective attention to category-relevant dimensions. This could be taken to support the idea that preferrers have sticky attention. Perhaps, there is nothing inherently different about these two groups, but rather the size preferrers randomly got stuck on size during the pretest triad while the brightness preferrers got stuck on the brightness. This would mean that learning a dimensional category organized by the opposite dimensions can get them to switch the dimension to which they attend. However, the results also suggest that brightness and size dimensions are not equivalent and that children who prefer attending to one dimension over the other are not equivalent. This would support the idea that size is the more difficult dimension to attend to. Those children that preferred size were able to quickly learn either brightness or size categories whereas those who preferred brightness were much quicker to learn brightness categories than size categories. However, it also seems as though these two dimensions are different from each other in ways that extend beyond the attentional maturity of the children that prefer them. Even size preferrers, for example, did not show increased category-relevant discrimination and decreased category-irrelevant discrimination if they learned size categories. Given these differences, I next ask, how do labels affect these children's selective attention to category-relevant dimensions.

Experiment 4

The purpose of this experiment was to examine the effects of redundant labels on preferrers' category learning and dimensional attention. The results of Experiment 2 indicated that incidental labels can scaffold selective attention even in holistic classifiers. The results of Experiment 3 suggest that while both brightness preferrers and size preferrers show an effect of category learning on their perceptual discrimination and subsequent similarity classification, they are not equally skilled at the initial category learning. In particular, brightness preferrers are much quicker to learn brightness categories than size categories, while size preferrers are equally quick to learn both kinds of categories and faster overall at category learning than brightness preferrers. This opens up the interesting possibility that size is a more difficult dimension to attend to, and when children prefer to attend to it they can easily selectively attend to either size or brightness. When children prefer to attend to brightness, on the other hand, they have difficulty attending to size instead and are much slower to learn size categories. The question that remains is will the presence of a label help brightness preferrers learn size categories as quickly as they learn brightness categories. If this is the case, it would suggest that labels enhance category representations allowing children to focus on category-relevant dimensions. However, recent work on adults' category learning suggests that when learners are biased to attend to a specific dimension, labels increase this bias making it more difficult to learn categories organized by another dimension (Brojde, Porter, &

Colunga, 2011). Thus, it is possible that labels could make it even more difficult for brightness preferrers to selectively attend to size in the category learning task. As size preferrers were equally good at learning both types of categories, it could be that labels facilitate category learning equally for both types.

Method

Participants

Participants were 28 5- to 8-year old children. 1 child did not compete the entire experiment and was excluded from these analyses. There were 27 children in the final group.

Procedure

The procedure was identical to that of Experiment 2—the label experiment. The only difference was that children in this experiment all used a preferrer strategy on the triad classification test. Children were divided into classifier groups based on performance in the triad pretest. There were 11 brightness preferrers and 16 size preferrers. See Table 1 for participant information.

Results and Discussion

Category learning

As can be seen in Figure 28, brightness preferrers were faster to learn brightness categories than size categories while size preferrers were equally quick to learn either type. This was supported by the results of a mixed linear regression model of the interaction between classifier type (brightness or size preferrer) and category structure (brightness or size) that revealed a significant effect of classifier type such that size preferrers were significantly faster to reach criterion than brightness preferrers, t=-31.71, p<.0001, a significant effect of category structure such that children were overall faster to learn brightness categories than size categories, t=-37.29, p<.0001, and a significant interaction such that size preferrers were significantly faster to learn size categories than brightness preferrers were, t=26.13, p<.0001. Clearly, the presence of a label did not support brightness preferrers' selective attention to size or facilitate their learning of size categories.

Discrimination

Overall, children were quite accurate in discriminating stimuli. In particular, the brightness preferrers who learned brightness categories, M=.79, $t(5)$=5.88, p<.01, the brightness preferrers who learned size categories, M=.84, $t(3)$=22.42, p<.0001, the size preferrers who learned brightness categories, M=.84, $t(6)$=8.78, p<.001, and the size preferrers who learned size categories, M=.84, $t(7)$=22.56, p<.0001, were all significantly better than chance at discriminating stimuli in this task.

If labels increase selective attention to category-relevant dimensions than I would expect children to show an advantage for relevant versus irrelevant discriminations. The results of Experiment 3 demonstrated that brightness preferrers who learned brightness categories showed this pattern. Thus it was expected that those in the current experiment should do so as well as they had learned their categories relatively quickly. However, because the brightness preferrers had more difficulty learning size categories, it was expected that they would not show this pattern and would instead, as in Experiment 3, actually show enhanced accuracy for discriminations along the irrelevant dimension. As can be seen in Figure 29a, brightness preferrers who learned brightness categories were in fact more accurate at making between-category discriminations than within-category discriminations along either the relevant or irrelevant dimension. This was supported by the results of a mixed logistic regression model of the effects of discrimination type (between, within-relevant, and within-irrelevant) on discrimination accuracy. This model showed that brightness preferrers who learned brightness categories were significantly more accurate at making between category discriminations than within category discriminations along either the relevant, z=3.74, p<.001 or irrelevant dimensions, z=3.31, p<.001. These children were also significantly more accurate at making within-category discriminations along the relevant than the irrelevant dimensions, z=-2.50, p<.05, demonstrating that, overall, they were more accurate at discriminating across the relevant than the irrelevant dimension. As can also be seen in the figure, brightness preferrers who learned size categories, on the other hand, were significantly more accurate at making discriminations on the irrelevant dimension than between category discriminations. This was supported by the results of a mixed logistic regression model of the effects of discrimination type on accuracy that revealed a significant effect of discrimination type such that children were more accurate at making within-category discriminations along the irrelevant dimension that between category discriminations, z=-2.17, p<.05, but equally accurate at making within-category discriminations along both irrelevant and irrelevant dimensions, z=1.54, *NS.* These children were no more accurate at making within-category discriminations along the relevant or irrelevant dimension, z=1.54, *NS.* Similar to the results of Experiment 3, this suggests that brightness preferrers are quite accurate at making discriminations along brightness regardless of whether it is the relevant or irrelevant dimension for their learned category.

It was predicted that because they had learned brightness and size categories equally quickly, size preferrers would show evidence of selective attention to relevant dimensions in the discrimination task. As can be seen in Figure 29b, size preferrers who learned brightness categories were more accurate at making between-category discriminations than within-category discriminations along the relevant but not irrelevant dimension. This was supported by a mixed logistic regression model of the effects of discrimination type on accuracy that revealed a significant effect of discrimination type such that these children were marginally more accurate at making between-category discriminations than within-category discriminations along the relevant dimension, *z*=1.73, *p*<.10, but equally accurate at making between category discriminations as within-category discriminations along the irrelevant dimension, *z*=-.96, *NS.* These children were no more accurate at making within-category discriminations along the relevant or irrelevant dimension, z=.28, *NS.* This suggests that size preferrers who learned brightness categories did show increases in attention to the relevant dimension, but also maintained a high degree of attention to the irrelevant dimension (which in this case happens to be their preferred dimension). As can also be seen in the figure, however, size preferrers who learned size categories did not show any differences in accuracy for different trial types. This was supported by a mixed logistic regression model of the effect of trial type on discrimination accuracy showing that these children were equally accurate at making between-category and within-category discriminations along both the relevant, *z*=.43, *NS,* and irrelevant dimensions, *z*=-.60, *NS*. These children were no more accurate at making within-category discriminations along the relevant or irrelevant dimension, z=.68, *NS.* Next I examine the results of the posttest triad task to see how far reaching these changes to dimensional attention were.

<u>Posttest triad</u>

If children learned to selectively attend to category-relevant dimensions, then it was expected that they would show increases in dimensional responding related to the category they had learned. Brightness preferrers who learned brightness categories should not increase in dimensional responding from pre- to posttest triad, but those who learned size categories might if they had learned to attend to size. Similarly, size preferrers who learned size categories should not increase in dimensional responding from pre- to posttest triad, but those who learned brightness categories might if they had learned to attend to brightness. In fact, as can be seen in Figure 30a, brightness preferrers who learned brightness categories did not show an increase in dimensional responding from pre- to posttest triad. This was supported by the results of a mixed logistic regression model of the effect of test time (pretest or posttest) on dimensional responding that revealed no effect of test time, $z=.13$, *NS.* Similarly, as can be seen in the figure, brightness preferrers who learned size categories also did now show an increase in dimensional responding from pre- to posttest triad. This was supported by the results of a mixed logistic regression model of the effect of test time on dimensional responding that revealed no effect of test time, $z=-.43$, *NS.* These results suggest that the presence of a label in category learning did not increase brightness preferrers' overall dimensional attention.

As can be seen in Figure 30b, size preferrers who learned brightness categories also did not show any increases in dimensional responding from pre- to posttest triad. This was supported by the results of a mixed logistic regression model of the effect of test time on dimensional responding that revealed no effect of test time, $z=.33$, *NS*. Similarly, as can also be seen in the figure, size preferrers who learned brightness categories did not show any increases in dimensional responding from pre- to posttest triad. This was supported by the results of a mixed logistic regression model of the effect of test time on dimensional responding that revealed no effect of test time, $z=.10$, *NS*. Together, these results suggest that the presence of a label in category learning did not increase size preferrers' overall dimensional attention. Next, I examined the specific pattern of dimensional responses for these groups of children to see if the specifics of the learned categories influenced their likelihood of selecting dimensional responses on category-relevant trials.

It was predicted that brightness preferrers who learned brightness categories would remain brightness preferrers. However, as can be seen in Figure 31a, these children actually showed a slight increase in dimensional responding on size trials and a decrease in dimensional responding on brightness trials. This was supported by a mixed logistic regression model of the interaction between test time (pretest or posttest) and trial type (brightness or size match) that revealed a significant effect of trial type such that children were overall more likely to select dimensional matches on brightness trials, z=-7.14, p<.0001, and a significant interaction such that they were more likely to select dimensional matches on size trials during the posttest than during the pretest, z=4.29, p<.0001. This result was completely unexpected. There does not appear to be a reason why brightness preferrers who had learned brightness categories would increase the number of dimensional matches they selected on size trials. As can be seen in Figure 31a, the brightness preferrers who learned size categories demonstrated this same effect (although it was the expected result for these children). This was supported by the results of a mixed logistic regression model of the interaction between test time and trial type that revealed a significant effect of trial type such that children were much more likely overall to select dimensional matches on brightness trials, z=-7.53, p<.0001, and a significant interaction such that children were more likely to select dimensional matches on size trials during the posttest, z=2.24, p<.05. On the one hand, if children who learned size categories learned to attend to dimensional similarity in size, we would expect them to be more likely to select dimensional matches on posttest size trials. On the other hand, however, even the brightness preferrers who learned brightness categories were more likely to select dimensional matches on posttest size trials. And these children should not have had a reason to increase selective attention to size. Thus it is not clear from this test whether either group of brightness preferrers had learned to selectively attend to the category-relevant dimension.

It was expected that the size preferrers who learned brightness categories might show an increase in dimensional responses on brightness trials, but that those who learned size categories would show no differences in dimensional responding on either trial type. As can be seen in Figure 31b, however, those size preferrers who learned brightness categories did not show any differences in their pattern of dimensional responding from pre- to posttest triad. This was supported by the results of a mixed logistic regression model of the interaction between test time and trial type that revealed no effect of trial type, z=.01, *NS*, and no interaction, z=-.01, *NS*. Similarly, as also can be seen in the figure, those size preferrers who learned size categories did not show any differences in their pattern of responding from pre- to posttest triad. This was supported by the results of a mixed logistic regression model of the interaction between test time and trial type that revealed no effect of trial type, z=.01, *NS.* These results demonstrate that size preferrers remained size preferrers during the posttest regardless of the category they had learned. The results of the posttest triad task are puzzling. Why, if preferrers had learned in the no label experiment to attend to category-relevant dimensions, would they not do so in the label experiment? Furthermore, the results of Experiment 2 demonstrated that the presence of a label increased holistic and dimensional classifiers' attention to category-relevant dimensions. Why should preferrers be any different? I next examined preferrers' performance in the categorization and discrimination tasks in hopes of clarifying these results.

In-depth analysis of dimensional attention

The results of Experiment 4 so far are mixed. Brightness preferrers are faster to learn brightness categories than size categories—just as they were in Experiment 3—and they show a tendency to be better at making discriminations along brightness, regardless of whether it is the category-relevant dimension—just as in Experiment 3. However, unlike in Experiment 3, brightness preferrers do not show a tendency to increase their dimensional responses on the posttest triad trials that correspond to the organization of the category they learned. Instead, both brightness preferrers who learned brightness categories and those who learned size categories showed an increase in selections of dimensional matches on size trials during posttest. How did the presence of a label change their attention in this way?

The size preferrers, on the other hand, were equally likely to learn both brightness and size categories quickly, but only showed effects of increased attention to category relevant dimensions when they learned brightness categories (as was the case in Experiment 3). I next took a more in-depth look at the results of categorization and discrimination to clarify the relationship between labels and dimensional attention in these preferrer children.

Category test

Overall, children in both groups who reached learning criterion were not very accurate in the categorization test relative to participants in the previous experiments. In particular, the brightness preferrers who learned brightness categories, M=.63, $t(5)$=1.48, *NS*, and the brightness preferrers who learned size categories, M=.64, $t(3)$=2.08, *NS*, were not significantly different than chance. The size preferrers who learned brightness categories were only marginally better than chance, M=.76, $t(6)$=2.24, p<.10, and only the size preferrers who learned size categories were significantly better than chance at categorizing stimuli in this task, M=.82, $t(7)$=9.53, p<.0001.

As no group of children in Experiments 1-3 showed any differences in accuracy when categorizing novel versus familiar stimuli, it was predicted that the children in Experiment 4 would not show any differences either. As can be seen in Figure 32a, brightness preferrers who learned brightness categories were equally accurate in categorizing novel and familiar stimuli. This was supported by the results of a mixed logistic regression model of the effect of trial type (familiar versus novel) on accuracy that revealed no significant effect of trial type, z=.34, *NS*. Similarly, as can be seen in the figure, brightness preferrers who learned size categories were equally accurate in categorizing both types of stimuli. This was supported by the results of a mixed logistic regression model of the effect of trial type on accuracy that revealed no significant effect of trial type, z=-.51, *NS*. As can be seen in Figure 32b, the same was true for size preferrers. This was supported by the results of mixed logistic regression models of the effect of trial type on accuracy that revealed no significant effect to trial type on the accuracy of either size preferrers who learned brightness categories, z=.18, *NS* or those who learned size categories, z=-.16, *NS*.

Next, I examined categorization with respect to accuracy categorizing novel stimuli from inside and outside the category space. It was expected that if children learned to attend to relevant dimension for categorization, rather than specific regions of the space, they should either be equally good at categorizing both types of stimuli or would actually be better at categorizing stimuli from outside of the space. As can be seen in Figure 33a, brightness preferrers who learned brightness categories were equally accurate in categorizing both types of stimuli. This was supported by the results of a mixed logistic regression model of the effects of stimulus type (novel inside or novel outside the category space) on accuracy that revealed no significant effect of stimulus type, z=.58, *NS*. However, as can be seen in the figure, those brightness preferrers who learned size categories were more accurate at categorizing stimuli from outside of the category space. This was supported by the results of a mixed logistic regression model of the effects of stimulus type on accuracy that revealed a marginal effect of stimulus type such that children were more accurate at categorizing stimuli from outside the category space, z=-1.73, p<.10. Together these results suggest that brightness preferrers might not have learned merely about specific regions of the category space, but instead may have learned about the relevant dimension for categorization.

As can be seen in Figure 33b, size preferrers who learned brightness categories also were more accurate at categorizing stimuli from outside the category space. This was supported by the results of a mixed logistic regression model of the effects of stimulus type on accuracy that revealed a significant effect of stimulus type such that children were more accurate at categorizing stimuli from outside the category space, z=-2.69, p<.01. Similarly, as can be seen in the figure, size preferrers who learned size categories also were more accurate at categorizing stimuli from outside the category space. This was supported by the results of a mixed logistic regression model of the effects of stimulus type on accuracy that reveal a significant effect of stimulus type such that children were more accurate at categorizing stimuli from outside the category space, z=-2.46, p<.05. This again suggests that preferrers might not merely have learned about specific stimuli or specific regions of the category space, but may instead have learned to selectively attend to category-relevant dimensions. The remaining analyses of the discrimination test should help clarify the extent to which children were selectively attending to dimensional similarity.

Strongly- versus weakly-learned areas of the category

No children in Experiments 1-3 showed any differences in discrimination accuracy for stimuli from strongly- versus weakly-learned areas of the category space. Thus, it was predicted that no children in Experiment 4 would either. In fact, no group did. This was supported by mixed logistic regression models of the interaction between stimulus location (strong- versus weakly-learned areas) and discrimination type (between, within-relevant, or within-irrelevant) on discrimination accuracy that revealed no significant effect of stimulus location or interaction on accuracy for brightness preferrers who learned brightness categories, z=-1.04, *NS*, brightness preferrers who learned size categories, z=-.20, *NS*, size preferrers who learned brightness categories, z=.55, *NS*, or size preferrers who learned size categories, z=-.08, *NS*. These results suggest that the children might not simply have learned about specific regions of the category space, but rather may have been learned something about the category-relevant dimensions.

Discrimination step size

The previous experiments showed a general trend for children to be more accurate at making discriminations when there were bigger differences between stimuli. Thus it was predicted that children in this experiment would show a similar effect. In fact, each group did show an increase in accuracy corresponding to an increase in threshold size. This was supported by the results of mixed logistic regression models of the interaction between threshold size (.5, 1, or 1.5x JND) and discrimination type (between, within-relevant, or within-irrelevant) that revealed significant effects of threshold size on accuracy (but no interactions) for brightness preferrers who learned brightness categories, z=3.76, p<.001, brightness preferrers who learned size categories, z=3.47, p<.001, size preferrers who learned brightness categories, z=5.06, p<.0001, and size preferrers who learned size categories, z=6.25, p<.0001. These results suggest that, as in previous experiments, children are more accurate at discriminating stimuli the more different they are. This does not clarify, however, whether the presence of a label had any effect on children's dimensional attention. Thus, in the final analysis, I examine the results of discrimination with respect to blocking to clarify this issue.

Blocking

If children are selectively attending to category-relevant dimensions, then they should show a cost for making discriminations within the context of mixed blocks. As can be seen in Figure 34a, brightness preferrers who learned brightness categories did not show any advantage for making discriminations in single-type blocks over mixed blocks. This was supported by mixed logistic regression models of the effects of block type (single-type or mixed) on discrimination accuracy that revealed no effects of block such that brightness preferrers who learned brightness categories were no more accurate at making between category discriminations in the context of single-type blocks than mixed blocks, z=.77, *NS*, no more accurate at making within-category discriminations along the relevant dimension in the context of single-type than mixed blocks, z=.39, *NS*, and no more accurate at making within-category discriminations along the irrelevant dimension in the context of single-type than mixed blocks, z=.84, *NS*. This would suggest that the brightness preferrers who learned brightness categories might not have been selectively attending to the relevant dimension.

On the other hand as can also be seen in the figure, those brightness preferrers who learned size categories did show a switch cost for making between-category discriminations, although they did not show a cost for within-category discriminations of either type. This was supported by the results of mixed logistic regression models of the effect of block on discrimination accuracy that revealed an effect of block such that these brightness preferrers who learned size categories were significantly more accurate at making between-category discriminations in the context of single-type blocks than mixed blocks, $z=2.50$, $p<.05$, but no more accurate at making within-category discriminations along the relevant dimension in single-type blocks than mixed blocks, $z=.23$, *NS*, or within-category discriminations along the irrelevant dimension in single-type blocks than mixed blocks, $z=.73$, *NS*. This suggests that brightness preferrers who learned size categories might have slightly increased their selective attention to category-relevant dimensions, as evidenced by their switch cost for between-category discriminations.

As can be seen in Figure 34b, size preferrers who learned brightness categories did not show any advantage for making discriminations in single-type blocks versus mixed blocks. This was supported by the results of mixed logistic regression models of the effect of block on discrimination accuracy that revealed that size preferrers who learned brightness categories were no more accurate at making between-category discriminations in the context of single-type blocks than mixed blocks, $z=.09$, *NS*, no more accurate at making within-category discriminations along the relevant dimension in the context of single-type blocks than mixed blocks, $z=1.32$, *NS*, and no more accurate at making within-category discriminations along the irrelevant dimension in the context of single-type blocks than mixed blocks, $z=-.26$, *NS*. As can be seen in Figure 34b, however, the size preferrers who learned size categories did show a switch cost for between-category discriminations, but not for either type of within-category discriminations (just as the brightness preferrers who learned size categories). This was supported by the results of mixed logistic regression models of the effects of block type on discrimination accuracy that revealed a significant effect of block such that size preferrers who learned size categories were significantly more accurate at making between-category discriminations in the context of single-type blocks than mixed blocks, z=2.67, p<.001, but equally accurate at making within-category discriminations along the relevant dimension in the context of single-type and mixed blocks, z=.03, *NS*, and equally accurate at making within-category discriminations along the irrelevant dimension in the context of single-type and mixed blocks, z=-.34, *NS*. This suggests that brightness preferrers who learned size categories might have slightly increased their selective attention to category-relevant dimensions, as evidenced by their switch cost for between-category discriminations.

Conclusions

Overall, the results of Experiment 4 suggest that labels do not qualitatively alter brightness or size preferrers' category learning or selective attention to dimensions. As in Experiment 3, the results show that size preferrers are quicker to learn both brightness and size categories than brightness preferrers. The results of both preferrer groups add to evidence from the first three experiments that there are differences in the way children learn about brightness and size categories that go above and beyond children's classification type. In particular, the results of Experiment 4 again demonstrated that only children who learned brightness categories show subsequent enhanced accuracy for between-category discriminations relative to within-category discriminations. The results of posttest triad were mixed however. In Experiment 3, preferrers who learned categories organized by the dimension opposite of their preference showed increases in dimensional responding corresponding to the structure of their learned category, suggesting preferrers learned to attend to the category-relevant dimension of similarity. However, in Experiment 4, brightness preferrers who learned about brightness actually showed an increase in dimensional responding on size trials while size preferrers who learned about brightness did not show any increases in dimensional responding on either trial type. Why would preferrers show an increase in dimensional attention in the no label experiment? I next directly compared the results of each preferrer group in the two experiments to clarify this apparent contradiction.

Comparison between Preferrers in Experiments 3 and 4

The prior analyses of Experiments 3 and 4 suggest that there are no qualitative differences in the way preferrers learn categories with or without a label. In particular, in both experiments, size preferrers were equally quick to learn both brightness and size categories, while brightness preferrers were significantly faster to learn brightness categories than size categories. Similarly, both brightness and size preferrers who learned brightness categories in either experiment showed increased accuracy for between-category relative to within-category discriminations. However, there do appear to be some differences in the pattern of results that each group shows with respect to the speed with which children learned dimensional categories and the direction of dimensional attention in the posttest triad task. Thus, direct comparison between the two experiments should be useful for understanding the nature of any differences.

Comparison Analyses and Discussion

As in the previous chapters, I first conducted analyses to answer my three primary questions regarding speed to reach criterion in category learning, overall accuracy in discrimination along relevant and irrelevant dimensions, and changes to dimensional responding in similarity classification. I then examined the more in-depth questions about accuracy in categorization and discrimination to explore potential differences in the extent to which children were selectively attending to dimensional similarity. Comparing the performance of children in each preferrer group between the two experiments should clarify what the role of labels are in supporting dimensional attention above and beyond category learning.

Category learning

The results of Experiment 1 and 2 demonstrated that holistic classifiers were significantly faster to learn dimensional categories in the presence of a label. Although there were no a priori predictions about exactly how labels would affect preferrers' category learning, it seemed possible that labels may facilitate their category learning too. However, the results of the direct comparison between preferrers in Experiments 3 and 4 revealed that the opposite was true. In fact, both preferrer groups, regardless of the type of category they learned, were slower to reach criterion in the label experiment than in the no label experiment. This was supported by the results of mixed linear regression models of the interaction between experiment (label versus no label) and category type (brightness or size) on the number of blocks it took children to reach criterion. In particular, a model of brightness preferrers' performance revealed a significant effect of experiment such that children were significantly slower to reach criterion in the label experiment, $t=23.11$, $p<.0001$, a significant effect of category type such that children were slower overall to reach criterion when learning size categories than brightness categories, $t=-63.92$, $p<.0001$, and a marginal interaction such that the difference in the number of trials to reach criterion for learning brightness and size categories was greater in the label condition, $t=1.89$, $p<.10$. Similarly, the equivalent model of size preferrers' performance revealed a significant effect of experiment such that children were significantly slower in the label experiment, $t=29.59$, $p<.0001$, a significant effect of category type such that children were faster overall to learn brightness categories than size categories, $t=-3.19$, $p<.01$, and a significant interaction such that the difference in the number of trials to learn size and brightness categories was larger in the label experiment, $t=5.92$, $p<.0001$. Together, the results of these models show that both preferrer groups are slower to learn categories in the label experiment and are especially slow to learn size categories in the label experiment.

On the one hand, that the brightness preferrers would have increased difficulty for learning size categories would seem to support the idea that labels can increase our existing attentional biases making it even harder to switch attention to another dimension (cf. Brojde et al., 2011). However, this would not explain why the size preferrers were also slower to learn size categories in the label experiment—unless both preferrer groups have some bias to attend to brightness that was exaggerated by the presence of a redundant label. This would support the idea that brightness is an easier dimension to attend to than size is. The size preferrers are good at attending to size, especially relative to the brightness preferrers, but it is still easier for them to attend to brightness. The results of the discrimination comparison might clarify what the nature of the differences between dimensions is.

Discrimination

The results of Experiment 3 and 4 demonstrated that both groups of preferrers showed enhanced between-category discrimination relative to within-category discrimination if they learned brightness categories. It was therefore thought that there would not be any quantitative differences between the two experiments for those who learned brightness categories. Similarly, I had no a priori predictions about differences between the groups in each experiment with respect to discrimination accuracy. Inasmuch as those in the label experiment took longer (and presumably had more difficulty learning the categories), I thought that perhaps they would show less of an advantage for between-category over within-category discriminations. The results of mixed logistic regression models of the interaction between experiment and discrimination type (between, within-relevant, and within-irrelevant) revealed that the brightness preferrers did not in fact show any differences in discrimination related to experiment whether they had learned brightness categories, z=-.66, *NS,* or size categories, z=.42, *NS.* Similarly, size preferrers who learned size categories did not show any differences related to experiment either,

z=.82, *NS.* This was not the case for size preferrers who learned brightness categories, however. These children showed a smaller increase in accuracy for between-category discriminations in the label experiment, z=-2.47, p<.05. It is unclear why this would be the case.

Posttest triad

The results of Experiment 3 suggested that both preferrer groups showed changes in their dimensional responding to the nonpreferred dimension on the posttest triad if they had learned categories organized by that dimension. However, in Experiment 4, children did not do so. Thus, it was thought that those in the no label experiment would show significantly larger changes in dimensional responding that related to the structure of the categories they had learned. In fact mixed logistic regression models of the effects of experiment on the change in dimensional responses from pre- to posttest triad revealed that both brightness preferrers who learned brightness categories, z=-4.04, p<.0001, and those who learned size categories, z=-3.88, p<.0001, showed significantly smaller increases in dimensional responding if they were in the label experiment. Similarly, a mixed logistic regression model of changes in the dimensional responding of size preferrers who learned brightness categories showed a significant effect of experiment such that those in the label condition showed significantly smaller increases in dimensional responding from pre- to posttest triad, z=-3.16, p<.01. However, a mixed logistic regression model of the changes in the dimensional responding of size preferrers who learned size categories showed a significant effect of experiment such that those in the label condition showed *larger* increases in dimensional responding from pre- to posttest triad, z=3.13, p<.01.

In-depth analysis of dimensional attention

Next I compared children's performance in the two experiments in the more in-depth measures of categorization and discrimination in order to clarify the extent to which children were selectively attending to dimensional similarity.

Categorization

The results of Experiment 3 and 4 demonstrated that all preferrers were equally accurate at categorizing familiar and novel stimuli. Thus, it was expected that there would be no differences between the two experiments with respect to this measure. In fact, mixed logistic regression models of the interaction between experiment (label versus no label) and trial type (familiar versus novel) revealed that brightness preferrers who learned brightness categories, z=-1.31, *NS,* brightness preferrers who learned size categories, z=-1.27, *NS*, size preferrers who learned brightness categories, z=-.37, *NS*, all were equally accurate at categorizing familiar and novel stimuli in the two experiments. However, size preferrers who learned size categories, were significantly more accurate overall at categorizing both kinds of stimuli if they were in the label experiment, z=2.15, p<.05.

I next compared accuracy across the two experiments with respect to categorizing novel stimuli from inside the category space versus those outside the category space. A mixed logistic regression model of the interaction between experiment (label versus no label) and trial type (novel inside or novel outside) on the categorization accuracy of brightness preferrers who learned brightness categories revealed a significant interaction such that children in the label experiment were less accurate at categorizing stimuli outside the category space, z=2.73, p<.01. Brightness preferrers who learned size categories, on the other hand, showed a significant effect of trial type such that in both experiments, children were more accurate at categorizing stimuli from outside the category space, z=-2.14, p<.05. The size preferrers who learned brightness categories showed a marginal interaction between experiment and trial type such that those in the label experiment showed less of an advantage for categorizing stimuli for outside versus inside the space, z=-1.81, p<.10. Finally the size preferrers who learned size categories showed a significant effect of experiment such that those children in the label experiment were overall more accurate at categorization than those in the no label experiment, z=3.18, p<.01, and a significant effect of trial type such the children in both experiments were less accurate at categorizing stimuli from inside the space than outside, z=-2.77, p<.01. While these results suggest that, in general, children's categorization accuracy was affected by the presence of a label, it is not clear exactly how this relates to whether or not they were selectively attending to dimensional similarity.

Strongly- versus weakly-learned areas of the space

All children in previous experiments were equally good at making discriminations in both strongly- and weakly-learned areas of the category space. Thus it was thought that there would be no differences between the two experiments with respect to preferrers' accuracy for making these discriminations. A mixed regression model of the interaction between stimuli location (strongly-learned or weakly-learned area) and discrimination type (between, within-relevant, or within-irrelevant) revealed that brightness preferrers who learned brightness categories were equally accurate at making discriminations in both locations in the two experiments, z=-.58, *NS.* Brightness preferrers who learned size categories were also equally accurate at making discriminations in both locations in the two experiments, z=.47, *NS.* Size preferrers who learned brightness categories, z=.42, *NS,* and those who learned size categories, z=.83, *NS,* were also both equally accurate at making discriminations in both locations in the two experiments. Together these results show that the presence of labels did not change children's accuracy with respect to making discriminations in strongly- and weakly-learned areas of the category space.

Discrimination size step

The results of the previous experiments have shown, in general, that as the size of the difference between stimuli increases, children are more accurate at discriminating them regardless of the presence or absence of a label. Thus, it was thought that there would be no effects of experiment on children's discrimination accuracy with respect to threshold size. In fact, mixed logistic regression models of the interaction between experiment (label versus no label) and threshold size (.5, 1, or 1.5x JND) revealed that both brightness preferrers who learned brightness categories, z=-.91, *NS,* and those who learned size categories, z=-.58, *NS,* were equally accurate at making discriminations in both experiments. The same was true for size preferrers who learned size categories, z=-.41, *NS*. However, size preferrers who learned brightness categories were significantly less accurate in making discriminations in the label experiment, z=-2.04, p<.05, but this interacted with discrimination size such that children in the label experiment showed a larger effect of discrimination size on accuracy, z=2.05, p<.05.

Blocking

The blocking analysis is supposed to assess whether children show a cost for having to switch between making relevant and irrelevant discriminations. Mixed logistic regression models of the interaction between experiment (label versus no label) and block type (single-type or mixed) revealed that brightness preferrers who learned brightness categories showed no effect of experiment on accuracy of making between-category discriminations, z=-.18, *NS* or on accuracy of making within-category discriminations along the relevant dimension, z=-1.33, *NS.* However, there was a marginal interaction between experiment and block in the model of these children's accuracy in making within-category discriminations along the irrelevant dimensions such that children in the label condition showed less of an advantage for making discriminations within the context of a single-type than a mixed block.

Conclusions

In general, the results of these comparison analyses are mixed, and unexpected in many ways. One possible explanation for the mixed and unexpected results is that there are differences in the way children learn about size and brightness, making it difficult to separate differences in the child and their attentional biases from differences inherent in the dimensions themselves. The results of the category learning comparison revealed that both brightness and size preferrers were slower to learn categories in the presence of a label. This finding differs markedly from the results of the holistic classifier comparison, which revealed that labels facilitated category learning in children who show very little evidence of dimensional attention. If we assume that preferrers do in fact represent an intermediate developmental stage between holistic and dimensional attention, then the results of Experiment 3 and 4 could be taken to suggest that labels direct attention to specific dimensions, but that when attention is already biased to a given dimension, the increase in selective attention to that dimension makes it much more difficult to instead attend instead to another dimension (cf. Brojde et al., 2011). This fits with the fact that size preferrers were equally good at learning categories organized by size or brightness without a label, but were slightly slower to learn size categories with the label, could suggest that size is a harder dimension to attend to in category learning. In that case, the label could be directing attention to the more salient or easy-to-attend-to dimension—brightness—and thereby slowing learning of the harder dimension.

Together these findings do lend some support to the idea that the developmental trajectory of selective attention is one from weak selective attention to selective but inflexible—or sticky—attention to selective and flexible attention. Here we see that preferrers are in fact quite sticky and that this appears to interact with inherent differences between size and brightness dimensions, as evidenced by the increased stickiness in the label experiment. To really explore how the differences between dimensions and differences between types of dimensional attention interact, further experiments will have to be conducted on preferrers' ability to categorize dimension categories organized by dimensions other than brightness and size (e.g. orientation, shape, color, etc.) or even novel dimensions. It could be the case that brightness and size are not equal in either salience or learnability, and exploring how children learn about other dimensions will help clarify this issue. Unfortunately, however, the results of the two preferrer experiments presented here, however, cannot disentangle the effects of a given child's dimensional bias from the inherent differences in the dimensions themselves.

CHAPTER VIII: GENERAL DISCUSSION

The role of category learning in dimensional attention

It had already been established that category learning influences adults' perceptual discrimination along category-relevant dimensions (Goldstone, 1994b). However, it was not previously clear whether category learning should influence children's perceptual discrimination abilities in the same way. Young children had previously been shown to have difficulty selectively attending to one dimension of similarity to the exclusion of others, and thus it was unclear whether they should even be able to learn dimensional categories in the same way that adults do. Furthermore, it was unclear whether learning to selectively attend to a category-relevant dimension would scaffold children's subsequent attention to dimensional similarity in a classification task that pit holistic similarity against dimensional similarity.

In Experiment 1, I showed that the influence of category learning on perceptual discrimination rests in part on children's ability to demonstrate attention to dimensional similarity. Children who classified holistically in a triad task were slower to learn dimensional categories. On the other hand, children of the same age who classified dimensionally were able to learn these categories quickly. Interestingly, both classifier groups showed changes in discrimination similar to the pattern demonstrated by adults in earlier studies (i.e. Goldstone, 1994b). In particular, both dimensional and holistic classifiers who learned brightness categories showed enhanced accuracy for between-category discriminations relative to within-category discriminations. However, only the dimensional classifiers demonstrated increased selective attention to category-relevant dimensions during the posttest triad task. Together these results suggest that while it is relatively difficult for holistic classifiers to learn dimensional categories, the process of doing so does influence their selective attention to category-relevant dimensions in the same way as it had dimensional classifiers or adults in previous experiments (e.g.

Goldstone, 1994b). Importantly, however, the influences of category learning do not appear to be strong enough to cause changes to children's subsequent similarity classification.

The role of labels in dimensional attention

If category learning is not enough to scaffold children's selective attention to dimensional similarity, what is? Evidence from young children's novel noun generalization suggests that word learning trains children's selective attention in the service of learning new words (Smith et al., 2002). Evidence from adults suggests that redundant labels facilitate category learning (Lupyan et al., 2007) possibly because they increase selective attention to category relevant dimensions (Brojde et al., 2011). Thus my proposal was that, over developmental time, word learning could train children's selective attention more broadly and thereby be the driving force in the development of selective attention processes. It was therefore predicted that, over the course of an experiment, labels would facilitate category learning—even for children who typically have difficulty learning dimensional categories—and support increased selective attention to dimensional similarity in subsequent classification.

In fact, in Experiment 2, the presence of a label enabled holistic classifiers to learn dimensional categories as quickly as dimensional classifiers did. Again both dimensional and holistic classifiers who learned brightness categories demonstrated increased accuracy for between-category relative to within-category discriminations. However, this time, the holistic classifiers also showed increased selective attention to category-relevant dimensions in the classification task. Furthermore, results of a direct comparison between holistic classifiers' performance in Experiments 1 and 2 revealed that not only were those in Experiment 2 (i.e. who heard labels) faster to learn dimensional categories, but they also showed a greater advantage for making between- versus within-category discriminations, and showed a larger increase in selective attention to dimensional similarity in a subsequent classification task. These second two results suggest that the advantage in categorization that holistic classifiers show in this experiment relative to their counterparts in Experiment 1 was connected to an increase in their changes in selective dimensional attention. Together, these results support my hypothesis that labels scaffold dimensional attention beyond the context of novel noun generalization.

Together, the results of Experiments 1 and 2 demonstrate that labels play a critical role in children's selective attention to dimensions. However, the results also suggest that category learning itself can influence selective attention. The fact that holistic classifiers can learn dimensional categories (even if it is relatively slowly), and that doing so influences their ability to make discriminations along the category-relevant dimension, demonstrates that the development of selective attention is not a discrete change. Rather, these data suggest that the emergence of attention to dimensional similarity is a continuous developmental change of increasingly selective and flexible attention.

Preferrers

Previous research demonstrated that a common, developmentally intermediate, pattern of responding in the triad classification task is to preferentially attend to one dimension of similarity on every trial regardless of whether this means selecting the dimensional or holistic match. At present it is still not understood why children should prefer one particular dimension over another. Nevertheless, in order to understand selective attention across development, it was important for me to examine the roles of category learning and labels on preferrers' selective attention to dimensional similarity.

Together, the results of Experiments 3 and 4 demonstrate that labels actually slow down dimensional category learning in preferrers. This was different from the results of the holistic classifier comparison, which revealed that labels facilitated category learning.

These results could be taken as evidence for the idea that labels direct attention to specific dimensions, but that when attention is already biased to a given dimension, this increase in selective attention to that dimension makes it quite difficult to attend instead to another dimension (cf. Brojde et al., 2011). Additionally, the fact that size preferrers were equally good at learning categories organized by their preferred and nonpreferred dimension without a label, but were slightly slower to learn size categories with the label, could suggest that size is a harder dimension to attend to in category learning. This could mean, among other possibilities, that even size preferrers are biased in some way to selectively attend to brightness—especially in the context of a label.

Differences between size and brightness dimensions

The results of the preferrer experiments add evidence to the idea that there is something inherently different about brightness as size dimensions. As Goldstone's (1994) study of the effects of adults' category learning on subsequent perceptual discrimination showed, participants who learned brightness, but not size categories, showed evidence of enhanced accuracy on discriminations along the relevant dimension relative to the irrelevant dimension. This pattern was generally replicated throughout the current study—children who learned brightness categories showed differences in their discrimination, but children who learned size categories did not. What is different about these dimensions and how might that help us understand what the results of the preferrer experiments?

One possibility is related to the difficulty that children have learning to use size terms, such as big and small, compared to color terms (Sandhofer & Smith, 1999). Part of this difficulty is that size terms need to be used relationally while color terms can be used categorically (Sandhofer & Smith, 2001). In fact, anecdotally it seemed that many children in the current studies who learned brightness categories tended to describe stimuli in terms of pre-existing color categories (often "white" and "brown" which are probably the closest labels some children have to describe values on a grayscale). It also seemed that many children learning size categories seemed to use words like “big” and “small” in an absolute sense rather than a relational sense to describe the stimuli.

Why would talking about size in an absolute sense make it more challenging to learn size categories? Labels such as “big” and “small” are relative terms: a mouse is big in the context of an ant but small in the context of an elephant. Inappropriately activating a familiar label like “small” to remember a specific absolute size before you have had time to get a sense of what the range of sizes within the stimulus set are, could make it hard to learn and remember where the category boundary is. It could help you correctly categorize stimuli at the furthest ends of the dimension, but could make it more difficult to categorize the stimuli in the middle. In addition, research suggests that using labels to describe objects tends to make our memories for specific object properties less precise (e.g. Lupyan, 2008). If children are remembering the smallest stimulus by calling it small, their memory for it may be exaggerated in the direction of it being even smaller than it is in reality. When they subsequently see a stimulus from the middle of the dimension and need to compare it to that small stimulus that they know for sure belongs to a particular category, the middle stimulus may seem more different from the memories for both the small and the big stimuli. Therefore, if children are remembering object-category pairings by saying “the small one goes here” their memory for the category boundary will be fairly imprecise. Further work will need to be done to figure out if the different ways in which children and adults talk about size and brightness relate to their ability to learn categories and make perceptual discriminations along these dimensions.

Another possibility is, even though the stimuli were calibrated to equate step size changes in size to changes in brightness (just as they had been for adults in Goldstone’s work), it could be that brightness is a more salient dimension than size. If so, it is possible that it would require many more size category learning blocks to affectively decrease children’s attention to brightness compared to the number of blocks it takes to decrease attention to size. Similarly, even if size preferrers *can* selectively attend to size well enough to learn size categories, it is may be so easy to attend to brightness that they can flexibly switch to attending to either during discrimination. Further work will need to be done to explore if this is possible or if attention to brightness can never really decrease.

Continuity in the development of dimensional attention

There is growing evidence that the development of selective dimensional attention is a protracted continuous change rather than a discrete shift. Over the preschool and early school years, we see children gradually increasing in their abilities to both selectively and flexibly attend to dimensional similarity (Hanania & Smith). A certain degree of selectivity is needed to correctly sort cards in the pre-switch phase of the DCCS task, and even more is need be able to flexibly shift and selectively attend to the opposite dimension in the post-switch phase. The amount of selective attention needed increases again to focus on a single dimension of similarity in the triad classification task. Finally, there is a further bump in the amount of selective attention is needed to be able to flexibly shift between attended dimensions to exclusively select dimensional matches, or as one child participant referred to it, as "adjusting between size and color."

The data from the current studies support this idea of continuity in the development of selective attention. First, even holistic classifiers in the no label experiment show an influence of category learning on their subsequent discrimination abilities such that they were significantly more accurate at making discriminations along the relevant the irrelevant dimension. This shows how, with a good deal of effort, children can selectively attend to dimensional similarity. However, this selective attention did not cascade into the triad classification task, demonstrating that it was still relatively weak compared to the attentional abilities of the dimensional classifiers, for example.

Second, preferrers appear to be demonstrating "sticky attention" not just in the triad task, but more generally. In particular I found that labels slow down both brightness and size preferrers' category learning—especially for size categories. This might mean that labels are more likely to direct attention to brightness than size, and that this gets preferrers stuck on brightness, which is okay if they are learning brightness categories, but detrimental to their learning of size categories. Sticky attention would suggest a continuous trajectory of selective attention development: children go from a preference to overall similarity, to selectively attending to one dimension in particular, to being able to flexibly switch back and forth to selectively attending to either dimension.

Third, and perhaps most obviously, the very fact that labels could scaffold dimensional attention in holistic classifiers, but not dimensional classifiers supports the continuity idea to some extent. Labels did not fundamentally change what the children were doing, but rather supported their weak attention enough to carry over into the subsequent tasks. Dimensional classifiers were strong enough in their selective attention to not benefit from this support. This is similar to Katz's 1963 finding that 7-year-olds' similarity perception was more affected by labeling patterns than 9-year-olds' similarity perception. Over development, children become increasingly skilled at directing their attention to dimensional similarity and thus become decreasingly influenced by labels—especially when the task at had is relatively easy (as was the category learning task in the current studies). Thus, an important direction for future research will be to further examine this developmental trajectory with respect to the role that labels can play in supporting dimensional similarity.

Category response labels versus novel incidental labels

In the current experiments, the category response itself was associated with a label: ocean or jungle. This is similar to prior work with adults, such as Lupyan and colleagues' aliens to approach or avoid. This means that, by one view, all the participants had a label that was redundant with the category information; and the participants in the "label" experiments had two. Why did the single known labels, not support the same degree of dimensional attention as extra novel labels did? One possibility is that the labels ocean and jungle, for example, are already familiar and loosely associated with many objects and categories. Thus, these labels are already associated with many things and not clearly relevant to the to-be-learned categories in the experiment. This would mean that it is more difficult to associate the category response label with the category-relevant dimension. Novel labels, on the other hand, are not associated with anything prior to the experiment. This could make them more easily associated with the novel stimuli and thus better serve to direct attention to the category-relevant dimension. Future work clearly needs to be done to delineate between these possibilities. One way to examine this issue would be to compare children's category learning with familiar and novel labels that were both redundant to a category response label. Additionally, familiar labels could be either relevant or irrelevant. For example, would children still be faster to learn a brightness distinction if the redundant labels were cat and dog or big and small as if they were light and dark or leebish and gracious?

While work with adults shows that labels are a better redundant cue to category structure than other cues, no such comparison was made in the current experiments. It is possible that, for children, any redundant cue could facilitate category learning and dimensional attention. Future work will need to be done to determine the extent to which it is the label or the redundancy that is helping children.

Other mechanisms of category learning

The primary goal of this dissertation was to examine the mechanisms driving the development of selective attention to dimensional similarity. Previous research with has suggested that category learning and labels play a role in adult's dimensional attention. Thus, these experiments were designed to measure changes to selective attention brought about by category learning with and without labels. For this reason, the results of the four experiments were discussed primarily in terms of the extent to which they demonstrate selective attention to dimensional similarity. However, it is critical to note that there are surely other mechanisms influencing children's category learning. For example, when learning a category distinction, children may associate individual stimuli or specific regions of a category space with the category response. Labels could still facilitate this process by acting as a redundant cue to category membership.

While the categorization and discrimination tests do not rule out this type of associative mechanism, they suggest that it is not the only one at play and that selective attention processes are playing a role in children's category learning. In particular, the analyses that examined differences in children's accuracy categorizing novel versus familiar stimuli demonstrate that this associative mechanism is not the only mechanism at play. If children had been significantly better at categorizing familiar relative to novel stimuli, this could suggest that they had learned something about specific regions of the category space rather than about the dimension of organization. Instead, I found that no group of children showed any difference in categorization accuracy for novel versus familiar stimuli. This suggests that children did not *only* learn about specific stimuli or regions of the space.

This finding does not rule out, however, whether associative mechanisms play an important role in category learning, especially early on in the time course of individual category learning. Further evidence against associative mechanisms being solely responsible for the category learning seen here comes from the results of the discrimination test. If children were only learning about regions of the category space they may have shown enhanced between-category discrimination and within-category discrimination that was equally decreased across both the relevant and irrelevant dimensions. Such a result would suggest that children had learned about categories but not necessarily about dimensions. Instead, I found that in every case where children showed categorical perception like effects in discrimination, they were significantly more accurate at making within-category discriminations along the relevant than the irrelevant dimension. This suggests that they did not just learn to ignore differences between some values (within the categories) and attend to differences at other values (between the categories), but instead learned which dimension was relevant. However, it is worth noting that they were also always more accurate at between-category discriminations than even within-category discriminations along the relevant dimension. This could potentially mean that both associative learning mechanisms and selective attention mechanisms influenced category learning. Additionally, it could mean that a simple associative mechanism is what builds up selective attention over the course of learning an individual category. Future research will need to be done to explore this issue further.

Future Directions

The effects of labels over multiple timescales

My proposal was that word learning trains children's selective attention to dimensional similarity, not just in the service of learning new words, but also more generally. The current studies support this proposal in that labels had supportive cascading effects on holistic classifier's attention in categorization, discrimination, and similarity classification. However, it is important to note that these effects were all found during the course of an experimental session and I did not measure whether or not children learned anything about the labels themselves. While these studies clearly support a role for labels and word learning in the development of dimensional attention, it remains unclear how exactly labels do this over a longer timescale.

One way to think about how word learning scaffolds dimensional attention over developmental time is to consider what we already know about the mechanisms driving selective attention in the context of word learning (see Smith et al., 2002; Samuelson, 2002; Perry et al., 2010). We know that the regularities present in the early linguistic environment drive children's selective attention to similarity in shape. Most words children learn early name solid objects in categories well-organized by similarity in shape (e.g. bucket). Children first associate individual labels with individual objects. A blue plastic bucket with a handle is called "bucket," an orange cloth bucket with a jack-o'-lantern face is called "bucket," and so on. These associations between the label and each individual object lead to an association between the label "bucket" and the category-relevant dimension: the abstract shape of buckets. As children learn more associations between labels and categories—many of which are also organized by similarity in shape—they develop a higher-order association between the naming context and past instances of selectively attending to shape. This association leads to a bias to selectively attend to shape in future instance of word learning. As children develop, they learn many new words: including nominal categories organized by similarity dimensions other than shape (e.g. pudding is a category organized by similarity in material) and dimensional adjectives and value labels (e.g. blue is a category organized by similarity in color). Children acquire the ability to flexibly shift their attention to different dimensions based on context: to material in the context of nonsolid substances and to color in the context of adjectival frames (Jones & Smith, 1993). The final step that connects this flexible attention to dimensions in different word learning contexts to flexible attention to dimensions more generally is still unclear. The body of work on word learning biases shows that the specific words children *already know* affect their ability to flexibly and selectively attend to dimensional similarity in novel noun generalization. One important direction for future research, therefore, will be to investigate how the specific words children already know affect their ability to selectively attend to dimensional similarity in non-linguistic contexts (see for example, Thom & Sandhofer, 2009; Sandhofer & Doumas, 2008; Sandhofer & Smith, 2001; Sandhofer & Smith, 1999).

What do labels really do?

Are all categories created equal?

The present studies showed that, overall, labels facilitated category learning and are associated with increases in selective attention. However, the directionality of these relationships is unclear. Do labels provide a cue to category membership thus enhancing category representations? Or, do labels support selective attention, enhancing attention to task-relevant dimensions and abstraction over others? If the latter is true, then not all category learning should benefit from the presence of labels (cf. Brojde et al., 2011). In particular, learning holistic categories, which require representing overall similarity rather than selecting specific dimensions, requires less selective attention (Lupyan et al., 2012; Ashby & Maddox, 2011; Sloutsky, 2010). If labels support selective attention, then they should only facilitate dimensional category learning (as in the present studies) but should not facilitate holistic category learning. If, however, labels provide an extra cue to category membership, making category representations stronger, then category structure should not matter—there will be facilitative effects for both holistic and dimensional categories.

How do labels become special?

It also seems possible that the role that labels play in category learning may change over time. Evidence from research on infants' categorization demonstrates that early on, labels are simply features that can be associated with objects, no different from any other associated feature (Sloutsky, 2009). The word "lime" is associated with a lime in the same way the color green is associated with a lime (see e.g. Gliozzi, Mayor, Hu, & Plunkett, 2009; Deng & Sloutsky, 2012; Sloutsky & Fisher, 2004). However, we know very clearly from research with adults, that eventually labels take on a different role and are no longer equivalent to other cues or features associated with objects. For example, work with adults demonstrates that while redundant labels facilitate category learning, other redundant cues, such as space do not (Lupyan et al., 2007). Similarly, labels facilitate adults' familiar object recognition while other cues highly associated with an object (such as a mooing sound associated with a cow) do not (Lupyan & Thompson-Schill, 2012). Therefore, while there is no *a priori* reason why labels should be "special" or different than any other source of information, they might *become* special over development as children learn to associate them with past instances of selective attention and they become more heavily weighted than other features (cf. Lupyan, 2012a).

An extremely important direction for future research, therefore, will be to examine the developmental changes that lead to this increasing "specialness" of labels. One way to do this would be to come back to this question of whether labels support the learning of both holistic and dimensional category structures. It is my hypothesis that labels drive selective attention to dimensional similarity. Labels should facilitate adults' *dimensional* category learning, but not their *holistic* category learning. However, it is likely that this would not be the case much earlier in development. Earlier in development, when labels are no different than any other feature, they should facilitate either type of category learning. Therefore examining both children and adults' holistic and dimensional category learning with and without labels will be critical to understanding when and how labels become special and label to direct attention to dimensional similarity.

Neural implications

A recent study demonstrates that examining neural connections between cognitive control and categorization is another useful way to examine the question of what labels do and how they might direct attention to dimensional similarity. This study asked how changing cortical activity over the left inferior frontal gyrus (LIFG) via transcranial direct current stimulation (tDCS, noninvasive electrical stimulation used to temporarily affect cortical activity) affected categorization (Lupyan et al., 2012). LIFG is associated both with cognitive control processes, such as those used in go/no-go tasks (see Novick et al., 2005) and language processes, such as speech production (see Gernsbacher & Kaschak, 2003). Changes in cortical activity induced by tDCS depend on polarity of stimulation: anodal stimulation increases cortical excitability and cathodal decreases it (Nitsche & Paulus, 2000). Thus, anodal stimulation to LIFG should increase cognitive control in categorization, and cathodal stimulation should decrease it. In fact, Lupyan and colleagues found that cathodal stimulation led to decreased accuracy for dimensional categories (e.g. "things that are green") relative to holistic categories (e.g. "things in a kitchen"), while increasing activity increased the likelihood of selecting items that were weakly associated with the category (Lupyan et al., 2012). This study shows how changing activity in LIFG affects categorization of familiar objects. To fully understand the connection between language and selective attention to dimensional similarity, future work will need to examine how LIFG activity affects cognitive control in *novel* category learning and assess the directionality of the neural relationship between these processes.

Words or a developmental history of word learning?

Does language affect cognitive control because of an immediately present label, or prior experience selecting informative dimensions? A long-term history of using words to refer to categories that require representing relevant dimensions and abstracting over irrelevant dimension should lead to a cognitive control system that can be activated even without immediate linguistic support. According to the Label Feedback Hypothesis, this is why learned labels have a top-down effect on cognition and perception (see Lupyan, 2012b). For example, learning the words "blue" and "green" means that when you see a green object, the word "green" is automatically activated and has a top-down effect that leads to a more categorical representation of green-ness. However, the developmental argument that suggests word-learning is what drives the flexible, selective attention we see in adult cognition suggests that top-down effects of labels do not just come from specific word-referent mappings but more general experience relying on words to support selective attention. A lexicon of labels associated with past instances of selectively attending to category-relevant dimensions leads to a higher-order association between labels and selective attention. This leads *novel* words to support selective attention to dimensions (as in this dissertation) and a cognitive control system that supports selective attention *without labels* (Smith, 1993). One way to examine this experimentally would be to alter cortical activity in the LIFG (via tDCS). If this affects cognitive control in category learning, even without a label present, this would suggest previously-demonstrated effects of labels on cognitive control stem from a longer history of using labels to support cognitive control over development.

Conclusions

Overall, the results of this dissertation support the idea that labels can not only facilitate children's dimensional category learning, but in doing so, they lead to cascading increases in selective attention to dimensional similarity in subsequent tasks. That labels could drive increases in selective attention online during the experiment provides support that labels and word learning could also be the driving force in increasing selective attention over development. Together, the results of these experiments help to clarify the processes involved in development of similarity perception and unify our understanding of attentional processes in word learning with those in a broader context.

REFERENCES

Ashby, F. G., & Maddox, W. T. (2011). Human category learnng 2.0. *Annals of the New York Academy of Sciences, 1224*, 147.

Baayen, R. H. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language, 59*(4), 390.

Brojde, C. L., Porter, C., & Colunga, E. (2011). "Words can slow down category learning": Erratum. *Psychonomic Bulletin & Review, 18*(5), 1033.

Cohen, R., Kelter, S., & Woll, G. (1980). Analytical competence and language impairment in aphasia. *Brain and Language, 10*(2), 331-347.

Cohen, R., Woll, G., & Ehrenstein, W. H. (1981). Recognition deficits resulting from focussed attention in aphasia. *Psychological Research, 43*(4), 391-405.

Cook, G. L., & Odom, R. D. (1992). Perception of multidimensional stimuli: A differential-sensitivity account of cognitive processing and development. *Journal of Experimental Child Psychology, 54*(2), 213-249.

Deng, W. S. ., & Sloutsky, V. M. (2012). Carrot eaters or moving heads: Inductive inference is better supported by salient features than by category labels. *Psychological Science, 23*(2), 178-186.

Edelman, S. (1999). *Representation and recognition in vision*

Frye, D., Zelazo, P. D., & Palfai, T. (1995). Theory of mind and rule-based reasoning. *Cognitive Development, 10*(4), 483-527.

Gastgeb, H. Z., Strauss, M. S., & Minshew, N. J. (2006). Do individuals with autism process categories differently? the effect of typicality and development. *Child Development, 77*(6), 1717-1729.

Gernsbacher, M. A., & Kaschak, M. P. (2003). Neuroimaging studies of language production and comprehension. *Annual Review of Psychology, 54*, 91-114.

Gliozzi, V., Mayor, J., Hu, J., & Plunkett, K. (2009). Labels as features (not names) for infant categorization: A neurocomputational approach. *Cognitive Science, 33*(4), 709-738.

Goldstone, R. (1994). The role of similarity in categorization: Providing a groundwork. *Cognition, 52*(2), 125.

Goldstone, R. L. (1994). Influences of categorization on perceptual discrimination. *Journal of Experimental Psychology: General, 123*(2), 178-200.

Greenberg, S. N., & MacGregor-Hannah, M. (2010). Disruptive effect of holistic bias on processing of other-race faces following face categorization. *Perceptual and Motor Skills, 110*(2), 567-579.

Hanania, R., & Smith, L. B. (2010). Selective attention and attention switching: Towards a unified developmental approach. *Developmental Science, 13*(4), 622-635.

Hanley, J. R., & Roberson, D. (2011). Categorical perception effects reflect differences in typicality on within-category trials. *Psychonomic Bulletin & Review, 18*(2), 355-363.

Jaeger, T. F. (2009, May 14). Random effects: Should I stay or should I go? [web log message]. Message posted to http://hlplab.wordpress.com/2009/05/14/random-effect-structure/

Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language, 59*(4), 434-446.

James, W. (1950). *The principles of psychology, vol. 1*. New York, NY: Dover Publications, Inc. (Original work published 1890).

Katz, P. A. (1963). Effects of labels on children's perception and discrimination learning. *Journal of Experimental Psychology, 66*(5), 423-428.

Kendler, T. S., Kendler, H. H., & Learnard, B. (1962). Mediated responses to size and brightness as a function of age. *The American Journal of Psychology, 75*(4), 571-586.

Kendler, T. S., Kendler, H. H., & Wells, D. (1960). Reversal and nonreversal shifts in nursery school children. *Journal of Comparative and Physiological Psychology, 53*(1), 83-88.

Kirkham, N. Z., Cruess, L., & Diamond, A. (2003). Helping children apply their knowledge to their behavior on a dimension-switching task. *Developmental Science, 6*(5), 449-476.

Kloos, H., & Sloutsky, V. M. (2008). What's behind different kinds of kinds: Effects of statistical density on learning and representation of categories. *Journal of Experimental Psychology: General, 137*(1), 52-72.

Landau, B., & Shipley, E. (2001). Labelling patterns and object naming. *Developmental Science, 4*(1), 109-118.

Landau, B., Smith, L. B., & Jones, S. S. (1988). The importance of shape in early lexical learning. *Cognitive Development, 3*(3), 299-321.

Lupyan, G. (2012a). "What do words do? towards a theory of language-augmented thought." in *The psychology of learning and motivation, vol. 57*, ed. B.H. Ross.

Lupyan, G. (2012b). Linguistically modulated perception and cognitions: The label-feedback hypothesis. *Frontiers in Psychology, 3*:54. *doi: 10.3389/fpsyg.2012.00054*

Lupyan, G., & Mirman, D. (Under review). Linking language and categorization: Evidence from aphasia.

Lupyan, G., Mirman, D., Hamilton, R., & Thompson-Schill, S. L. (2012). Categorization is modulated by transcranial direct current stimulation over left prefrontal cortex.

Lupyan, G. (2008a). The conceptual grouping effect: Categories matter (and named categories matter more). *Cognition, 108*(2), 566-577.

Lupyan, G. (2008b). From chair to "chair": A representational shift account of object labeling effects on memory. *Journal of Experimental Psychology: General, 137*(2), 348-369.

Lupyan, G. (2009). Extracommunicative functions of language: Verbal interference causes selective categorization impairments. *Psychonomic Bulletin & Review, 16*(4), 711-718.

Lupyan, G., Rakison, D. H., & McClelland, J. L. (2007). Language is not just for talking: Redundant labels facilitate learning of novel categories. *Psychological Science, 18*(12), 1077-1083.

Lupyan, G., & Thompson-Schill, S. (2012). The evocative power of words: Activation of concepts by verbal and nonverbal means. *Journal of Experimental Psychology: General, 141*(1), 170-186.

Medin, D. L., Goldstone, R. L., & Gentner, D. (1993). Respects for similarity. *Psychological Review, 100*(2), 254-278.

Minda, J. P., Desroches, A. S., & Church, B. A. (2008). Learning rule-described and non-rule-described categories: A comparison of children and adults. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34*(6), 1518-1533.

Novick, J. M., Trueswell, J. C., & Thompson-Schill, S. (2005). Cognitive control and parsing: Reexamining the role of broca's area in sentence comprehension. *Cognitive, Affective & Behavioral Neuroscience, 5*(3), 263-281.

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

Pisoni, D. B., & Tash, J. (1974). Reaction times to comparisons within and across phonetic categories. *Perception & Psychophysics, 15*(2), 285-290.

Samuelson, L. K., Schutte, A. R., & Horst, J. S. (2009). The dynamic nature of knowledge: Insights from a dynamic field model of children's novel noun generalizations. *Cognition, 110*(322), 345.

Samuelson, L. K. (2002). Statistical regularities in vocabulary guide language acquisition in connectionist models and 15-20-month-olds. *Developmental Psychology, 38*(6), 1016-1037.

Samuelson, L. K., & Smith, L. B. (1999). Early noun vocabularies: Do ontology, category organization and syntax correspond? *Cognition, 73*(1), 1-33.

Sandhofer, C. M., & Doumas, L. A. A. (2008). Order of presentation effects in learning color categories. *Journal of Cognition and Development, 9*(2), 194-221.

Sandhofer, C. M., & Smith, L. B. (1999). Learning color words involves learning a system of mappings. *Developmental Psychology, 35*(3), 668-679.

Sandhofer, C. M., & Smith, L. B. (2001). Why children learn color and size words so differently: Evidence from adults' learning of artificial terms. *Journal of Experimental Psychology: General, 130*(4), 600-617.

Sandhofer, C., & Smith, L. B. (2007). Learning adjectives in the real world: How learning nouns impedes learning adjectives. *Language Learning and Development, 3*(3), 233-267.

Simmering, V. R., & Spencer, J. P. (2008). Generality with specificity: The dynamic field theory generalizes across tasks and time scales. *Developmental Science, 11*(4), 541-555.

Simon, H. A. (1978). Information-processing theory of human problem solving. In W. K. Estes (Ed.), *Handbook of learning & cognitive processes* (pp. 271-295). Oxford, England: Lawrence Erlbaum.

Sloutsky, V. M. (2009). Theories about "theories": Where is the explanation? comment on waxman and gelman. *TRENDS in Cognitive Science, 13*, 331-332.

Sloutsky, V. M. (2010). From perceptual categories to concepts: What develops? *Cognitive Science: A Multidisciplinary Journal, 34*(7), 1244-1286.

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

Smith, L. B. (1983). Development of classification: The use of similarity and dimensional relations. *Journal of Experimental Child Psychology, 36*(1), 150-178.

Smith, L. B. (1993). The concept of same. *Advances in child development and behavior, vol. 24.* (pp. 215-252) San Diego, CA, US: Academic Press.

Smith, L. B., Gasser, M., & Sandhofer, C. (1997). Learning to talk about the properties of objects: A network model of the development of dimensions. In R. L. Gloldstone, D. L. Medin & P. G. Schyns (Eds.), *Perceptual learning* (pp. ???). San Diego: Academic Press.

Smith, L. B., Jones, S. S., Landau, B., Gershkoff-Stowe, L., & Samuelson, L. K. (2002). Object name learning provides on-the-job training for attention. *Psychological Science, 13*(1), 13-19.

Smith, L. B., & Kemler, D. G. (1977). Developmental trends in free classification: Evidence for a new conceptualization of perceptual development. *Journal of Experimental Child Psychology, 24*, 279-298.

Smith, L. B., & Kemler, D. G. (1978). Levels of experienced dimensionality in children and adults. *Cognitive Psychology, 10*(4), 502-532.

Thom, E. E., & Sandhofer, C. M. (2009). More is more: The relationship between vocabulary size and word extension. *Journal of Experimental Child Psychology, 104*(4), 466-473.

Vignolo, L. A. (1999). Disorders of conceptual thinking in aphasia. In G. Denes L. Pizzamiglio (Ed.), *Handbook of clinical and experimental neuropsychology* (pp. 273-288). Hove, England: Psychology Press/Erlbaum (UK) Taylor & Francis.

Yoshida, H., & Smith, L. B. (2005). Linguistic cues enhance the learning of perceptual cues. *Psychological Science, 16*(2), 90.

APPENDIX

TABLES AND FIGURES

Table A1. Participant information for each classifier type and category learning group in each experiment including: mean age and range, number of participants, number of females, mean socioeconomic status (as measured by maternal education where 1 is 8th grade degree and 6 is doctoral degree) and range of SES.

Participant Information

	Holistic Classifiers	Dimensional Classifiers	Brightness Preferrers	Size Preferrers
E1 & E3 (No Label): Brightness learners	*M:* 5:11 range:5:0-8:3 N=8, 6F SES: 4.9 (4-6)	*M:* 6:5 range: 5:0-8:1 N=8, 6F SES: 5.5 (5-7)	*M:* 7:9 range: 6:3-8:8 N=8, 4F SES: 5 (3-6)	*M:* 6:11 range: 5:3-8:6 N=8, 4F SES: 4.6 (3-6)
E1 & E3 (No Label): Size learners	*M:* 6:0 range: 5:0-8:2 N=8, 6F SES: 4.75 (3-5)	*M:* 6:10 range: 5:0-8:10 N=8, 4F SES: 4.4 (3-6)	*M:* 6:11 range: 5:3-8:4 N=8, 5F SES: 4.6 (3-6)	*M:* 6:7 range: 5:3-8:0 N=8, 2F SES: 4.5 (3-6)
E2 & E4 (Label): Brightness learners	*M:* 6:6 range:5:1-8:4 N=8, 2F SES: 4.2 (2-5)	*M:* 7:0 range:5:6-8:11 N=8, 5F SES: 5.6 (5-7)	*M:* 7:10 range:6:7-8:11 N=6, 4F SES: 5.6 (4-7)	*M:* 6:10 range: 5:1-8:7 N=8, 5F SES: 5 (4-6)
E2 & E4 (Label): Size learners	*M:* 6:6 range:5:0-7:11 N=8, 4F SES: 3.8 (3-6)	*M:* 7:7 range:6:10-8:9 N=8, 5F SES: 5.5 (5-7)	*M:* 7:0 range: 6:1-8:0 N=5, 4F SES: 4.6 (2-6)	*M:* 7:0 range: 6:2-8:1 N=8, 2F SES: 4.4 (3-6)

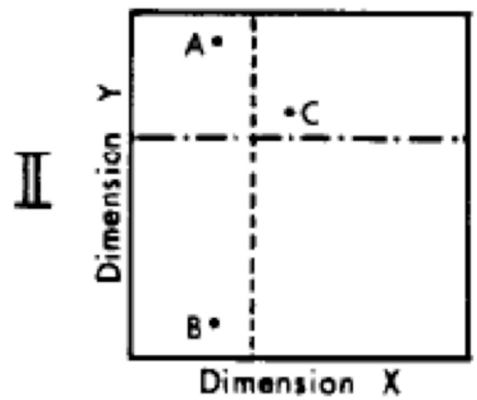


Figure A1. Schematic representation of stimuli used in Smith and Kemler's 1977 triad classification task.

Figure A2. Stimuli used in Experiments 1-4.

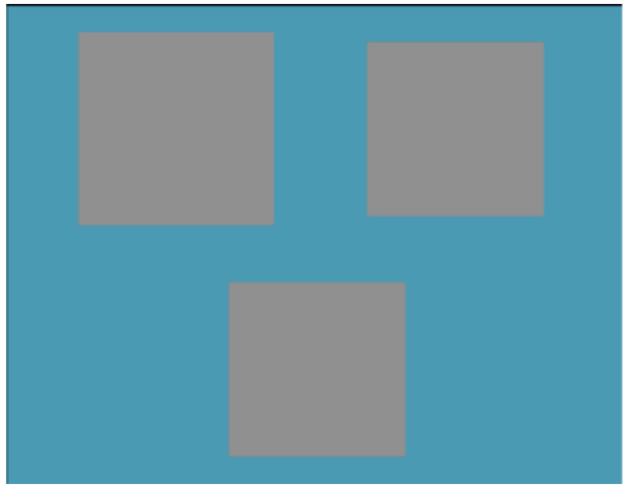


Figure A3. Discrimination task. Represents presentation of all three stimuli: target is at bottom, foil is top left, match is top right. Children had to touch the square represented on top that they believe was the same as the bottom one.

Figure A4. Example of stimuli used in triad task

Figure A5. Representation of a categorization trial.

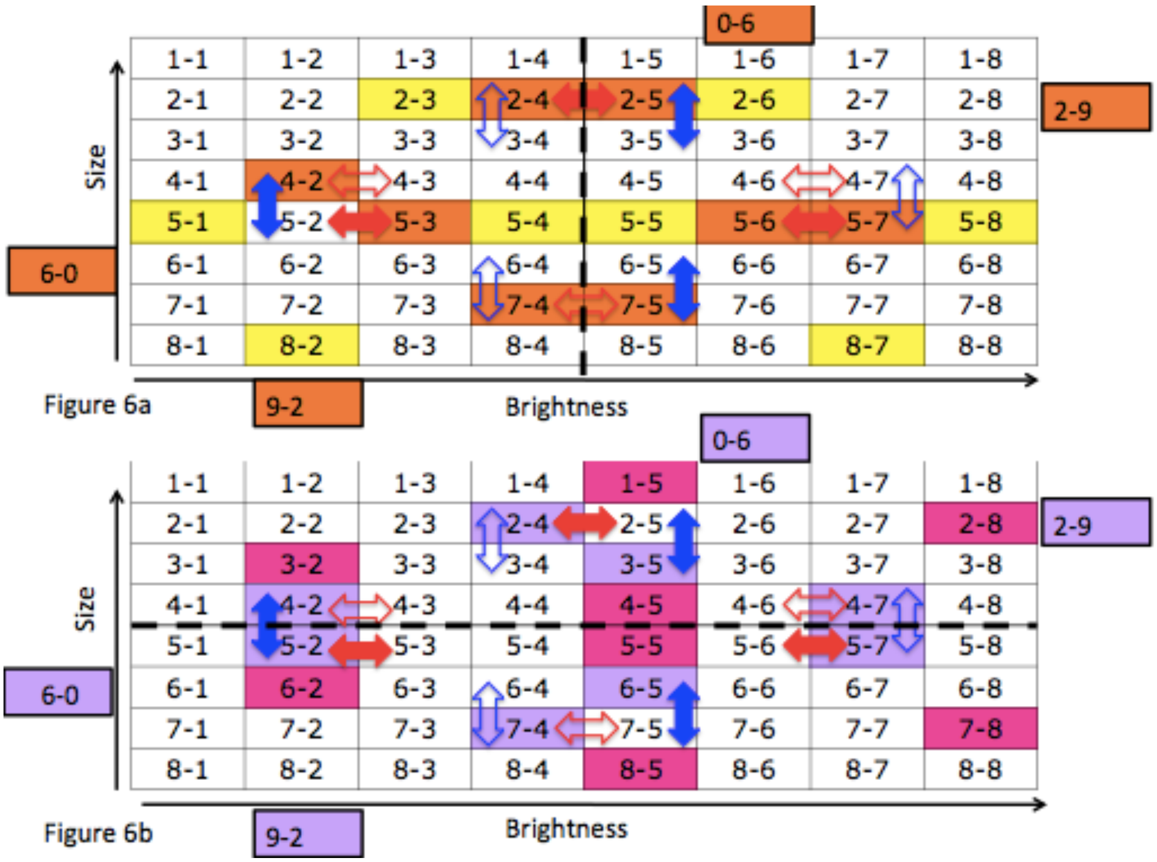


Figure A6. Stimulus space used for stimuli in Experiments 1-4. 6a shows structure of categories for brightness learners and 6b shows the structure of categories for size learners. Yellow and pink cells represent exemplars presented in category learning for each group. Orange and purple cells represent novel exemplars presented in category test for each group. Blue arrows represent locations at which size discriminations were made in the discrimination task and red arrows represent locations at which brightness discrimination were made. Filled arrows represent strongly-learned areas of the category space (for the learning group for whom the discrimination dimension is relevant). Open arrows represent weakly-learned areas of the category space (again for the learning group for whom the discrimination dimension is relevant).

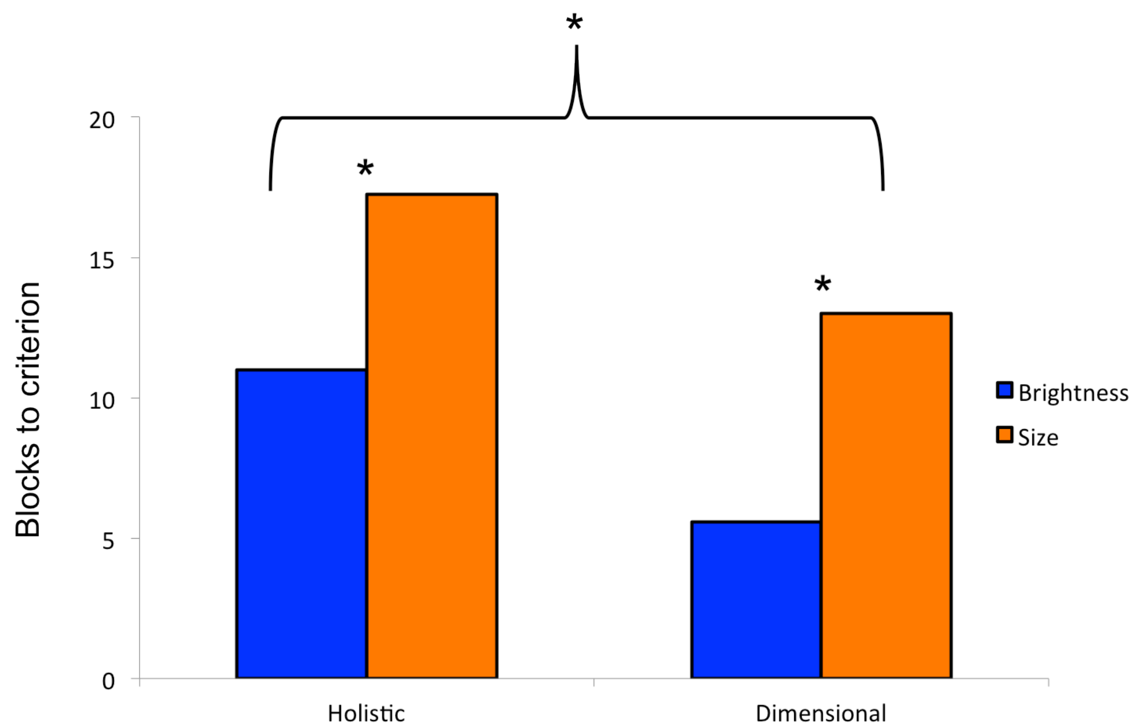


Figure A7. Average number of blocks it took children in each group to reach criterion (e.g. accurate categorization on 7 of 8 trials for 2 blocks in a row) in the category learning task of Experiment 1.

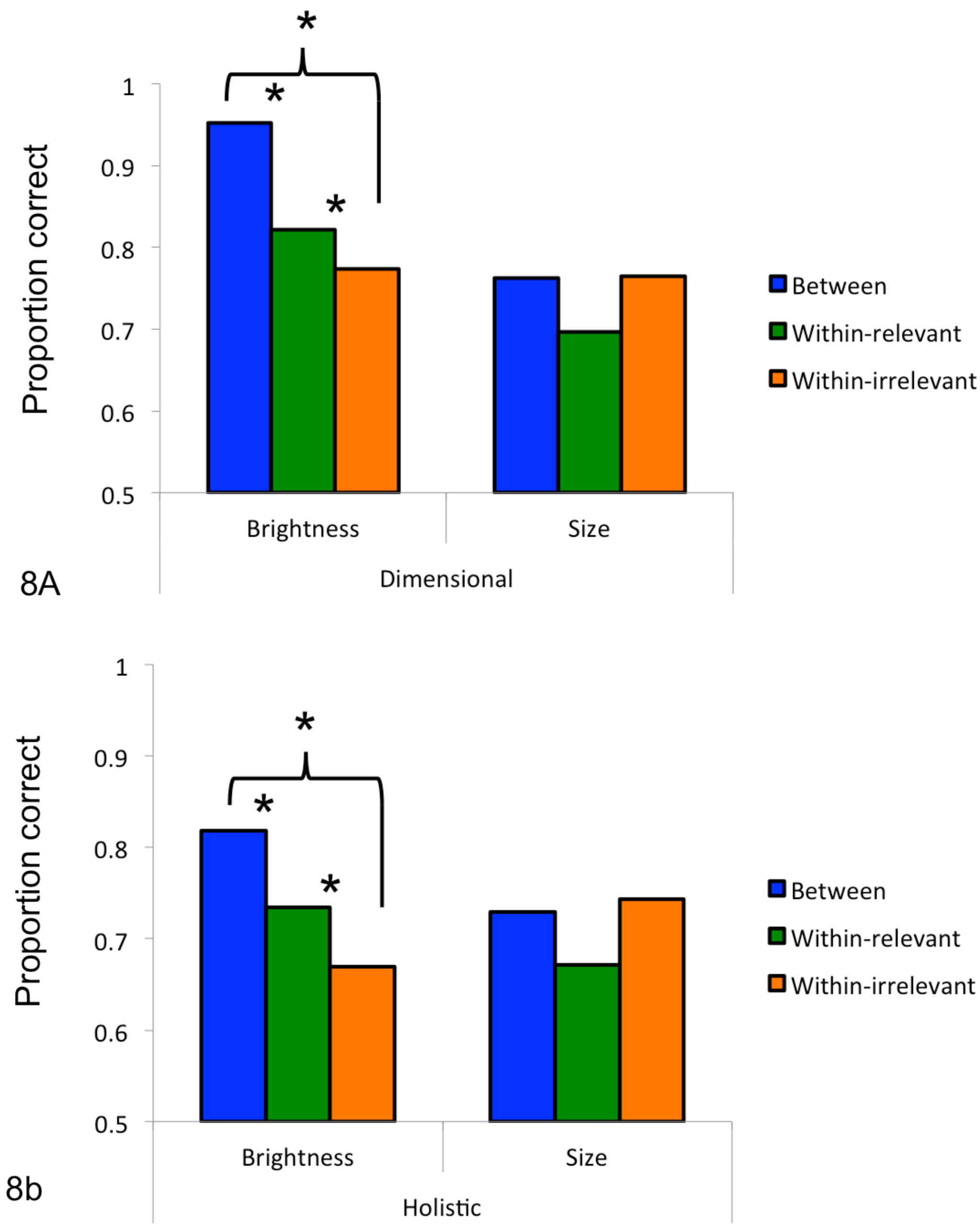


Figure A8. Accuracy for discriminations between categories, within categories along the relevant dimension, and within categories along the irrelevant dimension in Experiment 1. Panel A shows the results of children who were dimensional classifiers and Panel B shows the results of children who were holistic classifiers.

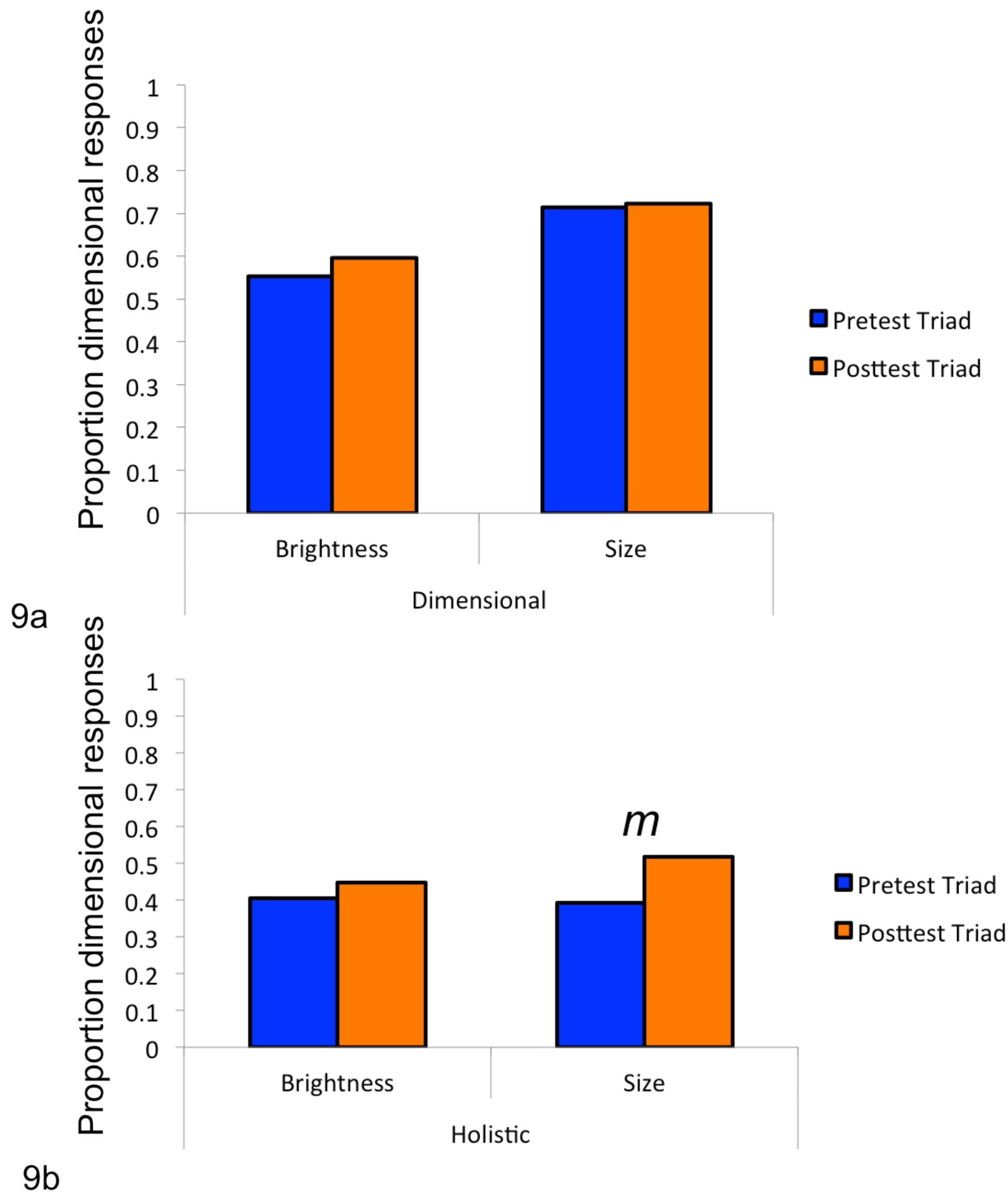


Figure A9. Proportion of dimensional responses that children made during pre- and posttest triad tasks in Experiment 1. Panel A shows the results of children who were dimensional classifiers and Panel B shows the results of children who were holistic classifiers.

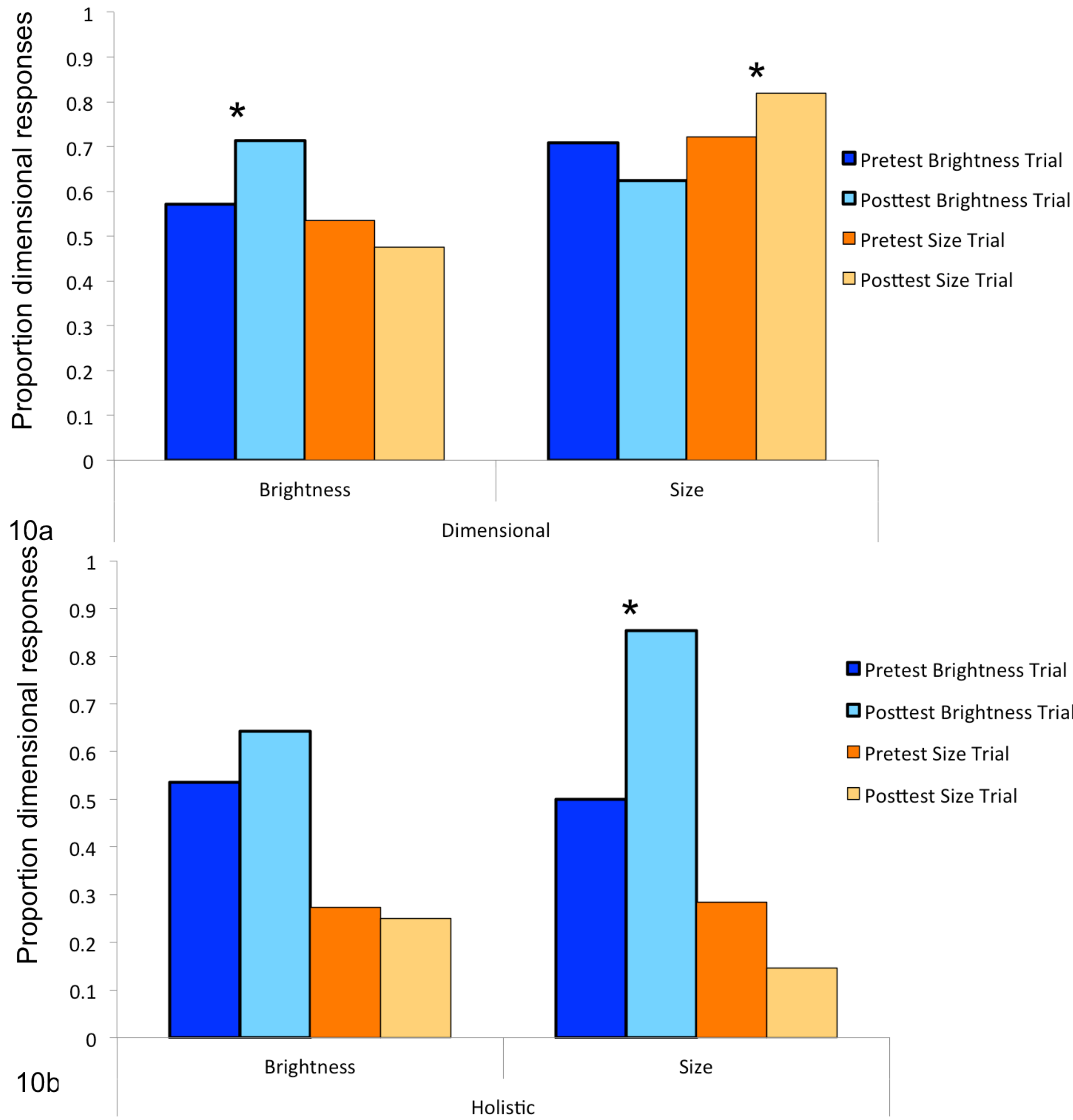


Figure A10. Proportion of dimensional responses that children made during pre- and posttest triad tasks for each trial type (whether the dimensional match was a match along the brightness dimension or the size dimensions) in Experiment 1. Panel A shows the results for children who were dimensional classifiers and Panel B shows the results for children who were holistic classifiers.

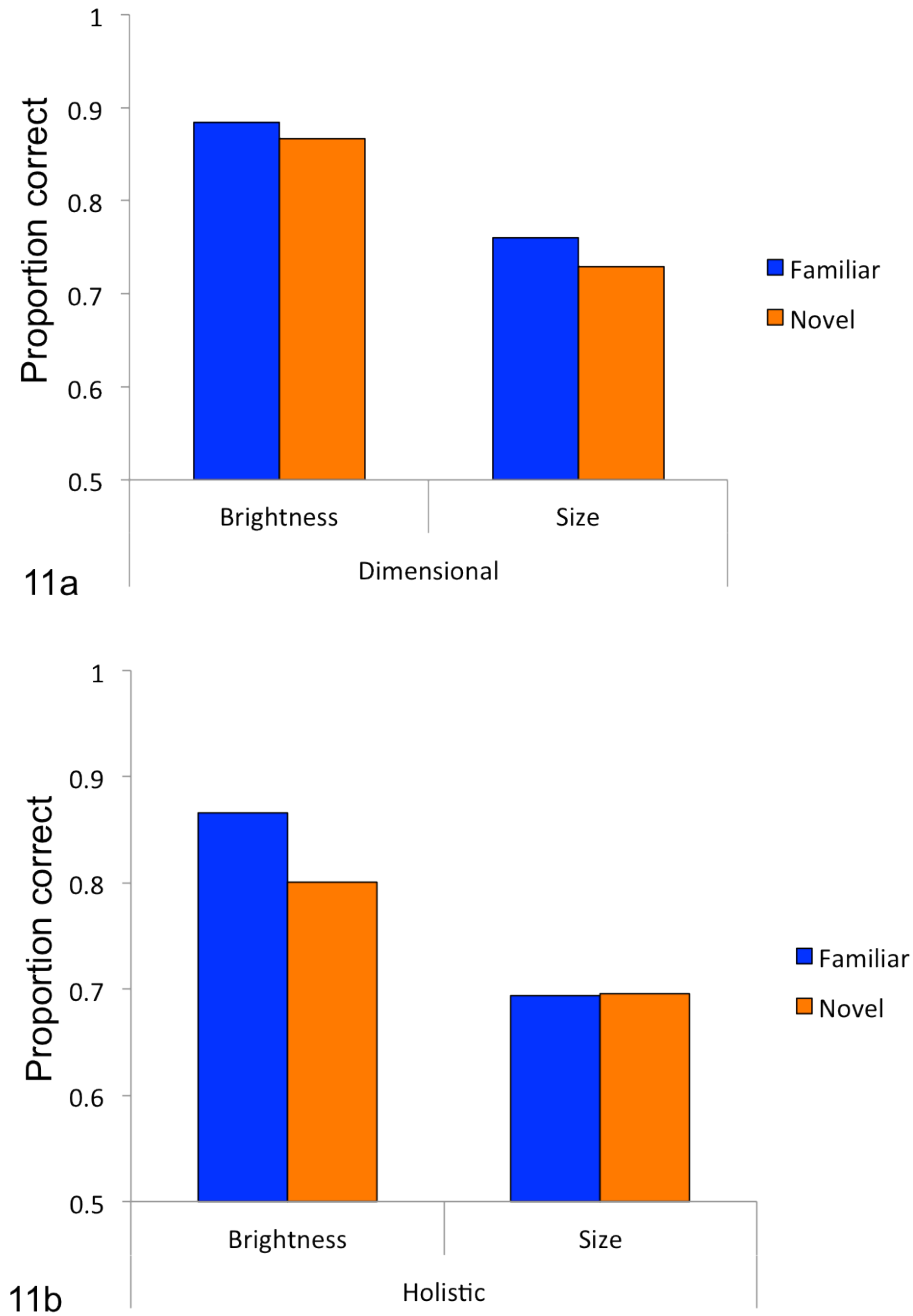


Figure A11. Average accuracy during category test on both familiar and novel trials for brightness learners (left bars) and size learners (right bars) in Experiment 1. Panel A shows the accuracy of children who were dimensional classifiers and Panel B shows the accuracy of children who were holistic classifiers.

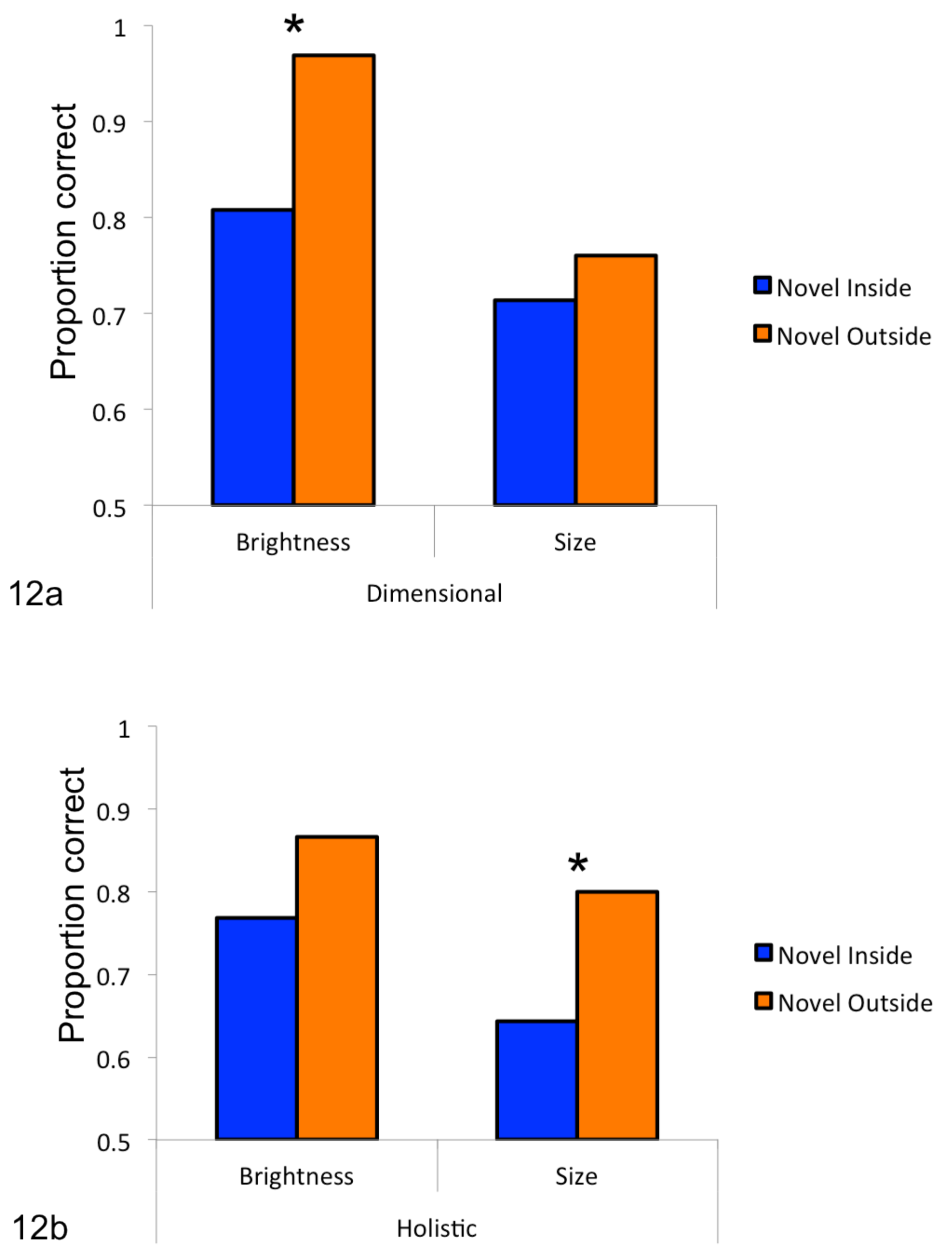
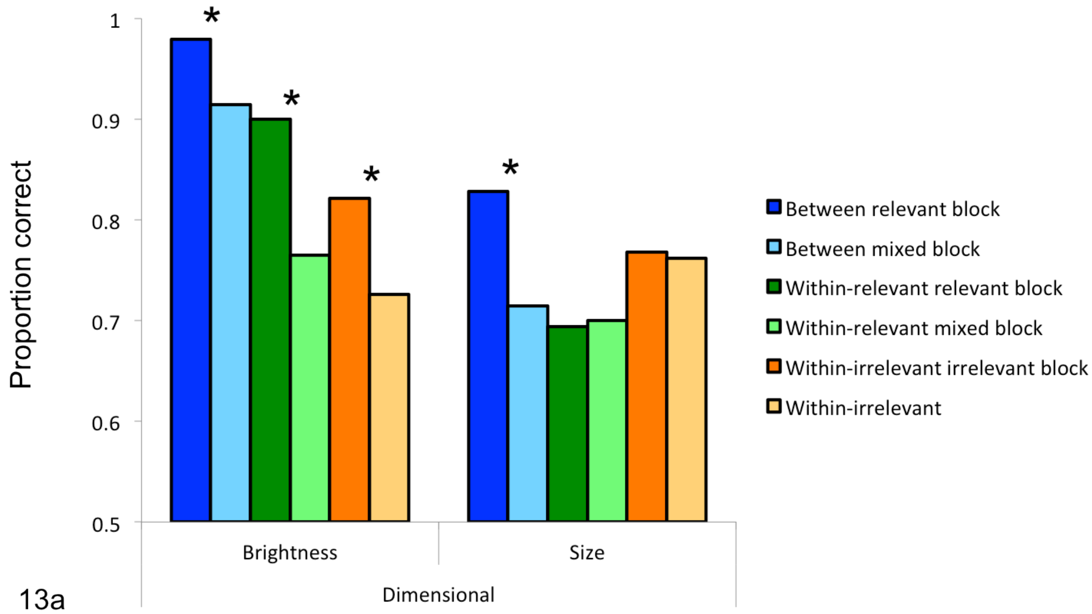


Figure A12. Average accuracy during category test for novel stimuli both inside and outside the category space in Experiment 1. Panel A shows the results for children who were dimensional classifiers and Panel B shows the results for children who were holistic classifiers.



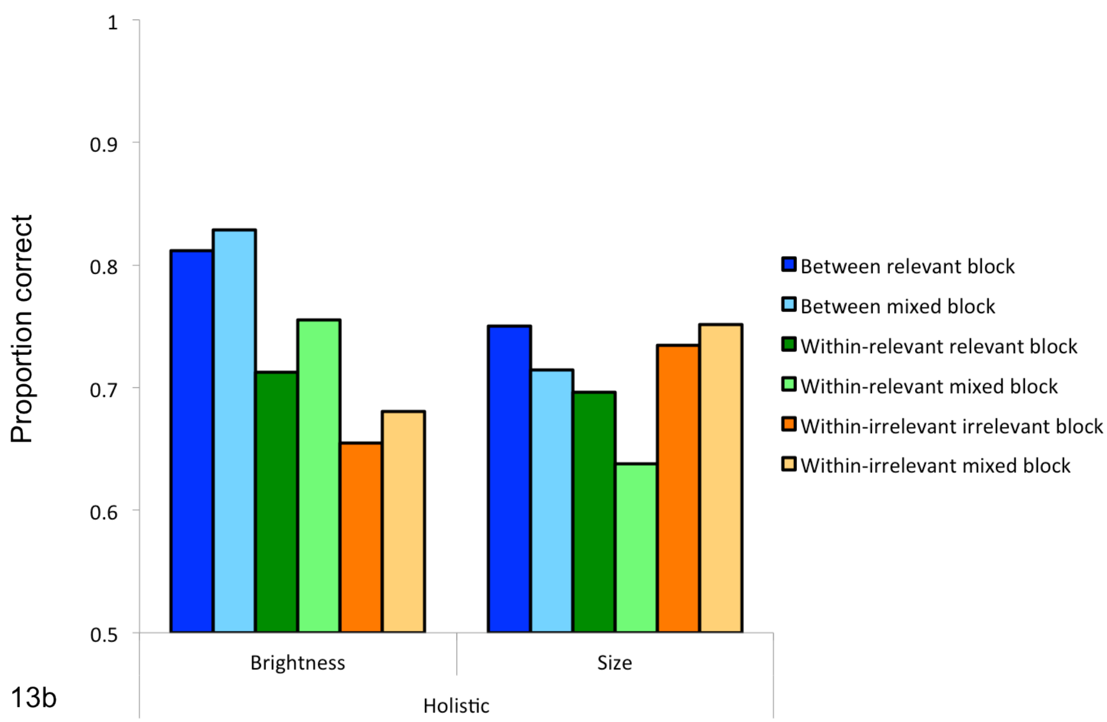


Figure A13. Accuracy for each type of discrimination depending on block structure (all relevant-dimension, all irrelevant-dimension, or mixed-dimension discriminations) in Experiment 1. Panel A shows results from children who were dimensional classifiers. Panel B shows results from children who were holistic classifiers.

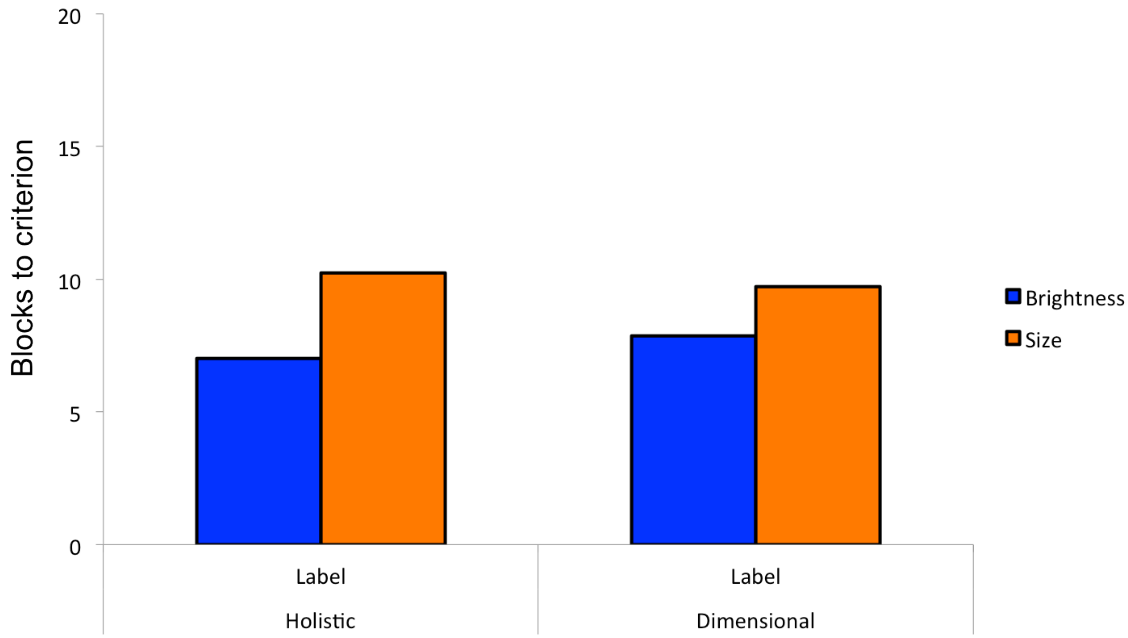


Figure A14. Average number of blocks it took children in each group in Experiment 2 to reach criterion (e.g. accurate categorization on 7 of 8 trials for 2 blocks in a row) in the category learning task.

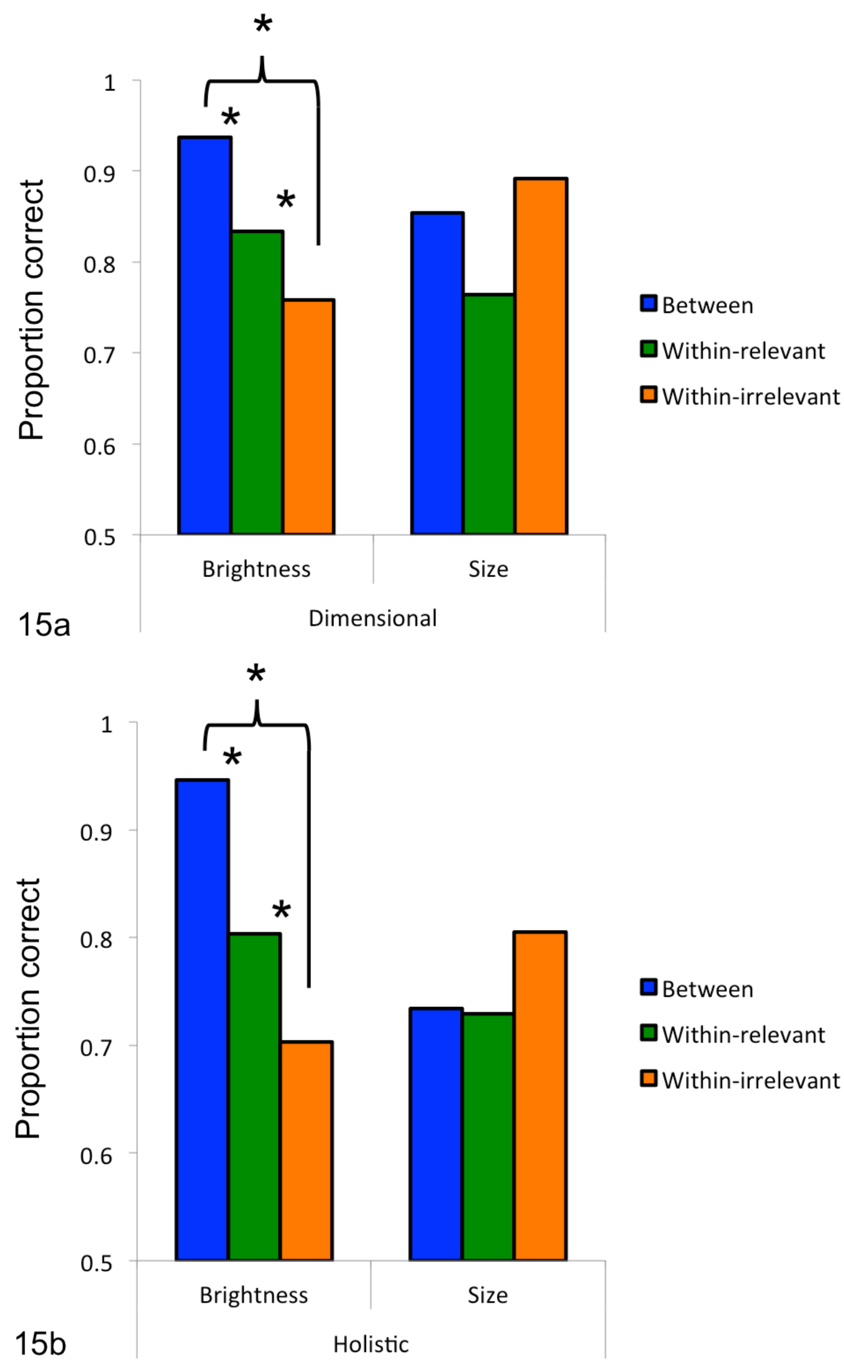


Figure A15. Accuracy for discriminations between categories, within categories along the relevant dimension, and within categories along the irrelevant dimension in Experiment 2. Panel A shows the results of children who learned were dimensional classifiers and Panel B shows the results of children who were holistic classifiers.

.

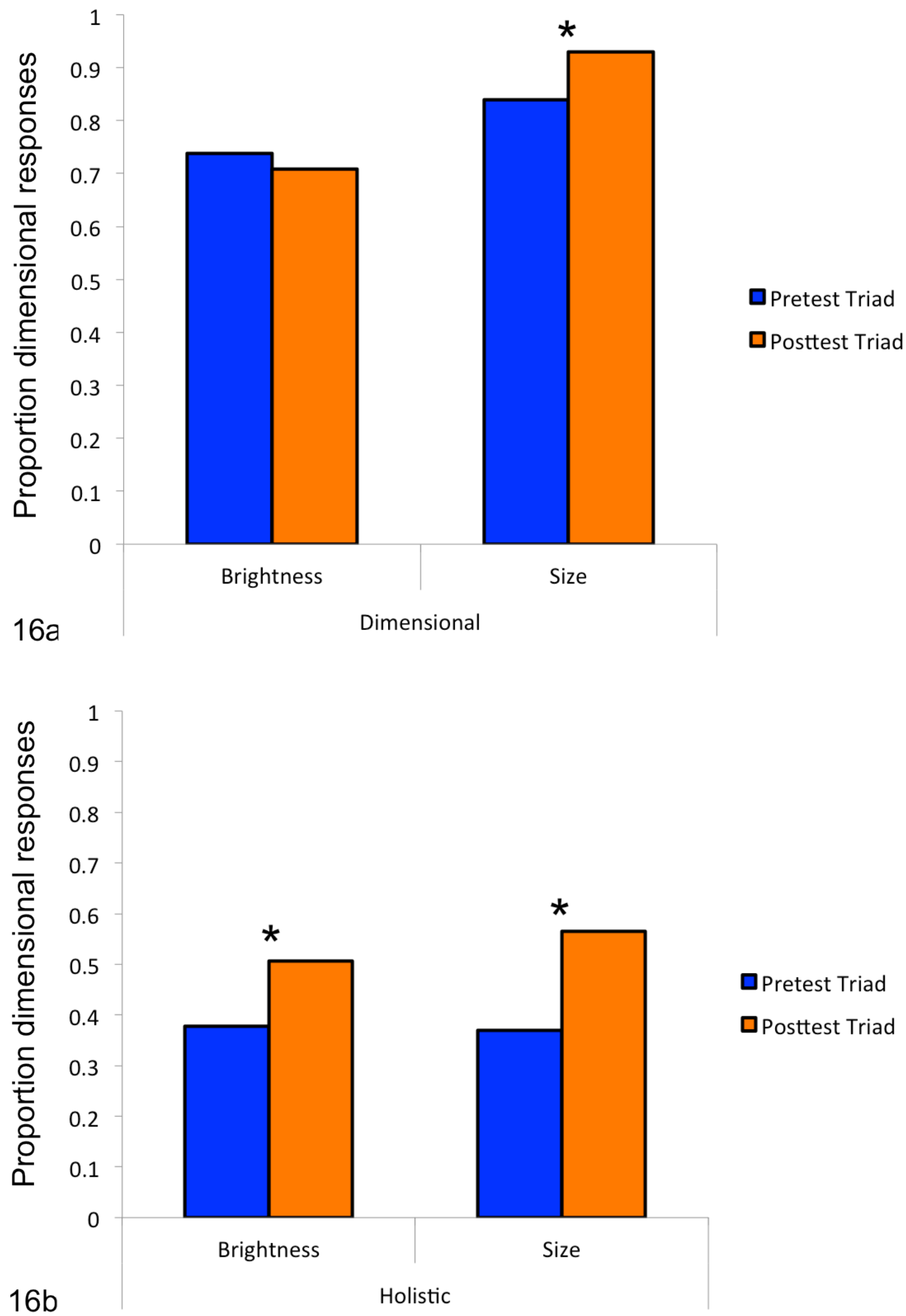
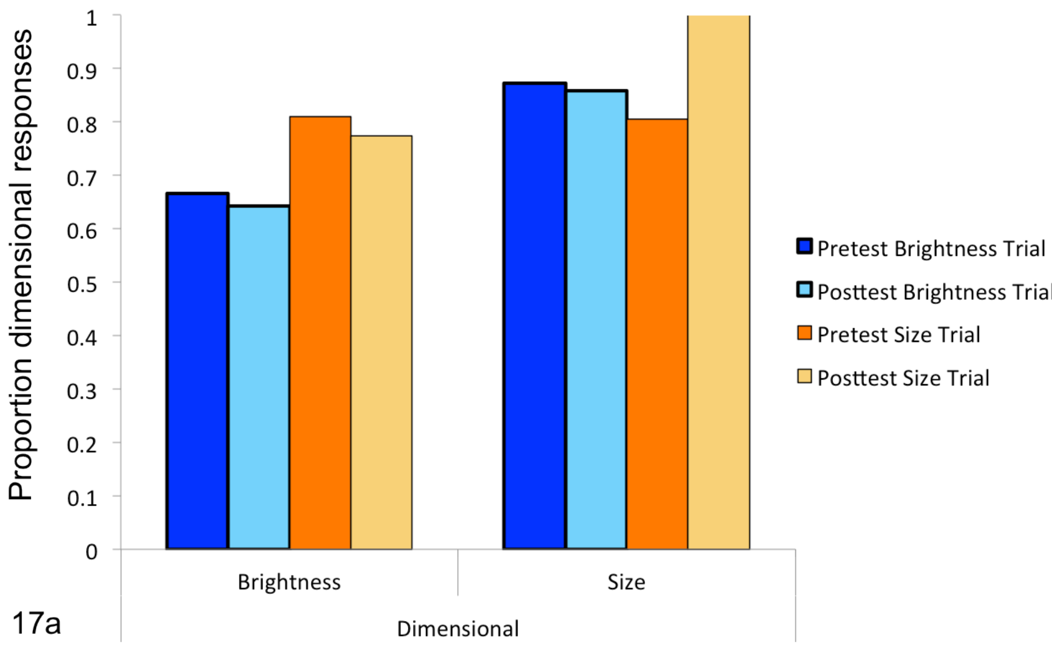
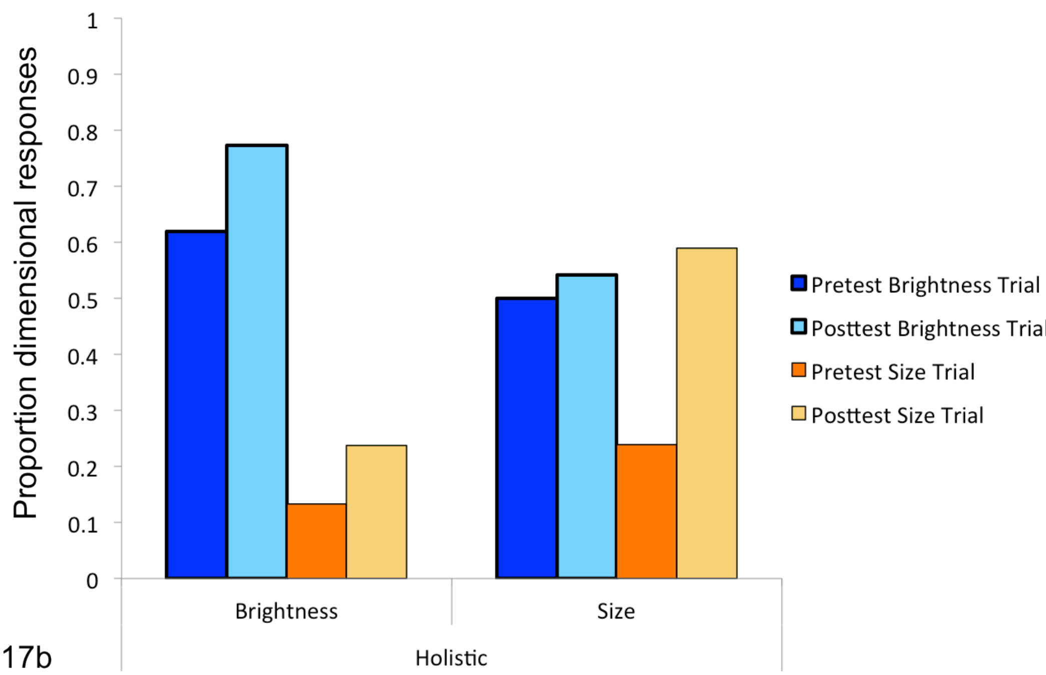


Figure A16. Proportion of dimensional responses that children made during pre- and posttest triad tasks in Experiment 2. Panel A shows the results of children who were dimensional classifiers and Panel B shows the results of children who were holistic classifiers.



17a



17b

Figure A17. Proportion of dimensional responses that children made during pre- and posttest triad tasks for each trial type (whether the dimensional match was a match along the brightness dimension or the size dimensions) in Experiment 2. Panel A shows the results for children who were dimensional classifiers and Panel B shows the results for children who were holistic classifiers.

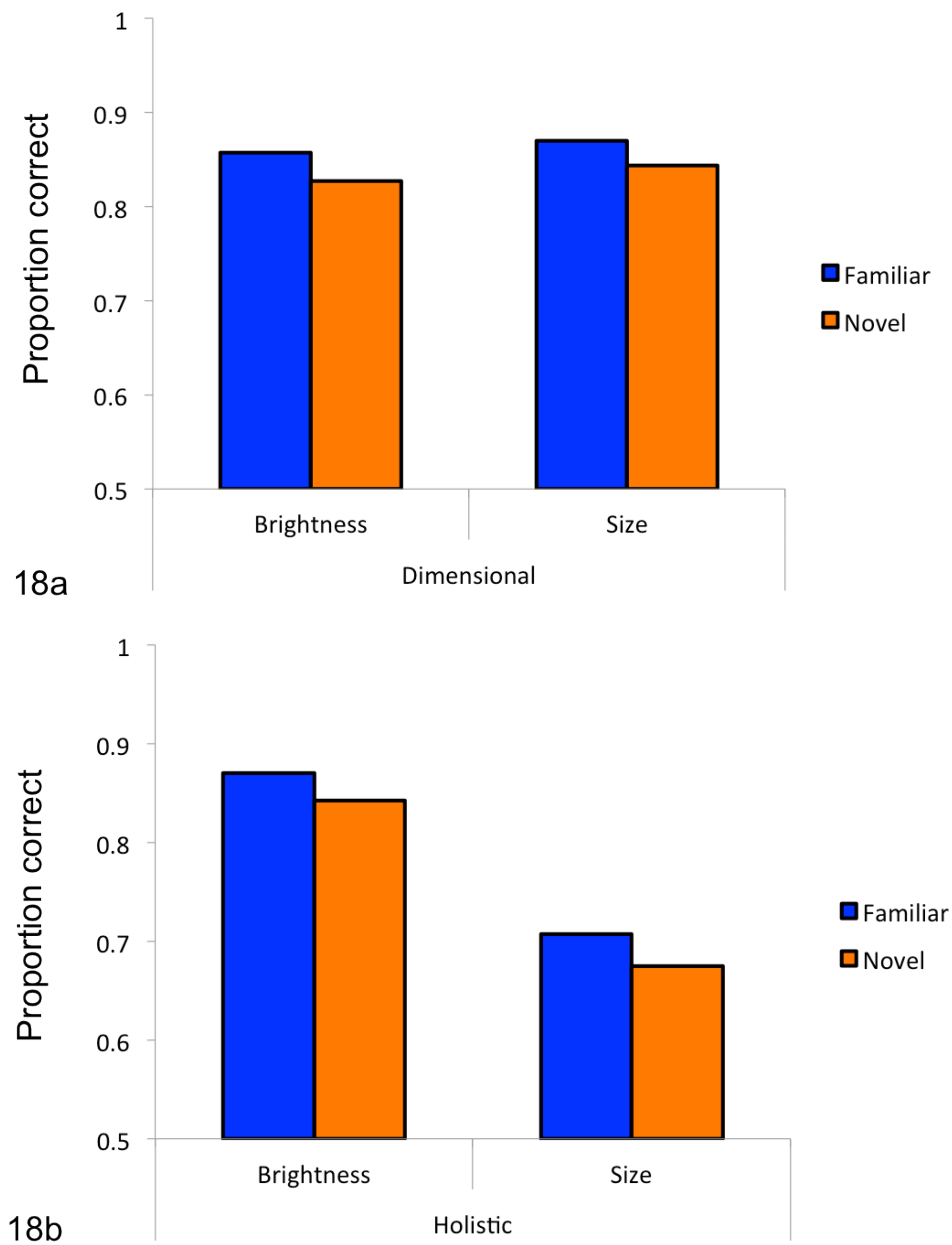


Figure A18. Average accuracy during category test on both familiar and novel trials for brightness learners (left bars) and size learners (right bars) in Experiment 2. Panel A shows the accuracy of children who were dimensional classifiers and Panel B shows the accuracy of children who were holistic classifiers.

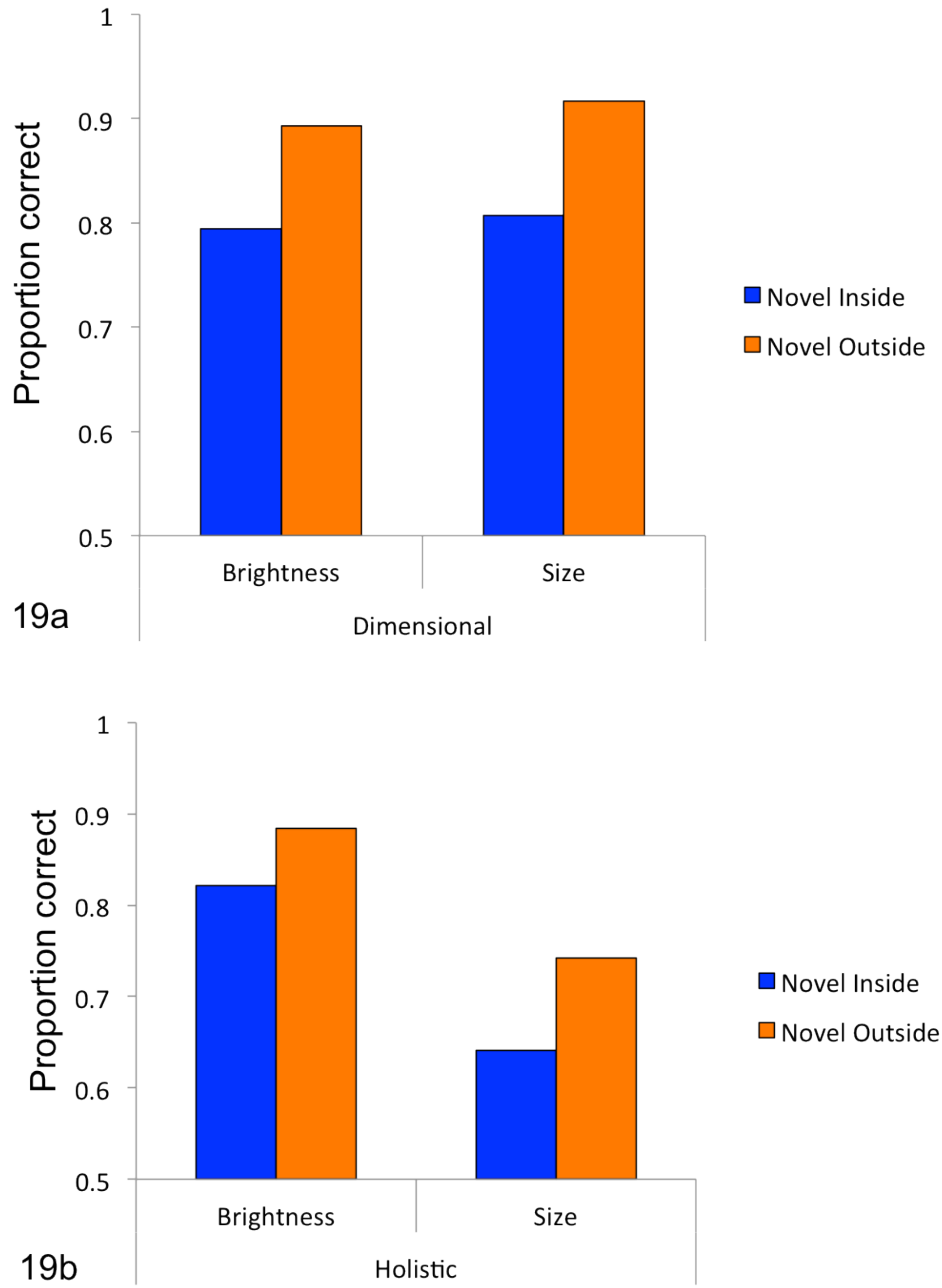


Figure A19. Average accuracy during category test for novel stimuli both inside and outside the category space in Experiment 2. Panel A shows the results for children who were dimensional classifiers and Panel B shows the results for children who were holistic classifiers.

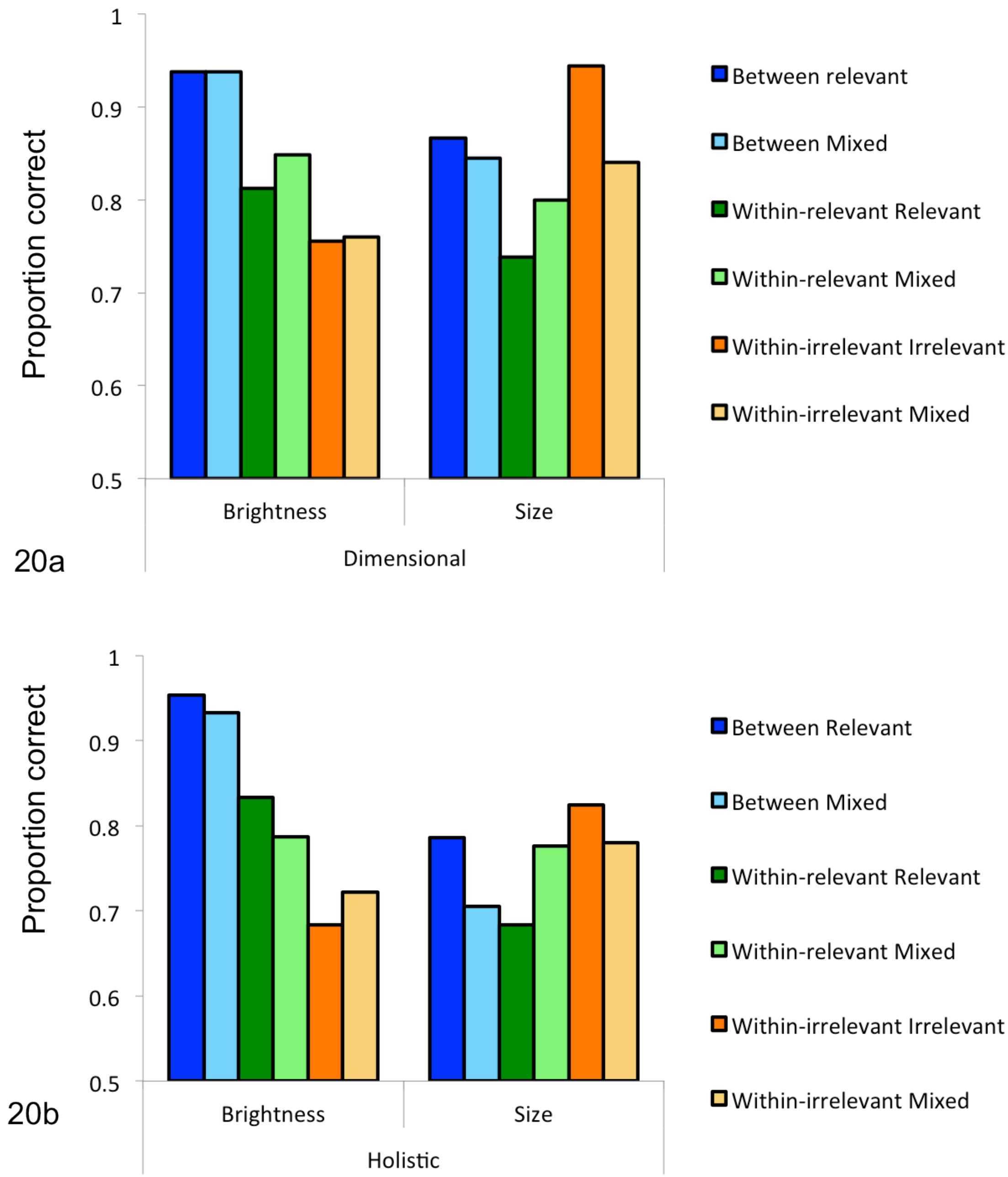


Figure A20. Accuracy for each type of discrimination depending on block structure (all relevant-dimension, all irrelevant-dimension, or mixed-dimension discriminations) in Experiment 2. Panel A shows results from children who were dimensional classifiers. Panel B shows results from children who were holistic classifiers.

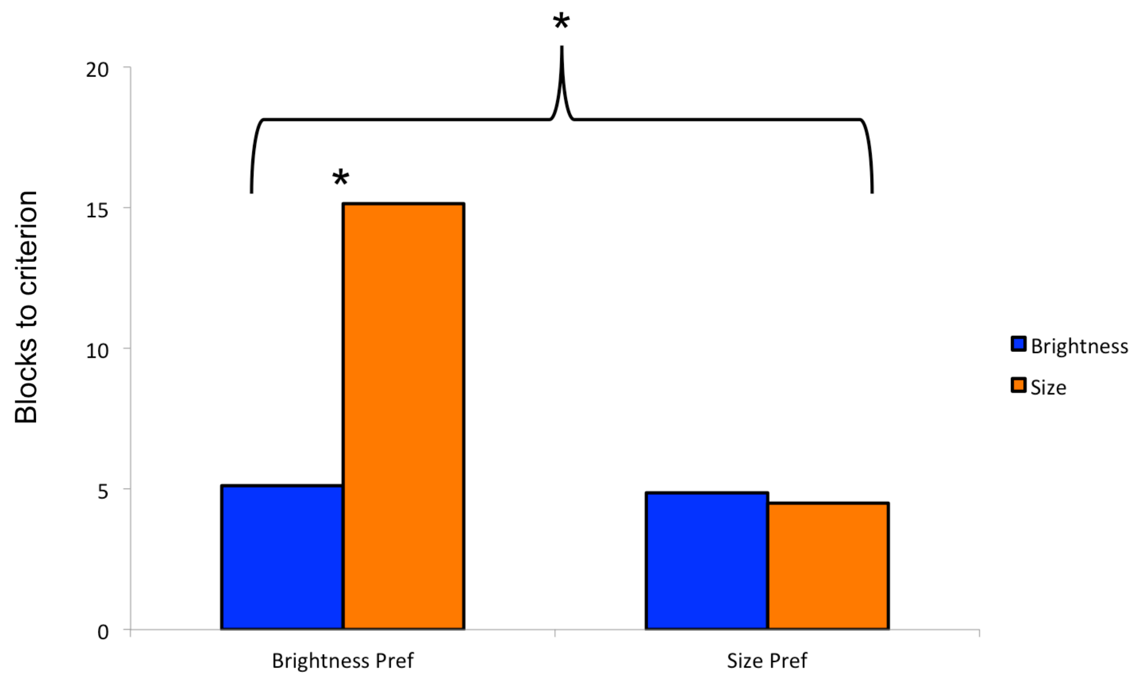


Figure A21. Average number of blocks it took children in each group to reach criterion (e.g. accurate categorization on 7 of 8 trials for 2 blocks in a row) in the category learning task of Experiment 3.

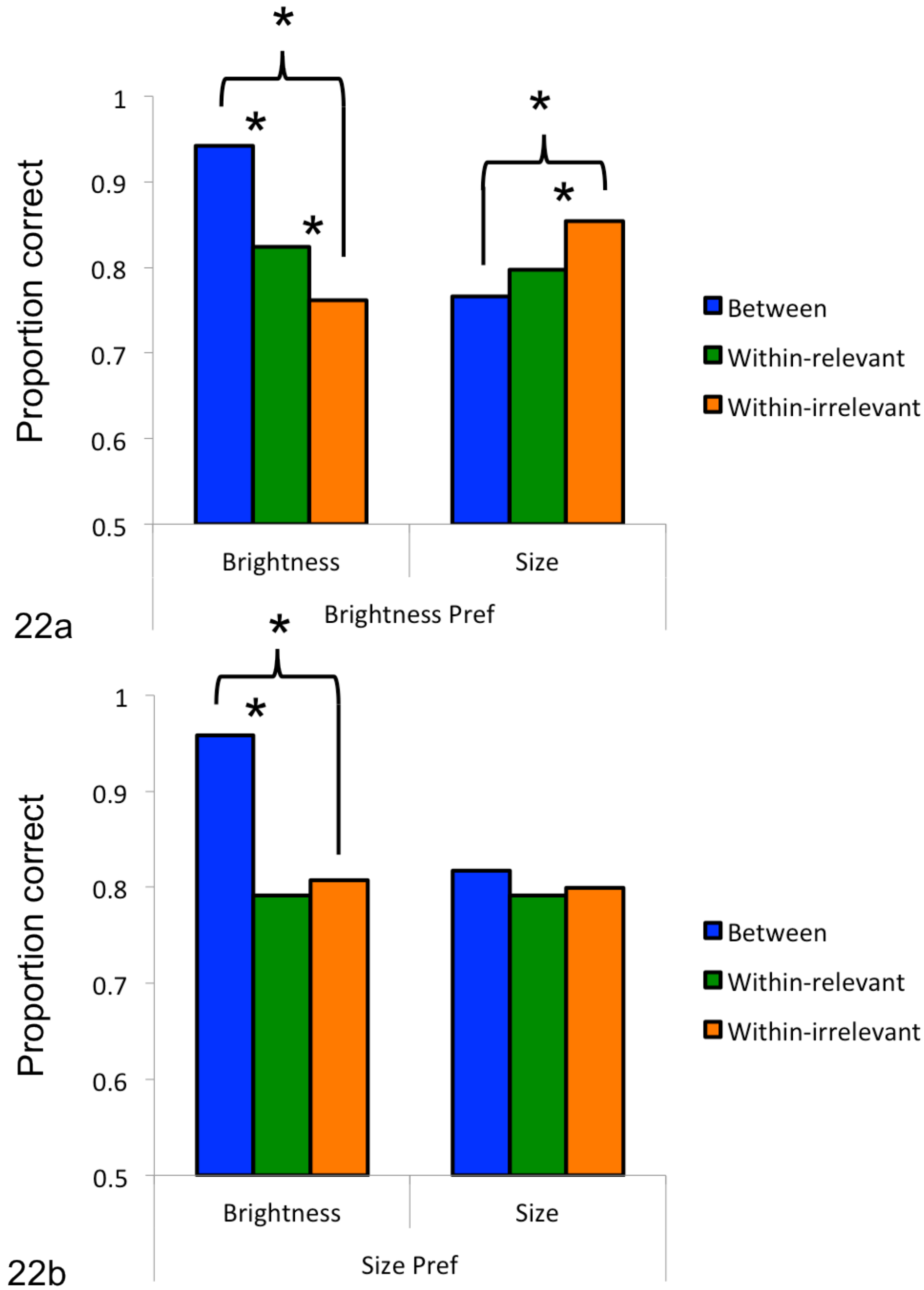


Figure A22. Accuracy for discriminations between categories, within categories along the relevant dimension, and within categories along the irrelevant dimension in Experiment 3. Panel A shows the results of children who learned were brightness preferrers and Panel B shows the results of children who were size preferrers.

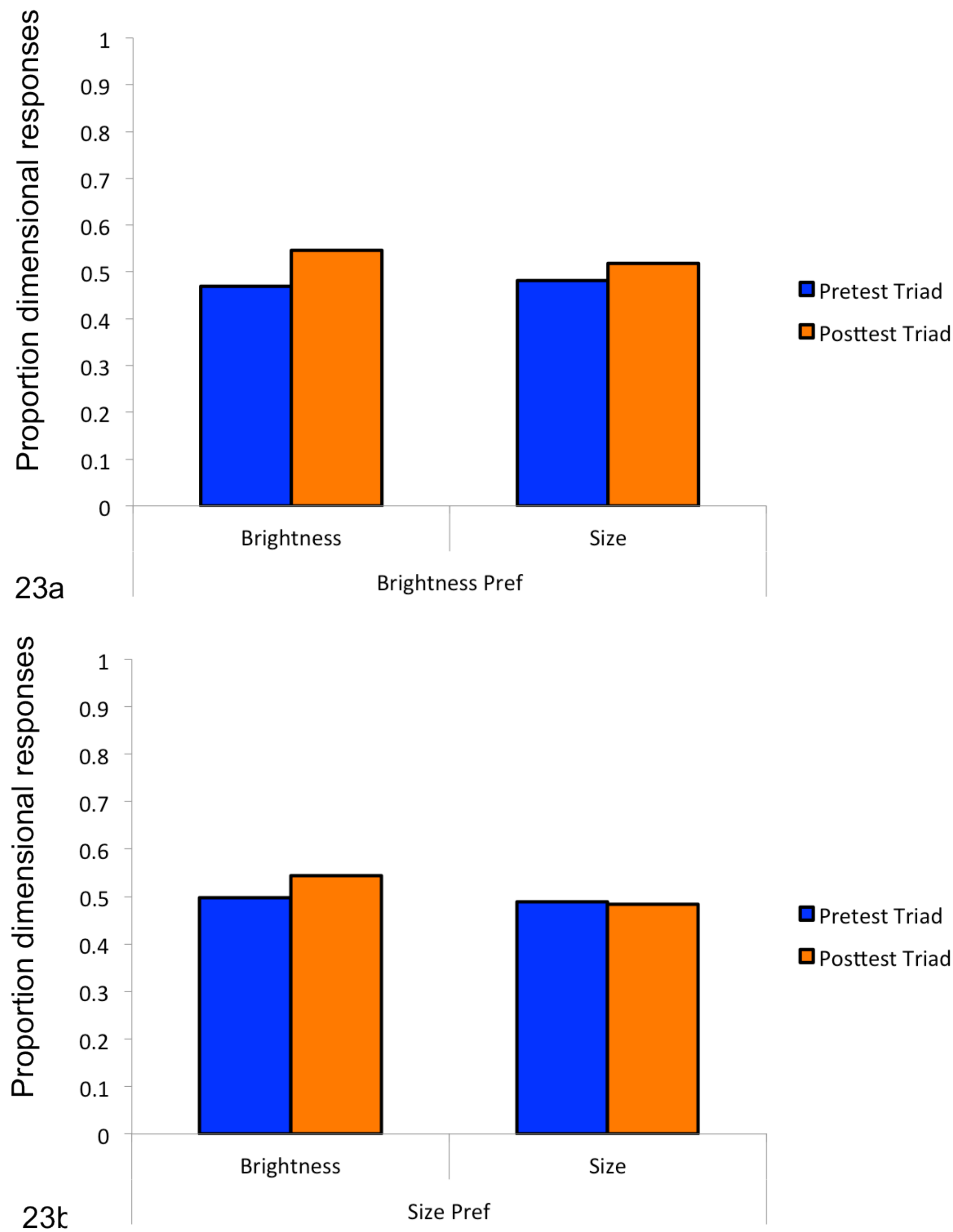


Figure A23. Proportion of dimensional responses that children made during pre- and posttest triad tasks in Experiment 3. Panel A shows the results of children who were brightness preferrers and Panel B shows the results of children who were size preferrers.

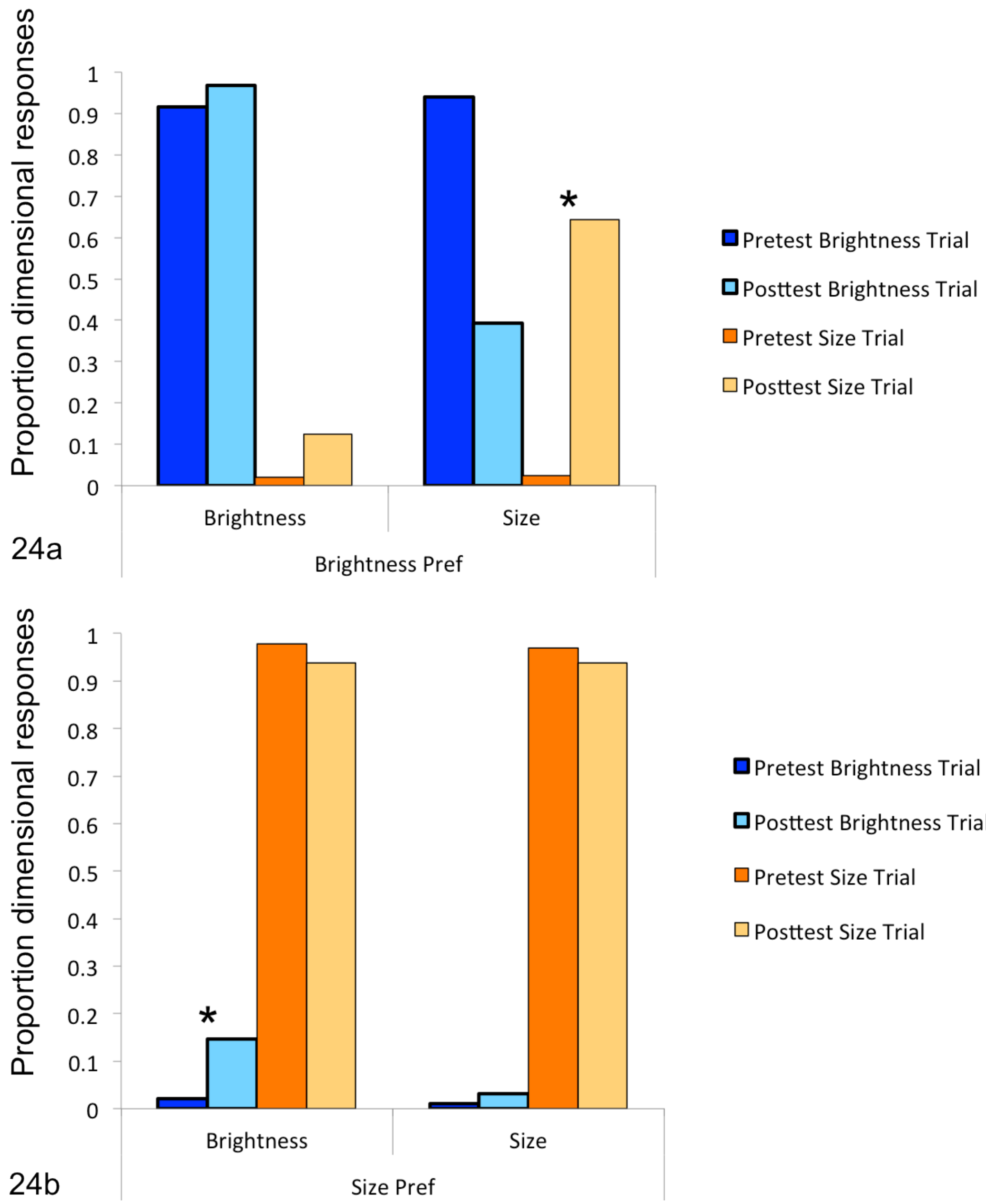


Figure A24. Proportion of dimensional responses that children made during pre- and posttest triad tasks for each trial type (whether the dimensional match was a match along the brightness dimension or the size dimensions) in Experiment 3. Panel A shows the results for children who were brightness preferrers and Panel B shows the results for children who were size preferrers.

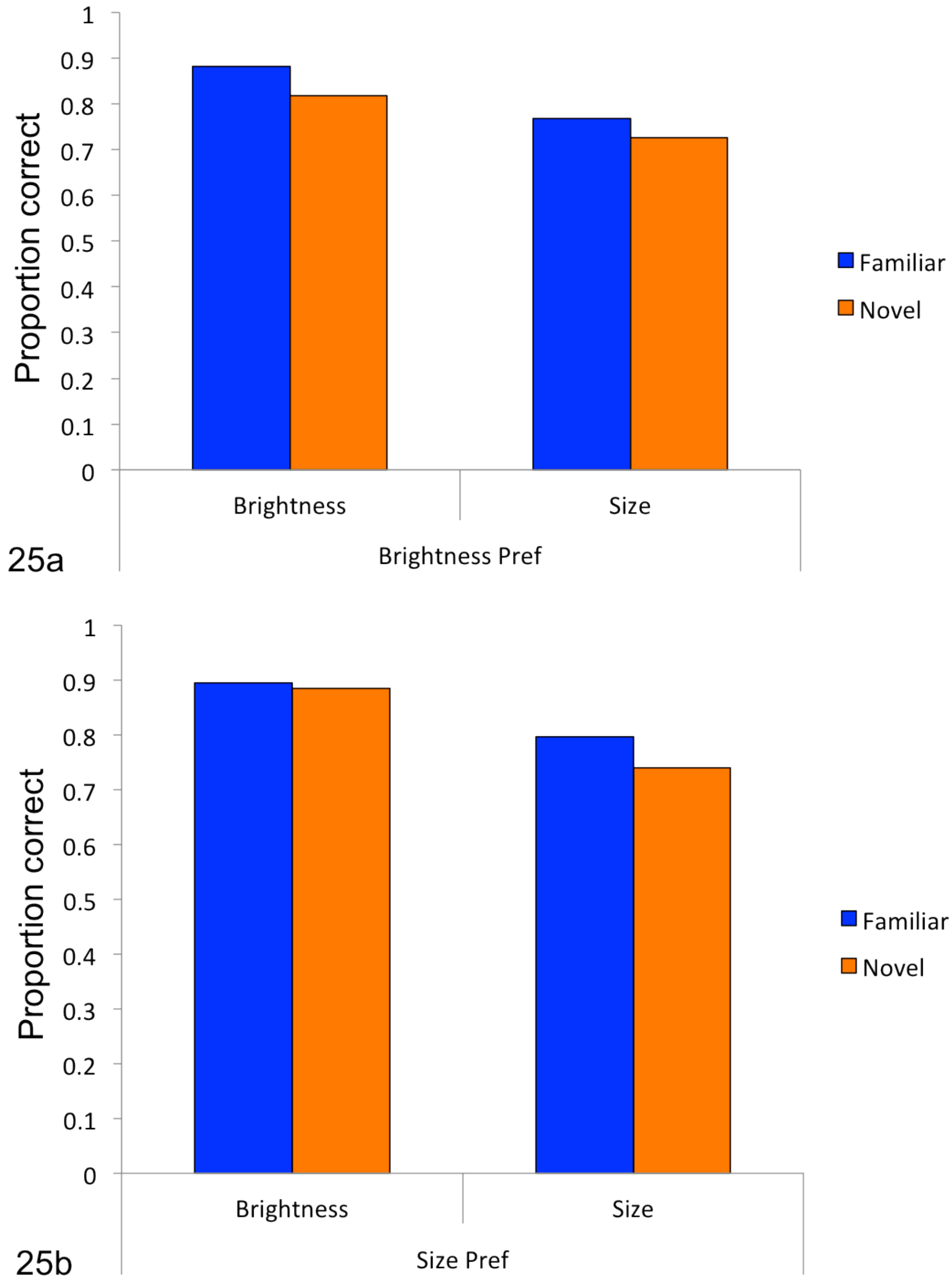
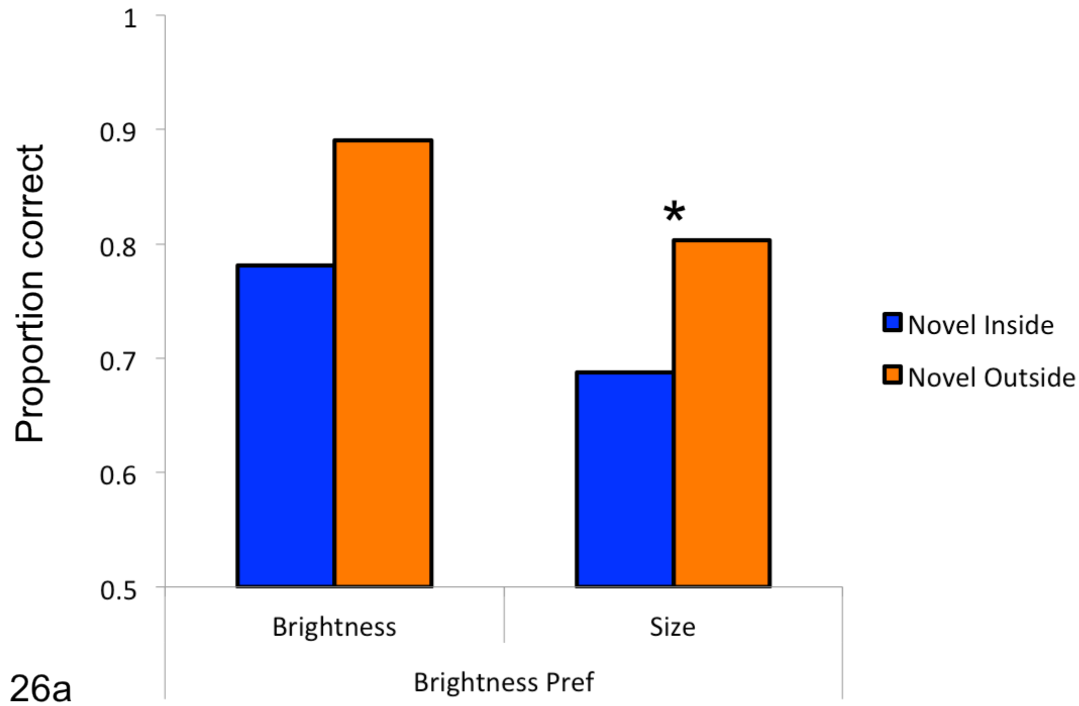


Figure A25. Average accuracy during category test on both familiar and novel trials for brightness learners (left bars) and size learners (right bars) in Experiment 3. Panel A shows the accuracy of children who were brightness preferrers and Panel B shows the accuracy of children who were size preferrers.



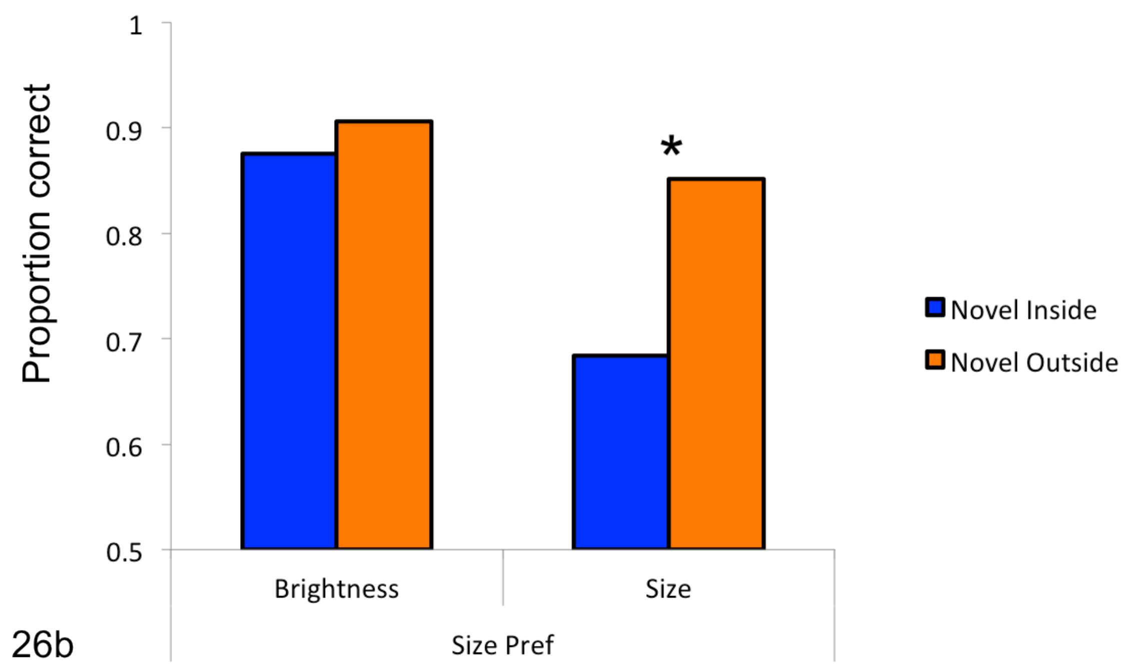


Figure A26. Average accuracy during category test for novel stimuli both inside and outside the category space in Experiment 3. Panel A shows the results for children who were brightness preferrers and Panel B shows the results for children who were size preferrers.

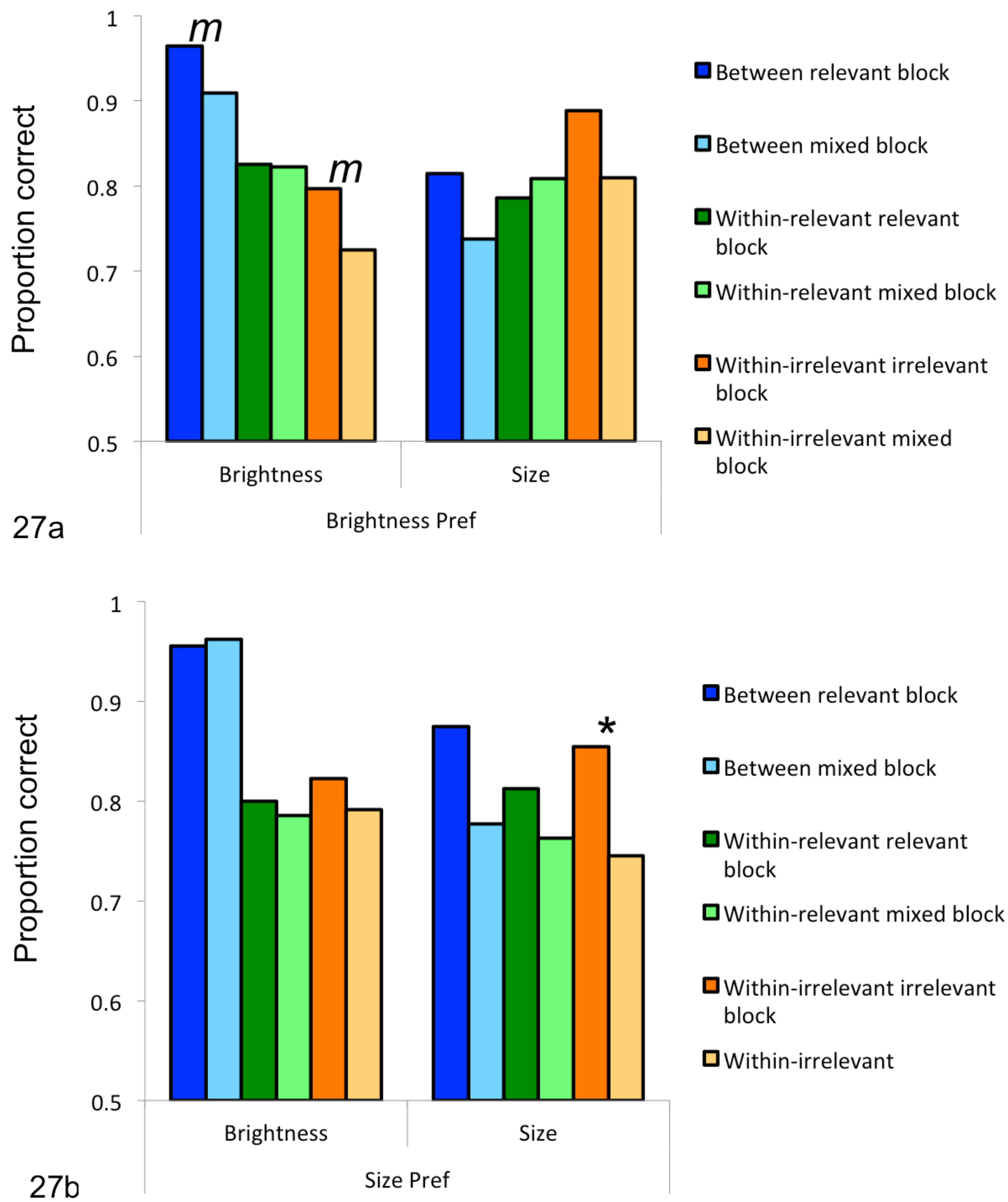


Figure A27. Accuracy for each type of discrimination depending on block structure (all relevant-dimension, all irrelevant-dimension, or mixed-dimension discriminations) in Experiment 3. Panel A shows results from children who were brightness preferrers. Panel B shows results from children who were size preferrers.

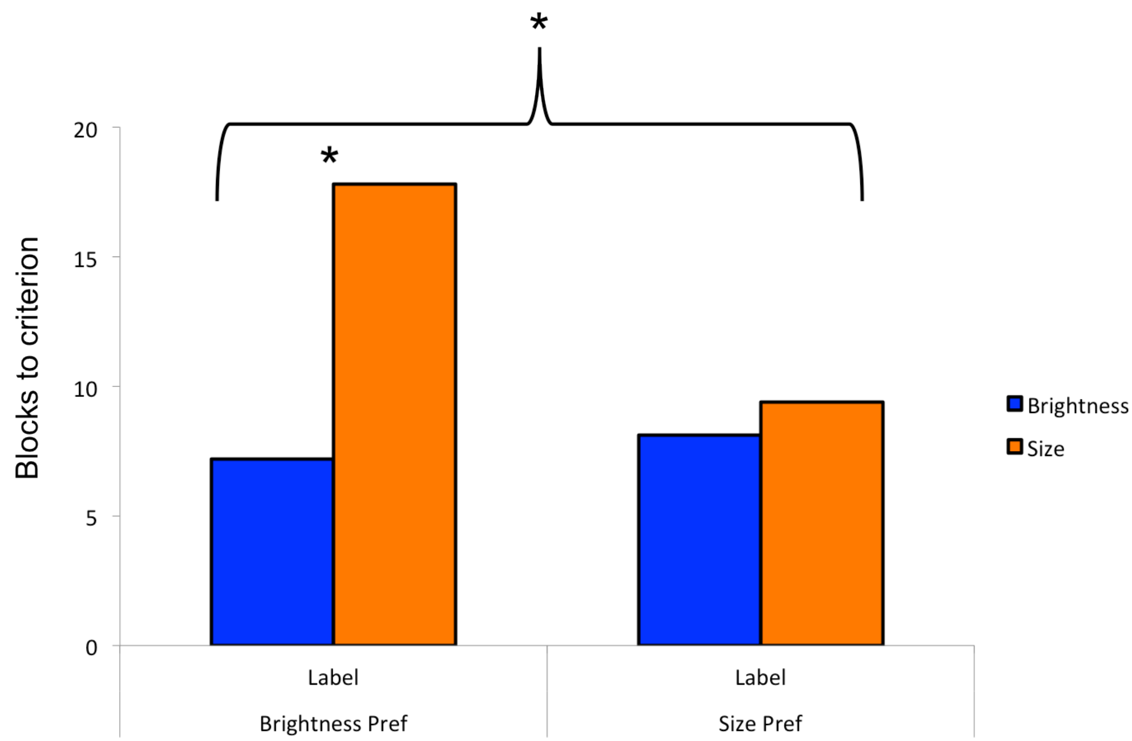


Figure A28. Average number of blocks it took children in each group to reach criterion (e.g. accurate categorization on 7 of 8 trials for 2 blocks in a row) in the category learning task of Experiment 4.

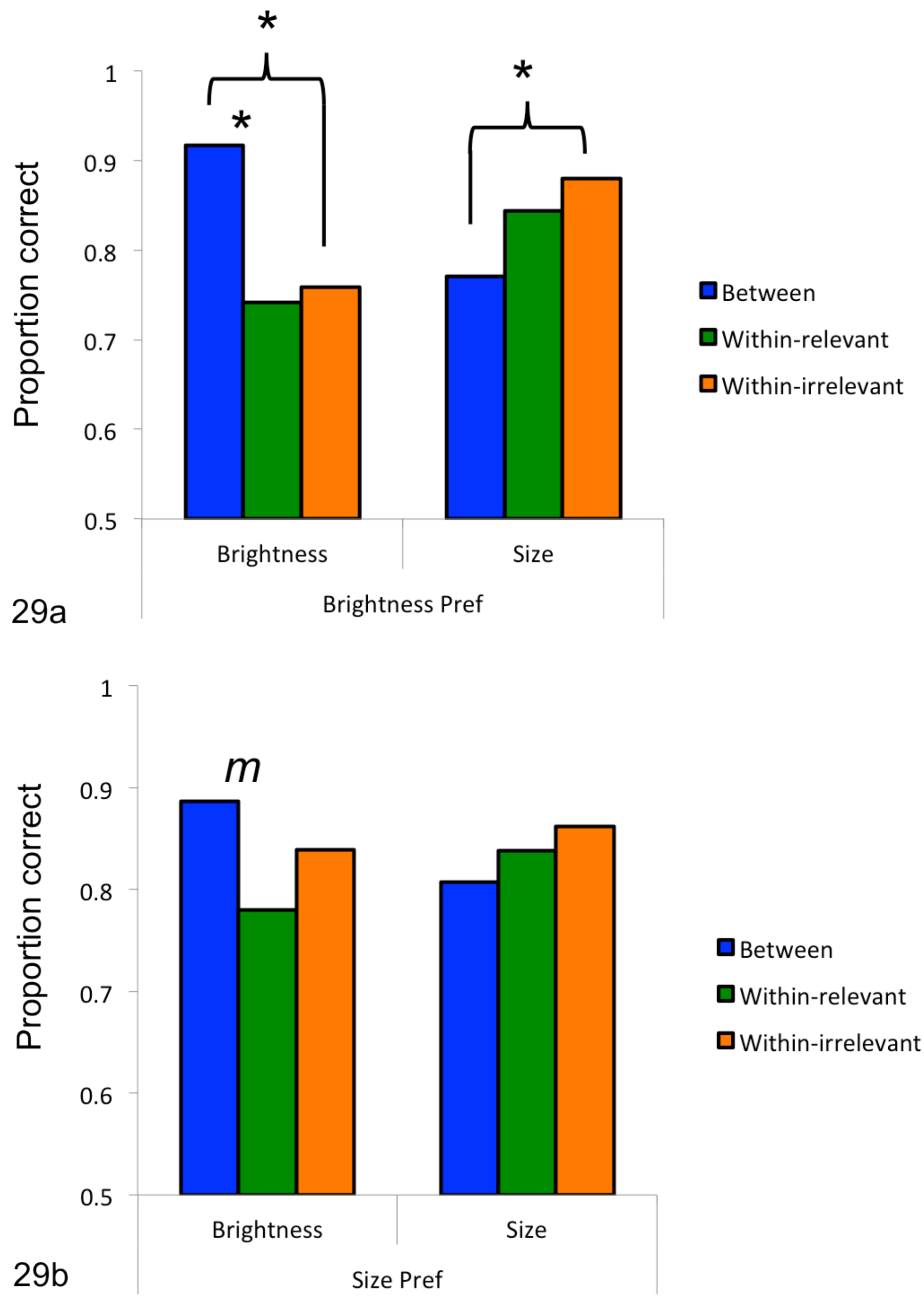
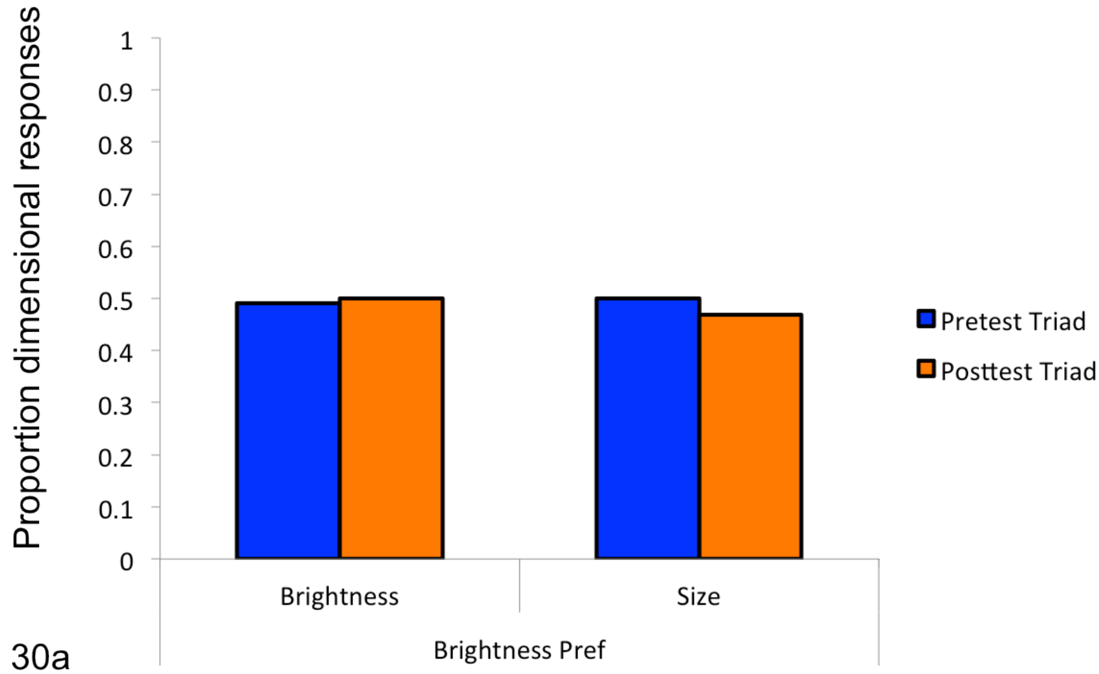
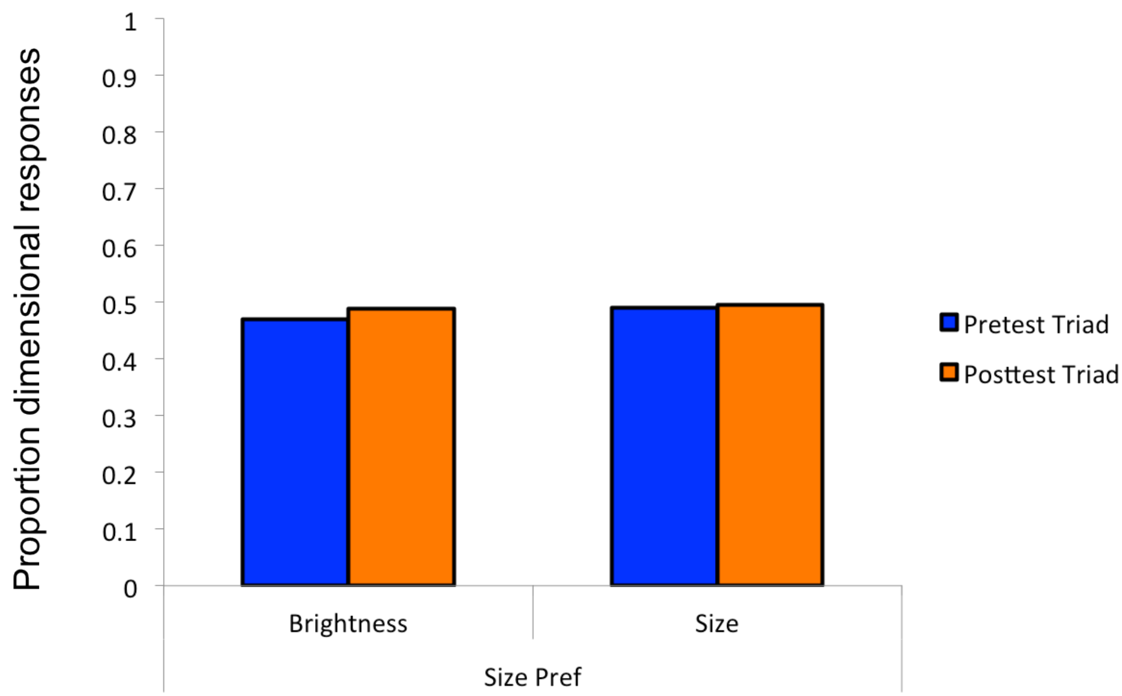


Figure A29. Accuracy for discriminations between categories, within categories along the relevant dimension, and within categories along the irrelevant dimension in Experiment 4. Panel A shows the results of children who learned were brightness preferrers and Panel B shows the results of children who were size preferrers.



30a



30b

Figure A29. Proportion of dimensional responses that children made during pre- and posttest triad tasks in Experiment 4. Panel A shows the results of children who were brightness preferrers and Panel B shows the results of children who were size preferrers.

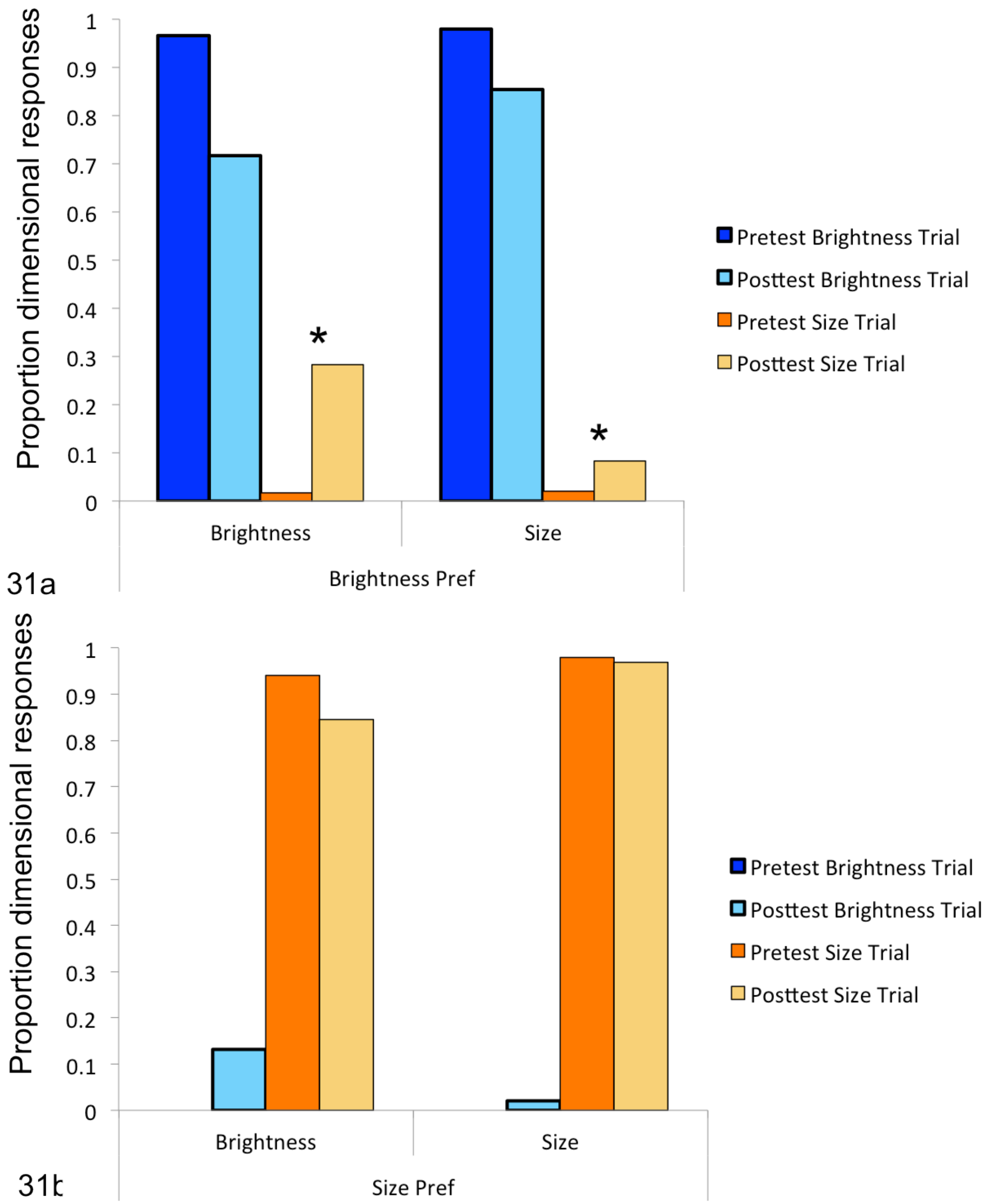


Figure A30. Proportion of dimensional responses that children made during pre- and posttest triad tasks for each trial type (whether the dimensional match was a match along the brightness dimension or the size dimensions) in Experiment 4. Panel A shows the results for children who were brightness preferrers and Panel B shows the results for children who were size preferrers.

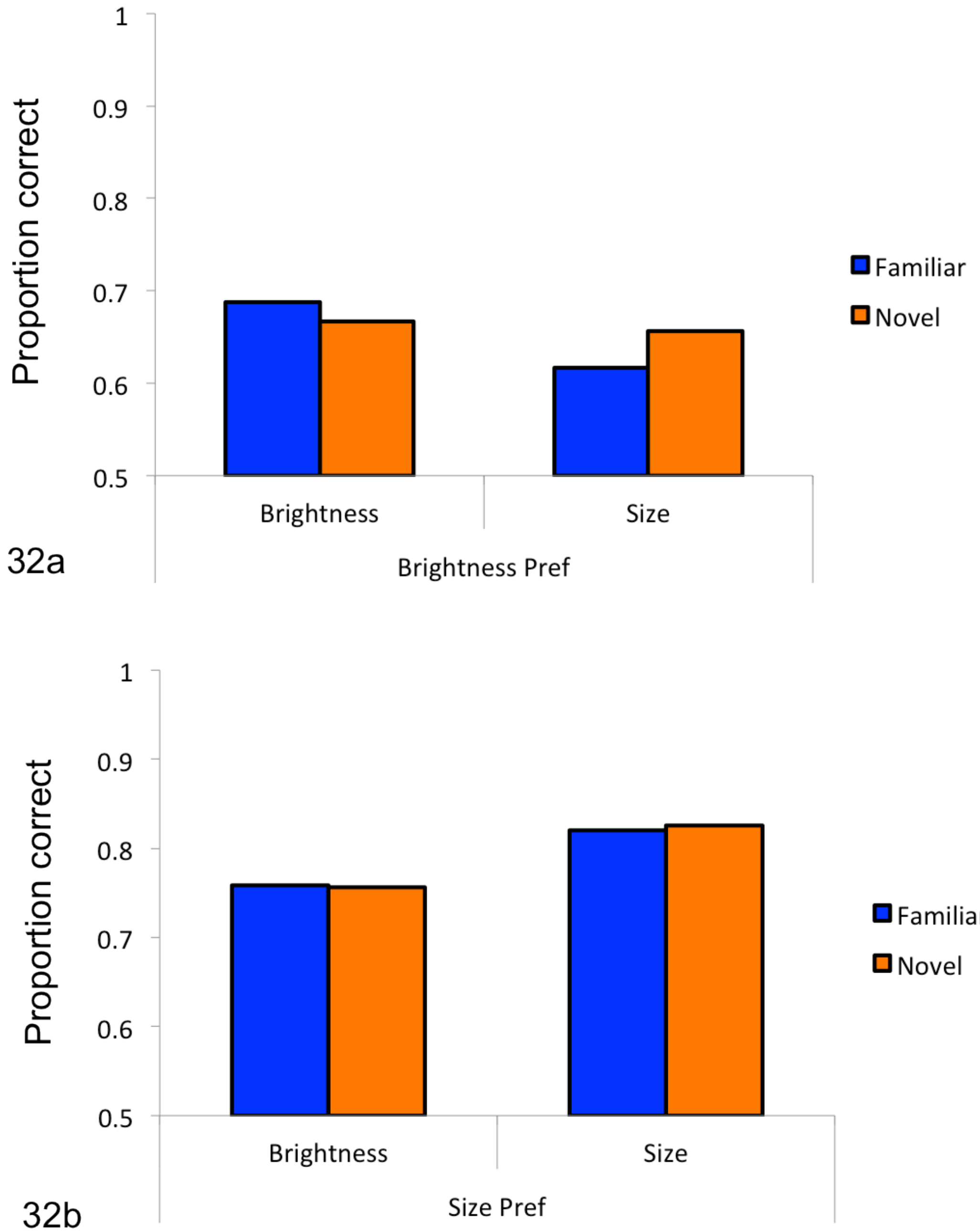


Figure A31. Average accuracy during category test on both familiar and novel trials for brightness learners (left bars) and size learners (right bars) in Experiment 4. Panel A shows the accuracy of children who were brightenss preferrers and Panel B shows the accuracy of children who were size preferrers.

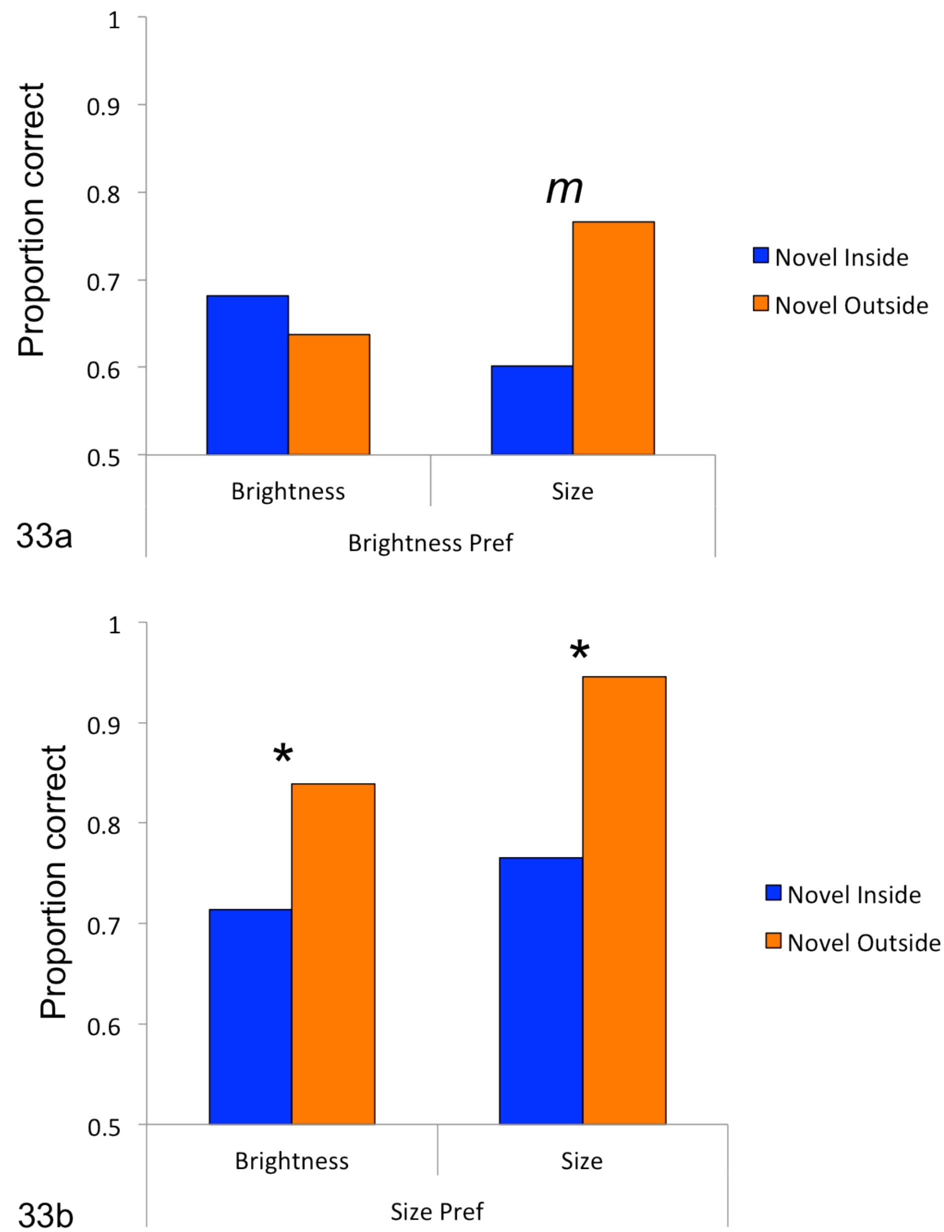
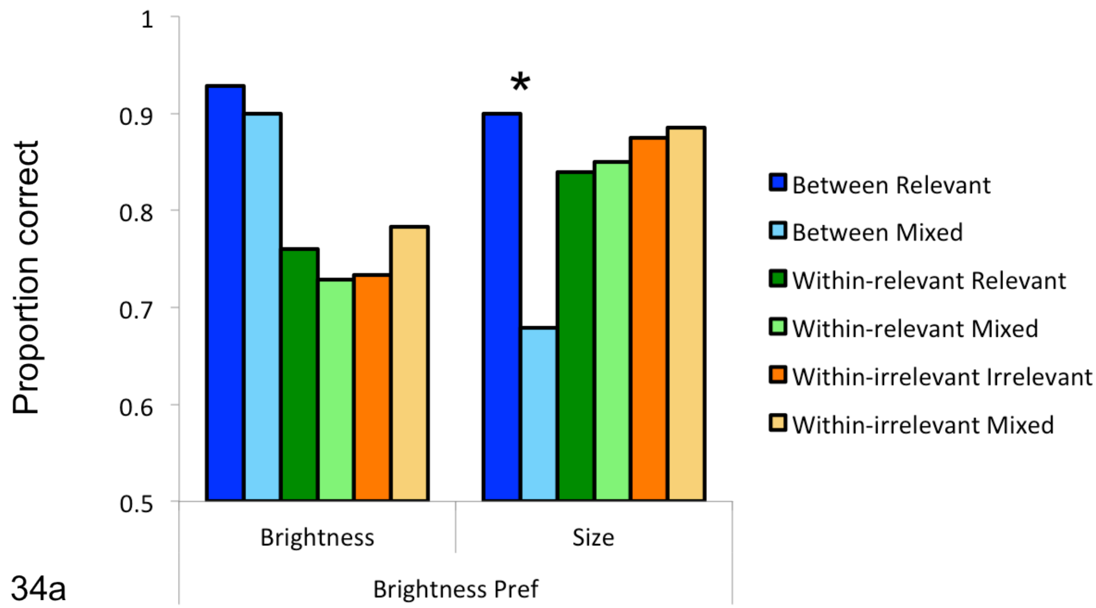
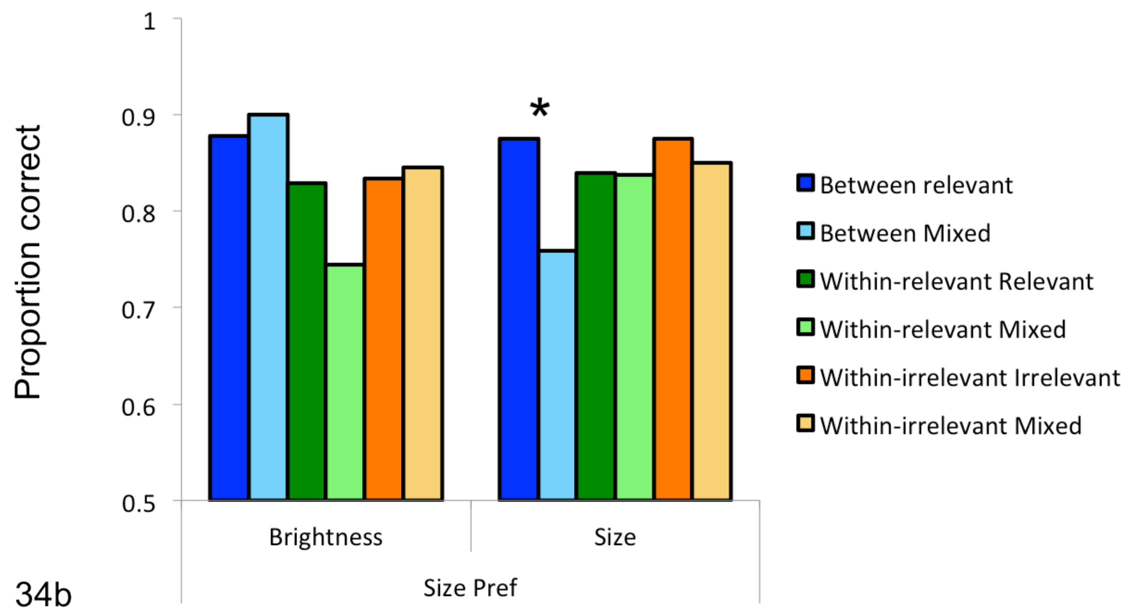


Figure A32. Average accuracy during category test for novel stimuli both inside and outside the category space in Experiment 4. Panel A shows the results for children who were brightness preferrers and Panel B shows the results for children who were size preferrers.



34a



34b

Figure A33. Accuracy for each type of discrimination depending on block structure (all relevant-dimension, all irrelevant-dimension, or mixed-dimension discriminations) in Experiment 4. Panel A shows results from children who were brightness preferrers. Panel B shows results from children who were size preferrers.