UNIVERSITY OF CALIFORNIA, SAN DIEGO

The Development of Color Representations from Infancy to Early Childhood

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

Doctor of Philosophy

in

Psychology

by

Katherine G Wagner

Committee in charge:

Professor David Barner, Chair
Professor Karen Dobkins, Co-Chair
Professor Jonathan Cohen
Professor Sarah Creel
Professor Donald MacLeod

2013

The Dissertation of Katherine G Wagner is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

Co- Chair

Chair

University of California, San Diego

2013

DEDICATION

This dissertation is dedicated to my grandmothers who have served as my inspiration:

Mary Emma Wagner, who as a mother of four received a PhD in Geology in 1972. I only recently realized what a feat that must have been.

Donna K. Chandler (born Sano Humiko), whose determination led her to leave school after sixth grade and do whatever it took to make sure that her eight younger siblings had a roof over their head, food to eat and an education.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my advisors, Karen Dobkins and Dave Barner. The last six years were a bit of a rough ride. Thank you so very much for your support that brought me through it. And of course, thank you both for your invaluable contribution to the research contained in this dissertation. Thank you to my committee and the many people in the department over the years who have been willing to lend a helping hand or a word of advice, and a particular thank you to Don MacLeod for his help on color spaces among other things. Thanks also to Enriqueta Canseco-Gonzalez, Allen Neuringer and Jennifer Corpus who first showed me how thrilling the world of research can be.

I am also very grateful to the numerous lab coordinators, post-docs, fellow grad students and research assistants in the Dobkins and Barner labs over the past six years. Some of who were indispensible in getting data collected, some of who provided statistical advice, some who provided guidance on grad school and some who forced me to get out, socialize and drink too much every now and then.

Thank you to all my friends, in particular Jess Thierman, Hao Ye and Chilezie Nnadi for keeping me sane, well fed and/or healthy.

Thank you to my parents for their valuable support and guidance and to my mom in particular for the occasional advice on probability calculations. And thank you to Adrienne, my sister, for providing me with some uplifting distractions the last few years. And finally, thank you to Madeleine, your cuteness supplied me with the dopamine necessary to get through the last few months of this thing. I hope to spend more time with all of you in the years to come.

Chapter 1, in full, is a reprint of the material as it appears in Synesthetic associations decrease during infancy. *Psychological Science, 22*, 1067-1072. Wagner, K., & Dobkins, K. (2011). The dissertation author was the primary investigator and author of this paper. Permissions for use of this material is granted through SAGE's Global Journal Author Reuse Policy.

Chapter 2, in full, is a reprint of the material as it appears in Slow mapping: Color word learning as a gradual inductive process. *Cognition,* 22, *127,* 307-317. Wagner, K., Dobkins, K. & Barner, D. (2013). The dissertation author was the primary investigator and author of this paper. Permissions for use of this material have been obtained from Elsevier Limited.

Chapter 3, in full, is currently being prepared for submission for publication of the material. Wagner, K, Jergens, J. & Barner, D. The dissertation author was the primary investigator and author of this paper.

CURRICULUM VITAE

2013 Doctor of Philosophy, Psychology
University of California, San Diego

2008 Master of Arts, Psychology
University of California, San Diego

2005 Bachelor of Science, Psychology
Reed College, Portland

PUBLICATIONS

Wagner, K., Dobkins, K & Barner, D. (2013). Slow mapping: Color word learning as a gradual inductive process. *Cognition, 127,* 307-317

Wagner, K. & Dobkins, K. (2011). Synesthetic associations decrease during infancy. *Psychological Science, 22*, 1067-1072.

Canseco-Gonzalez, E., Brehm, L., Brick, C., Brown-Schmidt, S., Fischer, K. & Wagner, K. (2010). Carpet or Cárcel: The effect of age of acquisition and language mode on bilingual lexical access. *Language and Cognitive Processes, 25*, 669-705.

Wagner, K. & Neuringer, A. (2006). Operant variability when reinforcement is delayed. *Learning & Behavior*, 34, 111-123.

Wagner, K., Brang, D., Ramachandran, V. & Dobkins, K. (in revision). Increased color contrast sensitivity in synesthetes.

Wagner, K., Jergens, J., & Barner, D (in prep). Color word comprehension precedes production.

Wagner, K. (in prep). Adaptive developmental decline: Anatomical, linguistic, perceptual.

Kimura, K, Wagner, K. & Barner, D. (in prep). Two for one? Transfer of conceptual content in bilingual number word learning.

Srinivarsan, M., Wagner, K., Frank, M., & Barner, D. (in prep). Abacus expertise and visual attention: Does expertise direct the allocation of attention within visual arrays?

Wagner, K., Cohen, S. & Dobkins, K. (in prep). Effects of personal familiarity on the other-race effect.

ABSTRACT OF THE DISSERTATION

The Development of Color Representations from Infancy to Early Childhood

by

Katherine G. Wagner

Doctor of Philosophy in Psychology

University of California, San Diego, 2013

Professor David Barner, Co-Chair
Professor Karen Dobkins, Co-Chair

This dissertation investigates color representations during early development. It provides new insight into the infant's perceptual experience as well as the process by which children map color words to color percepts.

Chapter 1 provides the first experimental evidence for the long-standing theory that a developmental period of exuberant neural connectivity followed by a retraction and reweighting of connections causes infants to experience increased synesthetic intermingling of their sensations that typical adults do not experience. We show that the presence of particular shapes influences color preferences in typical two- and three-month olds, but not in eight-month olds or adults. These data suggest that infants may experience a sensory phenomenon unlike anything experienced by typical adults but similar to the sensory experiences of adults with synesthesia, a rare sensory phenomenon that has been associated with exuberant neural connectivity and is characterized by strong

arbitrary associations between different sensations.

Chapter 2 examines children's ability to abstract color and assign labels to specific color categories. Children use color words incorrectly for many months before they converge on correct adult definitions. Most current accounts propose that the delay between children's first production of color words and adult-like understanding is due to problems abstracting color as a domain of meaning. Chapter 2 challenges this account with analyses of color word errors in two- to four-year-olds' speech and comprehension. We find that children's errors are systematic and best characterized as overextensions of adult meanings – a finding inconsistent with the notion that these children have not yet abstracted color as a domain of meaning.

Chapter 3 examines color word comprehension in younger children between 18- and 33-months and finds that on average, children construct preliminary meanings for some color words prior to producing any color words. In other words, color words are not an exception to the typical pattern of language comprehension before production. These results also imply that color word errors cannot be due to a failure to abstract color as a domain of linguistic meaning. Rather, the errors likely result from a process through which children gradually converge on correct language specific color word boundaries.

INTRODUCTION

The Development of Color Representations from Infancy to Early Childhood

Humans have the unique ability to supplement their perceptual and cognitive representations with language, a powerful tool that allows us to flexibly highlight the most relevant aspects of a particular experience and easily communicate it with others.

This dissertation focuses particularly on representations of color. Color is an attribute that is part of our every visual experience. We see blue skies, green apples, brown bears, and red lights but we never experience color as a separate entity – it is always an attribute of something else. Yet, we are capable of abstracting and mentally representing color as its own concept. This allows us to easily categorize colors and map language to these color categories.

The chapters of this dissertation explore the origins of our color representations, both linguistic and otherwise. Specifically, we examined the possibility that very early in infancy, color representations are intermingled with other perceptual representations and come to be represented separately through a developmental process that removes excess connectivity. Next, we examined color representations in young children that are beginning to produce color words in their speech but spend several months making many errors before converging on the correct adult-like meanings for these color words. We examine the root of these errors and whether they arise from an inability to abstract and represent color concepts as is frequently proposed (Franklin, 2006; Kowalksi & Zimiles, 2006; O'Hanlon & Roberson, 2006; Sandhofer & Smith, 1999) or if instead they arise out of an immature knowledge of a language's color category boundaries.

Background Information

Color perception and connectivity during infancy

Although color perception is quite poor at birth, it rapidly improves during the first six months of life. Though, blue/yellow color contrast sensitivity (mediated by the koniocellular pathway) is delayed relative to red/green color contrast sensitivity (mediated by the parvocellular pathway)(Crognale, Kelly, Weiss, & Teller, 1998; Suttle, Banks, & Graf, 2002).

Meanwhile, neural connectivity within visual cortex and between visual cortex and other sensory areas is proliferating at a rapid pace. This period of exuberant connectivity is followed by a period of pruning and reweighting of neural connections. Post-mortem synapse counts indicate that infants possess many more synapses than adults (e.g. Huttenlocher, de Courten, Garey, & Van der Loos, 1982), neuronal tracing studies in animals have indicated both longer and more numerous axonal projections during infancy (e.g. Rodman & Consuelos, 1994), and ERP data has revealed that the visual cortex of younger infants responds more to auditory stimuli than that of older infants (Mills, Coffey-Corina, & Neville, 1997; Neville, 1995).

This excess connectivity may in part explain the well-documented superior abilities of infants to discriminate other-race and other-species faces (Kelly et al., 2007; Pascalis, De Haan, & Nelson, 2002), and in the auditory domain, non-native speech sounds (Werker & Tees, 1984). Color perception may also pass through a period of increased connectivity during the normal course of development.

For example, in the adult visual system, motion processing relies primarily on luminance rather than chromatic cues, as the magnocellular pathway that feeds into MT (visual area required for motion processing) is only sensitive to luminance contrast and not to chromatic contrast (Lindsey & Teller, 1990). However, infants are equally capable of using chromatic and luminance information to determine motion (Dobkins & Teller, 1996), though between 2 and 4 months, this ability to use chromatic information for motion processing sharply declines (Dobkins & Anderson, 2002). This loss of ability to use color to determine motion can be understood in terms of a loss of connectivity between the parvocellular red/green system and either MT or the magnocellular system (Dobkins, 2009).

Researchers have suggested that there may be additional consequences of excess neural connectivity during infancy. Namely, the neonatal synesthesia hypothesis suggests that infants may experience a sensory phenomenon similar to those of adult synesthetes but unlike anything experienced by typical adults (Maurer, 1993; Spector & Maurer, 2009). Synesthesia is a rare sensory phenomenon that has been associated with exuberant neural connectivity and is characterized by strong arbitrary associations between different sensations. For example, the most well documented type of synesthesia involves strong associations between graphemes and colors. The neonatal synesthesia hypothesis has important implications for our understanding of color representation in early infancy. It suggests that in very early infancy, representations of color may be intertwined with other visual representations, such as form. As expressed by William James, young infants may experience a “blooming, buzzing confusion” (James, 1890, p. 488) due to the intermingling of their sensations.

Categorical perception of colors in infants

Although adults can discriminate between innumerable shades of colors, they often treat color categorically, and there is an ongoing debate as to whether these categories are perceptual in nature or supported by categorical color language. In support of the idea that color categories are at least in part perceptual in nature, is an increasing amount of evidence that pre-linguistic infants perceive colors categorically (e.g. Bornstein, 1976; Franklin et al., 2008; Franklin, Pilling & Davies, 2005). Furthermore, data thus far indicates that infant categorical perception of colors closely resembles that of adults, specifically English-speaking adults. However, only English color category boundaries have been tested so far. This leaves open the possibility that color categorical perception develops much like phonetic categorical perception where infants begin with many more category boundaries than adults but lose the category boundaries not highlighted in their native language (e.g. Werker & Tees, 1984), perhaps due to the pruning processes discussed above. Thus far however, the evidence indicates that infants do perceive color categorically and their category boundaries are believed to closely resemble that of adults. (Though see Goldstein, Davidoff & Robinson, 2009 for an opposing viewpoint that categorical perception of colors is caused by the acquisition of color words rather than existing pre-linguistically).

Abstraction of color in infancy and early childhood

Mature adult representations of color include the ability to symbolically represent color via language. Acquiring color language not only requires the ability to perceive color and represent it categorically but it also requires the ability to attend to color and abstract it conceptually.

Infants attend earlier to shape and more to shape than to surface properties like color (Wilcox, 1999). However, there is also evidence that depending on the nature of the task, that infants *can* attend to and abstract color by 5 months (Catherwood, Crassini, & Freiberg, 1989), 9.5 months (Wilcox, Woods & Chapa, 2008) or 11.5 months (Wilcox, 1999). For example, Wilcox, Woods & Chapa (2008) found that although 9.5 month olds naturally attend to shape over color, if they are trained that color predicts function with multiple color exemplars, they can learn to attend to color over shape.

In light of such infant data, it is surprising that there is very mixed evidence as to whether toddlers are capable of conceptualizing color prior to their acquisition of color terms. Soja (1994) shows that children who do not yet know color terms are capable of matching by color and using color to predict ownership or distinguish one character from another. However, children fail to match colors in more difficult tasks that require encoding color in memory and attending to color over shape (Sandhofer & Smith, 1999; Kowalski & Zimiles, 2006) until around the time that they master color word meanings.

The acquisition of color language

Children's categorical perception of colors and their ability to abstract color properties clearly have important implications for our understanding of the color word acquisition process. Color words are thought to be particularly difficult to acquire because when asked questions like *What color is it?,* children use incorrect but domain appropriate color words (e.g. *red* when referencing a yellow item) for several months before they converge on adult-like definitions. There are a number of reasons children may have difficulty with color terms but they can mostly be classified into two categories, and these two types of difficulties predict very different types of errors during the delay between color word production and adult like mastery of color word meanings. 1) *Difficulties abstracting color:* Children have a difficult time abstracting color properties and identifying them as a domain of linguistic meaning. In this case, children's color word errors should be random and unrelated to the color properties being labeled. 2) *Difficulty determining the perceptual boundaries of color terms*: Children take time to converge on the necessary color category boundaries. In this case, color word errors should be systematically related to the color properties being labeled. For example, a child that has an immature understanding of the boundaries of *red,* may be more likely to mistakenly label an orange item *red* than a blue item because orange is perceptually more similar to red than blue.

Current Directions

The following dissertation chapters contribute new important pieces to our understanding of the development of color representations. In Chapter 1, we explore the possibility that in very early infancy, representations of color may be intertwined with other visual representations, such as form. We test whether young infants experience increased synesthetic associations between color and shapes relative to older infants and adults.

In Chapter 2, we carefully analyzed the errors that children make to determine whether these errors reflect a failure to abstract color or a difficulty determining the appropriate perceptual boundaries for color terms.

In Chapter 3, we examine color word comprehension in a younger group of children that includes a subset of children who have yet to produce any color words. This allows us to determine whether color words are an exception to the general rule in language acquisition that comprehension precedes production.

Chapter 1

Synaesthetic Associations Decrease During Infancy

Katherine Wagner and Karen R. Dobkins

University of California, San Diego

Supplemental Material

Additional supporting information may be found at

http://pss.sagepub.com/content/by/supplemental-data

Corresponding Author:

Karen R. Dobkins

Department of Psychology, 0109

University of California, San Diego

La Jolla, CA 92093

E-mail: kdobkins@ucsd.edu

Abstract

Early development is characterized by a period of exuberant neural connectivity followed by a retraction and reweighting of connections over the course of development. It has been proposed that this connectivity may facilitate arbitrary sensory experiences in infants that are unlike anything experienced by typical adults but are similar to the sensory experiences of adults with synaesthesia, a rare sensory phenomenon that has been associated with exuberant neural connectivity and is characterized by strong arbitrary associations between different sensations. We provide the first evidence for this infant-synaesthesia hypothesis by showing that the presence of particular shapes influences color preferences in typical 2- and 3-month-olds, but not in 8-month-olds or adults. These results are consistent with two possibilities: first, that exuberant neural connectivity facilitates synaesthetic associations during infancy, but that these associations are typically eliminated during development, and, second, that a failure of the retraction process leads in rare cases to synaesthesia in adults.

Synaesthetic Associations Decrease During Infancy

The neonatal-synaesthesia hypothesis proposes that synaesthesia, a condition wherein stimulation of one sense involuntarily evokes an additional arbitrary stimulation of another sense, may be a universal experience in young infants, but then dissipates during the normal course of development (Maurer, 1993). The hypothesis suggests that synaesthesia may occur at high rates during infancy because of increased functional connectivity in the infant brain relative to the adult brain (e.g., Mills, Coffey-Corina, & Neville, 1997; Neville, 1995). The neural underpinning of synaesthesia may be the well-documented exuberant anatomical connectivity that occurs during infant development (e.g., Huttenlocher, de Courten, Garey, & Van der Loos, 1982; Rodman & Consuelos, 1994) or the potential decreased inhibition of feedback projections early in development (see Spector & Maurer, 2009; Eagleman & Goodale, 2009, for discussion). Furthermore, if this increased functional connectivity does not dissipate over the course of development, it may result in synaesthesia in adults (e.g. Ramachandran & Hubbard, 2001; Rouw & Scholte, 2007). Although the neonatal-synaesthesia hypothesis has been discussed for nearly 20 years (Maurer, 1993), increased rates of synaesthesia in young infants have yet to be systematically demonstrated.

Recently, there have been a growing number of studies investigating synaesthetic-like associations in infants and children. For example, 2- to 3-year-old children associate high-pitched sounds with small and light objects (Mondloch & Maurer, 2004), preliterate toddlers associate Xs with the color black and Os with white (Spector & Maurer, 2008), and 3-month-old infants associate high-pitched sounds with tall and sharp visual objects (Walker et al., 2010). Some of these associations seem to be arbitrary, and this leads us to propose that, as opposed to being learned from the environment, these associations are remnants of neonatal synaesthesia. However, the associations in infants and children reported in previous studies have also been observed with similar prevalence in typical adults (Marks, Hammeal, & Bornstein, 1987; Melara & O'Brien, 1987; Spector & Maurer, 2008; Ward, Huckstep, & Tsakanikos, 2006). Thus, such results in infants and children do not fulfill one of the critical core predictions of the neonatal-synaesthesia hypothesis, namely, that synaesthetic associations decline with age as exuberant connections are pruned.

In contrast with previous studies, the study reported here examined the age-related decline in exuberant connectivity by testing associations in infancy that are thought to be very rare in adults. Specifically, our study focused on the best-documented form of synaesthesia in adults: grapheme-color synaesthesia. In grapheme-color synaesthesia, which is seen in only about 1% of adults, specific letters or numbers evoke idiosyncratic, largely individualized sensations of specific colors (Ramachandran & Hubbard, 2001).

Rather than using graphemes, which are complex and likely to be unfamiliar to young infants, we chose to use simple shapes. Our reasoning was that shape recognition in infants may be a precursor to grapheme recognition in adults, as both are thought to be processed in the same region of visual cortex (McCandliss, Cohen, & Dahaene, 2003); in addition, it has recently been shown that the shapes of graphemes influence grapheme-color synaesthetic associations (Brang, Coulson, & Ramachandran, in review).

In our study, we presented shapes (circles or triangles) on backgrounds of two separate colors, either red versus green (see Fig. 1) or blue versus yellow. If a participant possessed shape-color associations, we expected the presence of shapes (either circles or triangles) to differentially affect the participant's background color preferences. This could happen in one of two ways. Suppose that a given infant associates triangles with red. First, the infant may prefer to look more to the red than to the green side of the screen when triangles are presented because red is congruent with this infant's associations. The alternative is that if the triangles actually "appear" red for this infant, they might stand out more (and thus be easier to detect) on the green rather than the red background. This would result in the infant preferring to look more often to the green side of the screen (see Smilek, 2001, for evidence of this second type of phenomenon in an adult synaesthete).

Although our data cannot show which of these two scenarios is more accurate (which means our data cannot help us determine specific shape-color associations made by individual infants), this is not of concern; the objective of the study was to determine whether associations (regardless of their nature) are more prevalent in younger infants than in older infants. Furthermore, our methodology was designed to accommodate the likelihood that, like the highly individualized grapheme-color associations in adult synaesthetes (e.g., one adult synaesthete might associate "E" with green, and another might associate "E" with red), shape-color associations made by infants should also be highly individualized (e.g., one infant might associate triangles with red, and another might associate triangles with green).

We also hypothesized that synaesthetic pairings of colors mediated primarily by the koniocellular pathway (blue and yellow) may occur later in development than pairings mediated by the parvocellular pathway (red and green). This prediction was made because sensitivity to red and green is present by 2 months of age (e.g., Dobkins, Anderson & Kelly, 2001), but sensitivity to blue and yellow is thought to emerge around 3 to 4 months of age (Crognale, Kelly, Weiss, & Teller, 1998; Suttle, Banks, & Graf, 2002). Thus, we chose to study 2- to 3-month-olds because color pathways are still actively developing in this age range, and in addition, they are developing differentially for red and green versus blue and yellow. Thus, synaesthetic associations for these two sets of colors may develop over different time courses. We also tested 8-month-olds and adults as older comparison groups, because we predicted that shape-color associations would be much more rare past 3 months of age.

Method

Participants

Infant participants were recruited by mass mailings to new parents in San Diego County. All infants were full term, between 38 and 42 weeks gestation at birth, and participated within 1 week of their 2-month, 3-month, or 8-month birthday. Adult participants were 18 to 23 years old (M = 20.4 years) and received course credit for psychology classes at the University of California, San Diego. All participants were screened for color-blindness via family-history interview (infants) or Ishihara color plates (adults).

Subjects were excluded from analysis if they completed less than 25 trials per condition (one 2-month-old, one 3-month-old, three 8-month-olds) or performed contrary to expectation on 15% or more of catch trials (five 2-month-olds, five 3-month-olds, two 8-month-olds, 6 adults). Most adult exclusions resulted from participants wearing glasses, which made it difficult to see their eye movements. Additionally, one 3-month-old was excluded because of parent interference with the task, and one 3-month-old was excluded because he had torticollis, making his eye and head movements difficult to interpret. After exclusions, participants were fifteen 2-month-olds (9 males, 6 females), fifteen 3-month-olds (9 males, 6 females), fifteen 8-month-olds (4 males, 11 females), and 16 adults (3 males, 13 females).

Stimuli

Stimuli were programmed in MATLAB (The MathWorks, Natick, MA) using Psychtoolbox 2 (Brainard, 1997) and were presented on a 17-in EIZO monitor (Ishikawa, Japan) driven by a Dell PC laptop with an 8-bit ATI Radeon graphics card (AMD, Sunnyvale, CA). Each trial consisted of a field of 24 hollow circles or triangles (each subtending ~6.4°, line width = 1.7°, entire display = 56° × 42°) presented on a background consisting of separate colors on the left and right side of the display.

We chose to limit our investigation to four hues considered unique according to the opponent-process theory of color (red, green, blue, and yellow; Sternheim & Boynton, 1966); these colors were presented as two "chromatically opponent" pairs—red versus green and blue versus yellow—for three reasons. First, red-green contrast and blue-yellow contrast are mostly processed in separate visual pathways, and this phenomenon allowed us to compare the results of our different age groups with the known developmental time course of these pathways. Second, previous research on adult synaesthetes has shown that color naming is facilitated when graphemes are presented in a color congruent with each synaesthete's associations, and interference is greatest when the graphemes are presented in a color chromatically opponent to each synaesthete's associations (Nikolíc, Lichti, & Singer, 2007). Thus, we felt that synaesthetic shape-color associations would affect looking behavior in participants to the greatest degree if we presented the shapes with both congruent and chromatically opponent colors simultaneously. Finally, chromatically opponent colors are far apart in perceptual color space. Thus, their use is expected to minimize the number of possible synaesthetic colors of the triangles and circles that would be perceptually equidistant from the two background colors, which if occurred, would be predicted to yield no effect on color preference.

The luminance levels of the red-green and blue-yellow background pairings were judged by the experimenters to be of approximately equal brightness. The resulting brightness matches produced variations in luminance by hue that were similar to previously observed brightness-to-luminance-ratio patterns (Burns, Smith, Pokorny, & Elsner, 1982; Teller, Civian, & Bronson-Castain, 2004). Although the different colors that adult synaesthetes experience with different graphemes are neither equally luminous nor equally bright (e.g., Beeli, Esslen, & Jänecke, 2007), we nonetheless thought a good starting point would be to present colors that were equally bright. The stimulus values employed are presented in Table 1. Shape type (triangle, circle), background-color pairing (red-green, blue-yellow), and left/right position of the background colors were randomized across trials.

Procedure

Color preferences in infants and adults were measured by an experimenter who could not see the display. The experimenter used a variant of the forced-choice preferential-looking (FPL) technique (e.g. Dobkins et al., 2001) to judge which side of the display the participant looked toward first on each trial. Infants were tested individually in 2 to 3 sessions (15–60 min each) over 2 to 3 days within a single week. Each session lasted until the infant became bored or fussy or until all trials were completed (maximum of 60 trials per shape condition). Adults were tested in the same FPL format as infants were (McDonough, Choi, & Mandler, 2003) and completed the experiment in a single session lasting approximately 30 min. The design of the study was cross-sectional: No subject was included in more than one age group.

Catch trials

Thirty-two catch trials were interspersed throughout the experiment. They consisted of the same stimuli as in the main experiment, with two differences. First, the entire background was a single color (red, green, blue, or yellow), and, second, the shapes (circles or triangles) were presented on only one side of the video monitor (i.e., the other side of the video monitor was colored but contained no shapes).

These catch trials were employed to ensure that participants were paying attention. As noted previously, only data from participants who looked toward the side of the monitor containing the shapes on more than 85% of catch trials were retained for further analysis. For retained participants, averaged across all trial types (i.e., different color backgrounds and different shapes), the mean percentage of correct responses for catch trials was 94.6% for 2-month-olds, 96.8% for 3-month-olds, 93.2% for 8-month-olds, and 92.1% for adults. The fact that these values are all above 90% indicates that participants were engaged in the FPL task.

Analysis

To represent the magnitude of shape influence on color preference, we calculated a synaesthesia index (SI) for each participant. For red-green trials, we subtracted the proportion of trials in which a participant preferred (i.e., looked toward) red (as opposed to green) on the circle trials from the proportion of trials in which that participant preferred red on triangle trials, and then took the absolute value of this difference, as follows: $SI = |p(red)_{triangles} - p(red)_{circles}|$.

The same procedure was repeated for blue-yellow trials. Note that the last step of the SI calculation (i.e., taking the absolute value) necessarily throws out information about the direction of the preference (e.g., whether red was preferred more when triangles were presented than when circles were presented or vice versa). This step was necessary because we assumed that, as is the case for adult synaesthetes, infants have individualized shape-color pairings. That is, half of the infants could prefer red more when presented with triangles than with circles, and the other half could prefer red more when presented with circles than with triangles. In this scenario, if our SI did not have an absolute value, then averaging across individuals would cancel out any true effect. Of course, in taking an absolute value, we transformed our distribution in a way that violated the normality assumptions of most standard statistical tests. Thus, to determine whether the mean of these absolute values was greater than would be predicted by chance, we used a nonparametric analysis, a Monte Carlo simulation.

The Monte Carlo simulation created a distribution of mean SIs that would be expected from chance given no shape influence on color preference, for each age group and color combination tested. First, we created “simulated participants,” with the number of simulated participants matching the actual number of participants for each age group.

Then, we calculated observed color preferences, disregarding shape condition, and used the mean and standard deviation of these color preferences to create a distribution of simulated color preferences that was similar to the observed distribution of color preferences (disregarding shape condition). This was done separately for each age group and color condition (see Average Color Preferences in the Supplemental Material available online for actual values). For each simulated participant, we then drew from this distribution to assign a simulated color preference, disregarding shape condition (e.g., 65% red vs. green preference). Our null hypothesis was that participants would have the same color preference for both triangle and circle conditions, thus the next step was to simulate for each participant a triangle and circle color preference, both drawn from the same distribution.

To accomplish this, for each simulated trial, the model randomly assigned the trial as a circle or triangle trial, and then drew from a binomial distribution defined by each simulated participant's color preference, which assigned that trial as a red or green preference. The resulting shape-condition color preferences, one for each of the circle and triangle conditions, allowed us to calculate individual SIs, and then group mean SIs. This simulation was repeated 100,000 times per each age group and color-pair condition. We then determined whether each observed group mean SI had less than a 5% chance of being drawn from the simulated distribution of group mean SI values.

Results

Figure 2 shows the group mean SI for each age group in each color condition. The results of the Monte Carlo simulation indicated that the group mean SI was greater than would be expected by chance (dashed lines in Fig. 2) in two cases: the red-green condition in 2-month-olds (p = .015) and the blue-yellow condition in 3-month-olds (p = .006). Note that there was also a clear pattern in the data sets; the group mean SI for the red-green condition declined monotonically past 2 months, and for the blue-yellow condition, it peaked at 3 months, declining steadily thereafter.

Note that because the mean number of trials differed somewhat from one age group to another, the mean of the simulated distributions varied somewhat across ages. To address the possibility that the observed age differences were an artifact of trial number, we repeated the analysis equating trial number across participants but using only the first 25 trials per condition from each participant. The pattern of results was very similar to the pattern in our main analysis (group mean SI was greater than would be expected by chance in the red-green condition in 2-month-olds, p = .039, and in the blue-yellow condition in 3-month-olds, p = .014).

We conducted another analysis that determined whether, within an individual participant, there was a significant difference in the proportion of red versus green preference and blue versus yellow preference between the circle and triangle conditions. The findings from this analysis nicely mirror those of our Monte Carlo simulation (for details, see Individual-Level Analysis and Fig. S1 in the Supplemental Material).

Discussion

The results from this study provide the first evidence that fulfills both of the core predictions of the neonatal-synaesthesia hypothesis: They demonstrate synaesthetic-like associations early in life and demonstrate that they decline with age. Furthermore, the observed synaesthetic shape-color associations occurred later in development with blue-yellow than with red-green stimuli. This could result from the fact that blue-yellow sensitivity, mediated by the koniocellular subcortical pathway, develops later than red-green sensitivity, mediated by the parvocellular subcortical pathway (e.g., Crognale et al., 1998; Suttle et al., 2002). As part of its delayed development, the blue-yellow pathway may also retract later from shape-grapheme areas than the pathway underlying red-green processing does. Thus, the time course of shape-color associations may be later for the blue-yellow than for the red-green condition. Furthermore, in the red-green condition, the observed decline in synaesthetic associations observed between 2 and 3 months is consistent with prior evidence that the red-green pathway may retract from other cortical areas (specifically, motion areas) around the same time, between 2 and 4 months (Dobkins & Anderson, 2002).

It is important to point out that the supporting evidence for this conclusion, namely, the age-related decline in shape-color associations, cannot be accounted for by age-related improvements in shape discrimination (Slater, Mattock, & Brown, 1981) or color discrimination (Dobkins et al., 2001); if anything, such improvements would predict an increase in shape-color associations rather than a decline with age. Also note that the study reported here focused on the four unique colors: red, green, blue, and yellow. However, many associations made by adult synaesthetes involve other, nonunique colors (such as purple). In our paradigm, a synaesthetic experience of even a nonunique color will be perceptually more similar to a synaesthetic experience of one of the two colors in our pairs (i.e., more similar to blue than to yellow), and thus we expected our design to be sensitive to some of these nonunique associations. We recognize, of course, that our paradigm may not be sensitive enough to pick up on all nonunique color associations, nor would it be sensitive in circumstances in which an individual makes similar color associations with both shapes (triangles and circles). Thus, the results reported here, if anything, likely show an underestimation of the prevalence of neonatal synaesthesia. Especially in light of these limitations, we believe that our results provide strong evidence for synaesthetic shape-color associations in young infants.

The results reported here join a wealth of previous data demonstrating behavioral abilities declining during infancy, including discrimination of nonnative phonetic categories (Werker & Tees, 1984), nonnative spatial categories (Choi, 2006; Hespos & Spelke, 2004), other-race faces (Kelly et al., 2007), and other-species faces (Pascalis, De Haan, & Nelson, 2002). Age-related declines have also been reported in the ability to match other-species vocalizations to mouth movements (Lewkowicz, Sowinski, & Place, 2008) and use of color information for motion processing (Dobkins & Anderson, 2002). Our results demonstrate that this domain-general developmental learning process, in which extraneous abilities are lost, might also produce arbitrary synaesthetic associations in infancy that dissipate with age.

In sum, our finding of an age-related decline in shape-color associations is consistent with the previously proposed possibility that neonatal synaesthesia is caused by exuberant anatomical connectivity, and that a failure of the retraction process leads in rare cases to synaesthesia in adults (Mondloch & Maurer, 1993; Spector & Maurer, 2009). Furthermore, even in nonsynaesthetes, retraction may be not be 100% complete, leaving them with a weaker form of synaesthesia, as previously reported in typical children and adults (Spector & Maurer, 2008, 2009). In addition to the hypothesis of anatomical exuberance and retraction, another recent suggestion is that synaesthesia results from decreased inhibition from feedback projections (e.g., Eagleman & Goodale, 2009), and thus it is also possible that the age-related decline in shape-color associations observed in the study reported here reflects age-related changes in inhibitory feedback connections (for discussion, see Spector & Maurer, 2009). More information about the time course of anatomical exuberance and retraction (e.g., Huttenlocher et al., 1982) and development of feedback pathways (e.g., Burkhalter, 1993) will be necessary to choose between these possibilities.

On a final note, the presence of widespread synaesthetic associations has important implications for typical development. It suggests that the infant perceptual experience is fundamentally different from that of typical adults, in that young infants experience a “blooming, buzzing confusion” (James, 1890) from intermingling of sensations. Although this intermingling would not appear to be beneficial in and of itself, it may be that exuberance followed by retraction is the most efficient way to form neural connections (and sensory associations) that are ultimately useful (Checkik, Meilijson, & Ruppin, 1999; and see Dobkins, 2009). Thus, although this intermingling might create a confusing experience for infants, this confusion may be dwarfed by the advantage gained in the ability to learn most efficiently about their world.

References

Beeli, G., Esslen, M., & Jänecke, L. (2007). Frequency correlates in grapheme-color synaesthesia. *Psychological Science, 18,* 788–792.

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision, 10,* 433–436.

Brang, D., Rouw, R., Ramachandran, V. S., & Coulson, S. (2011). Similarly shaped letters evoke similar colors in grapheme-color synesthesia. *Neuropsychologia, 49,* 1355–1358.

Burkhalter, A. (1993). Development of forward and feedback connections between areas V1 and V2 in human visual cortex. *Cerebral Cortex, 5,* 476–487.

Burns, S. A., Smith, V. C., Pokorny, J., & Elsner, A. E. (1982). Brightness of equal-luminance lights. *Journal of the Optical Society of America, 72,* 1225–1231.

Checkik, G., Meilijson, I., & Ruppin, E. (1999). Neuronal regulation: A mechanism for synaptic pruning during brain maturation. *Neural Computation, 11,* 2061–2080.

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

Crognale, M., Kelly, J., Weiss, A., & Teller, D. (1998). Development of the spatiochromatic visual evoked potential (VEP): A longitudinal study. *Vision Research, 38,* 3283–3292.

Dobkins, K. R. (2009). Does visual modularity increase over the course of development? *Optometry and Vision Science, 86,* 583–588.

Dobkins, K. R., & Anderson, C. (2002). Color-based motion processing is stronger in infants than in adults. *Psychological Science, 13,* 76–80.

Dobkins, K. R., Anderson, C., & Kelly, J. (2001). Development of psychophysically derived detection contours in L- and M-cone contrast space. *Vision Research, 41,* 791–807.

Eagleman, D. M., & Goodale, M. A. (2009). Why color synesthesia involves more than color. *Trends in Cognitive Sciences, 13,* 288–292.

Hespos, S., & Spelke, E. (2004). Conceptual precursors to language. *Nature, 430,* 453–456.

Huttenlocher, P., de Courten, C., Garey, L., & Van der Loos, H. (1982). Synaptogenesis in human visual cortex: Evidence for synapse elimination during normal development. *Neuroscience Letters, 33,* 247–252.

James, W. (1890). *The principles of psychology. Volume 1* New York, NY: Henry Holt & Co.

Kelly, D., Quinn P., Slater, M., Lee, K., Ge, L., & Pascalis, O. (2007). The other-race effect develops during infancy: Evidence of perceptual narrowing. *Psychological Science, 18,* 1084–1089.

Lewkowicz, D., Sowinski, R., & Place, S. (2008). The decline of cross-species intersensory perception in human infants: Underlying mechanisms and its developmental persistence. *Brain Research, 1242,* 291–302.

Marks, L. E., Hammeal, R. J., & Bornstein, M. H. (1987). Perceiving similarity and comprehending metaphor [Target article and commentary]. *Monographs of the Society for Research in Child Development, 52,* 1–102.

Maurer, D. (1993). Neonatal synesthesia: Implications for the processing of speech and faces. In B. de Boysson-Bardies, S. de Schonen, P. Jusczyk, P. McNeilage, & J. Morton (Eds.), *Developmental neurocognition: Speech and face processing in the first year of life* (pp. 109–124). Dordrecht, The Netherlands: Kluwer Academic Publishers.

McCandliss, B., Cohen, L. & Dahaene, S. (2003). The visual word form area: Expertise for reading in the fusiform gyrus. *Trends in Cognitive Sciences, 7,* 293–298.

McDonough, L., Choi, S., & Mandler, J. (2003). Understanding spatial relations: Flexible infants, lexical adults. *Cognitive Psychology, 46,* 229–259.

Melara, R., & O'Brien, T. (1987). Interaction between synesthetically corresponding dimensions. *Journal of Experimental Psychology: General, 116,* 323–336.

Mills, D., Coffey-Corina, S., & Neville, H. (1997). Language comprehension and cerebral specialization from 13 to 20 months. *Developmental Neuropsychology, 13,* 397–445.

Mondloch, C., & Maurer, D. (2004). Do small white balls squeak? Pitch-object correspondences in young children. *Cognitive, Affective, & Behavioral Neuroscience, 4,* 133–136.

Neville, H. J. (1995). Developmental specificity in neurocognitive development in humans. In M. Gazzaniga (Ed.), *The cognitive neuroscience* (pp. 219–234). Cambridge, MA: MIT Press.

Nikolíc, D., Lichti, P., & Singer, W. (2007). Color opponency in synaesthetic experiences. *Psychological Science, 18,* 481–486.

Pascalis, O., de Haan, M., & Nelson, C. (2002). Is face processing species-specific during the first year of life? *Science, 296,* 1321–1323.

Ramachandran, V., & Hubbard, E. (2001). Synaesthesia: A window into perception, thought and language. *Journal of Consciousness Studies, 8,* 3–34.

Rodman, H., & Consuelos, M. (1994). Cortical projections to anterior inferior temporal cortex in infant macaque monkeys. *Visual Neuroscience, 11,* 119–133.

Rouw, R., & Scholte, H. (2007). Increased structural connectivity in grapheme-color synesthesia. *Nature Neuroscience, 10,* 792–797.

Slater, A., Mattock, A., & Brown, E (1981). Form perception at birth: Cohen & Younger (1984) revisited. *Journal of Experimental Child Psychology, 51,* 395–406.

Smilek, D., Dixon, M., Cudahy, C. & Merikle, P. (2001). Synaesthetic photisms influence visual perception. *Journal of Cognitive Psychology, 13,* 930-6.

Spector, F., & Maurer, D. (2008). The colour of Os: Naturally biased association between shape and colour. *Perception, 37,* 841–847.

Spector, F., & Maurer, D. (2009). Synesthesia: A new approach to understanding the development of perception. *Developmental Psychology, 45,* 175–189.

Sternheim, C. E., & Boynton, M. (1966). Uniqueness of perceived hues investigated with a continuous judgmental technique. *Journal of Experimental Psychology, 72,* 770–776.

Suttle, C., Banks, M., & Graf, E. (2002). FPL and sweep VEP to tritan stimuli in young human infants. *Vision Research, 42,* 2879–2891.

Teller, D., Civian, A., & Bronson-Castain, K. (2004). Infants' spontaneous color preferences are not due to adult-like brightness variations. *Visual Neuroscience, 21,* 397–401.

Walker, P., Bremner, J. G., Mason, U., Spring, J., Mattock, K., Slater, A., & Johnson, S. P. (2010). Preverbal infants' sensitivity to synaesthetic cross-modality correspondences. *Psychological Science, 21,* 21–25.

Ward, J., Huckstep, B., & Tsakanikos, E. (2006). Sound-colour synaesthesia: To what extent does it use cross-modal mechanisms common to us all? *Cortex, 42,* 264–280.

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

Table 1.1: Chromaticity and Luminance Values of the Colors Used in the Experiment

	Chromaticity		
Color	*x*	*y*	Luminance (cd/m^2)
Red	0.612	0.34	8.69
Green	0.287	0.597	13.9
Blue	0.153	0.066	4.94
Yellow	0.457	0.463	15.6

Note: Chromaticity was defined according to the *x* and *y* coordinates on the Commission Internationale de l'Éclairage (CIE) 1931 chromaticity chart.

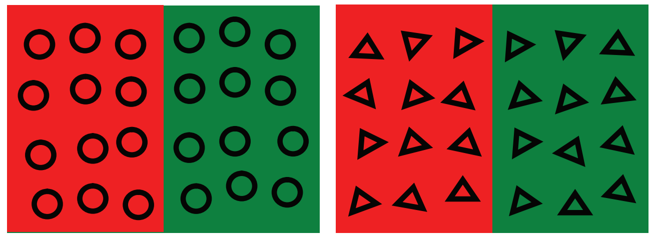

Figure 1.1: Example stimuli used in the experiment. Twenty-four shapes (either circles or triangles) were presented on a background that was either red on one side and green on the other (shown here) or blue on one side and yellow on the other.

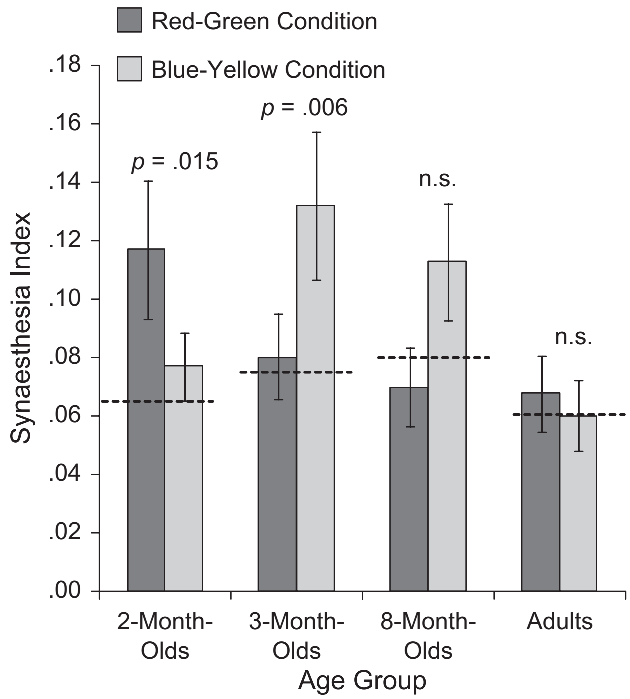

Figure 1.2 Mean synaesthesia index as a function of age group and color condition. Synaesthesia indices were calculated by subtracting the proportion of trials in which a participant preferred red or blue (as opposed to green or yellow, respectively) in the circle condition from the proportion of trials in which that participant preferred red or blue in the triangle condition, and then taking the absolute value of this difference. The dashed lines denote the mean values that would be expected by chance, as determined by a Monte Carlo simulation. Means that are significantly greater than the simulated distributions are indicated. Error bars denote standard errors of the mean.

Chapter 1, in full, is a reprint of the material as it appears in Synesthetic associations decrease during infancy. *Psychological Science, 22*, 1067-1072. Wagner, K., & Dobkins, K. (2011). The dissertation author was the primary investigator and author of this paper. Permission for use of this material is granted through SAGE's Global Journal Author Reuse Policy.

Chapter 2

Slow mapping: Color word learning as a gradual inductive process

Katherine Wagner, Karen Dobkins and David Barner

University of California, San Diego

Corresponding Author

Katherine Wagner

Department of Psychology, 0109

University of California, San Diego

La Jolla, CA 92093

kgwagner@ucsd.edu

Abstract

Most current accounts of color word acquisition propose that the delay between children's first production of color words and adult-like understanding is due to problems abstracting color as a domain of meaning. Here we present evidence against this hypothesis, and show that, from the time children produce color words in a labeling task they use them to represent color. In Experiment 1, an analysis of early color word errors finds that, before acquiring adult-like understanding, children make systematic hypotheses about color word meanings, which are best characterized as overextensions of adult meanings. Using a comprehension task, Experiment 2 finds that these overextensions are due to overly broad color categories, rather than a communicative strategy. These results indicate that the delay between production and adult-like understanding of color words is not due to difficulties abstracting color, but is largely attributable to the problem of determining the color boundaries marked by specific languages.

Slow mapping: Color word learning as a gradual inductive process

Color words like *red, green,* and *blue* pose a difficult problem to children learning language. According to early reports, children at the turn of the 20th century did not acquire the meanings of color words until as late as 8 years of age. Recent reports suggest that children now acquire color words earlier, around 3 or 4 years of age, but nevertheless struggle to learn them (e.g., Backscheider & Shatz, 1993; Sandhofer & Smith, 1999). The primary evidence of children's difficulty is that, similar to the domains of number (Wynn, 1990) and time (Shatz, Tare, Nguyen, & Young, 2010), children typically produce color words well before they use them with adult-like meanings. Also, it's often argued that the use of color words is initially "haphazard and inconsistent" (p.70, Pitchford & Mullen, 2003). By most current accounts, this delay between production and adult-like understanding is caused by a difficulty abstracting color as a dimension of linguistic meaning. Here we challenge this idea and present evidence that children's initial use of color words is in fact systematic rather than haphazard, and that they have abstracted color by the time they begin using color words. We argue that the main source of children's delay is the problem of inferring category boundaries for color words.[1]

When children learn words that describe number, time, space, and color, they typically produce the words, recognize them as belonging to distinct lexical classes, and even use them in response to questions like "What color is this?", well before they acquire their adult-like meanings (for review, see Shatz et al., 2010). In the case of color, the focus of the current study, researchers have argued that the lag between production

[1] Note that the term "category boundary" does not entail the existence of crisp boundaries. Instead, we assume that category boundaries are graded - e.g., as in the cases of gradable adjectives, like big, tall, hairy, etc. (see Barner & Snedeker, 2008).

and adult-like comprehension stems primarily from difficulties abstracting color as a domain of linguistic meaning. Despite quickly learning to produce and associate color words with one another, children, according to some accounts, have difficulty identifying color as the particular dimension of experience that they encode.

In support of this hypothesis, some have argued that color is not as salient to children as other properties like shape, function, and kind, and that they therefore pay less attention to color (for discussion, see O'Hanlon & Roberson, 2006). Others have argued that color may be salient to children, but still more difficult to abstract than other types of content (Kowalksi & Zimiles, 2006; Sandhofer & Smith, 1999). By either account, it is typically assumed that once children have identified color as a domain relevant to word meaning, the mapping of color words to their target color categories proceeds somewhat quickly, and resembles a conceptual epiphany. For example, according to Franklin (2006), "Children seem to struggle with their first color word yet learn most of the other basic terms fairly rapidly over the next several months…. This seems to suggest that there is some kind of 'switch' for children's ability to learn and map color words correctly" (p. 324). On this view, although children have considerable difficulty acquiring the adult meaning of their first color word, once they have done so the mapping of other words to their adult-like meanings is relatively simple and fast (for discussion, see Franklin, 2006; Soja, 1994).

The idea that abstraction is the primary problem of color word learning derives in large part from evidence that infants (Bornstein, Kessen & Weiskopf, 1976; Franklin; Pilling & Davies 2005) and pre-schoolers who do not yet know color words (Bonnardel & Pitchford, 2006; Franklin et al., 2008) possess perceptual color categories like those of adults. According to some, if children have adult-like perceptual color categories when they begin acquisition, then color word learning may reduce to a problem of mapping words to these categories, once color is identified as the relevant domain of meaning. For example, according to Shatz (1996), "on perceptual tasks, infants treat the continuous dimension of hue categorically much as adults do… Thus, the apparent difficulty children have with colour term acquisition cannot be primarily because the perceptual domain is continuous whereas the lexical domain is discontinuous" (p.178). Similarly, according to Pitchford and Mullen (2003), "Developmental studies have shown young children's perceptual colour space is organized in a similar manner to that of the adult… Thus, when children engage in the learning of colour terms, they already possess colour percepts on which colour concepts can be mapped." (p.53) And according to Franklin (2006), "A common theme in explaining children's difficulty in color naming is the idea that children find it difficult to learn color names because they need to learn the boundaries of colors… [However] perceptual categories are in place even at 4 months of age" (p. 324-325). The implication of such arguments is that, because infants have color categories prelinguistically, the lag between production and adult-like understanding must not be due to the problem of determining boundaries of individual color words. Instead, the delay must be due to the prior problem of identifying color as a domain of linguistic meaning.

However, pre-linguistic color categories notwithstanding, there are good reasons to believe that the acquisition of color words is not a simple mapping problem. Perhaps most important is evidence that languages differ substantially in how they carve up perceptual color space. Languages vary both in the number of basic color words they have (from 2 to 12) and also in how these particular words divide up color space. According to the World Color Survey (Kay et al., 2009), languages that feature only two or three color words organize perceptual space in a way unlike languages with more color words, frequently grouping warm colors (e.g., white, red, yellow) under one label and cool colors (e.g., black, green, blue) under another. For example, two of the five color categories used in Berinmo, a tribal language spoken in Papa New Guinea, are *nol* (green, blue and purple) and *wor* (green, yellow, orange and brown). Thus, Berinmo marks at least one color boundary that is missing in English, while failing to mark other boundaries that are found in English (Roberson, Davidoff, Davies & Shapiro, 2005). Differences like these are not explained purely by differences in the number of color words a language provides. Some languages that have four basic color terms mark a category boundary between red and yellow (e.g., Culina, spoken in Peru; Waorini, spoken in Ecuador) whereas others do not (e.g., Chácobo, spoken in Bolivia; Múra-Pirahã, spoken in Brazil; Kay et al., 2009). Also, although Russian, Korean and English have roughly the same number of basic color terms (11-12; Berlin & Kay, 1969), each language divides the blue-green region of color space differently (e.g. Roberson, Hanley & Pak, 2009; Winawer, Witthoft, Frank, Wu, & Boroditsky, 2007). In sum, while infants may perceive color like adults, and surely use prelinguistic categories as inputs to learning, the categories encoded by language are not fully determined by perception. This gap between perception and language suggests that inductive learning – i.e., a process of constructing categories from experience with a subset of possible exemplars – must play a significant role in color word learning. However, despite being the focus of research in most other domains of word learning since Quine (1960), the role of inductive learning has received little attention in the color word learning literature.

In the current study, we explored the idea that linguistic color categories are constructed via a gradual inductive process. We argue that this gradual process, and not the problem of abstracting color as a domain, is the primary cause of delay between the onset of color word production and adult-like understanding. Our suggestion is that children acquire preliminary meanings for color words well before they converge on adult-like meanings, and thus abstract the domain of color much earlier in the acquisition process than typically thought. On this hypothesis, a significant component of the delay between color word production and adult-like understanding is due to a "slow-mapping" process, whereby children gradually determine the language-specific boundaries of color words. This process was first described by Carey and Bartlett (1978), in their influential paper on "fast-mapping" – the process of associating a label with a particular referent in a single learning trial. In their paper, they argued that "fast mapping is only a small fraction of the total information that will constitute a full learning of the words" and that the "fast-mapping" process is followed by a "second phase, the long drawn out mapping" (p. 2). The basis for this idea was that, although some children in their study could map a recently learned color word to its original referent, children were also prone to overextend the word, or to confuse it with other words that labeled proximal colors. However, because their study (and another study by Bartlett, 1978) focused on a small group of children who had, for the most part, already acquired at least one adult-like color word meaning (e.g., using *red* to exclusively label red objects), it left open when the slow mapping process begins, and thus whether it can account for the long delay between children's onset of color word production and acquiring adult-like meanings of color words.

More generally, past studies have typically failed to address the nature of the mapping problem because of how they have characterized children's color word meanings. For example, researchers have often focused on whether children have knowledge of adult-like meanings – e.g., using *red* to exclusively label red objects – without testing whether children first acquire preliminary, non-adult-like meanings (e.g., Kowalski & Zimiles 2006; Soja, 1994; O'Hanlon & Roberson, 2006). In doing so, such studies likely underestimate children's early knowledge of color words, and thus the point at which they first abstract color. Further, this method may create the appearance that children acquire all color words at once. If the meanings of color words are mutually constraining, they may all become adult-like in synchrony despite evolving gradually over the course of many months in development.

Consistent with this concern, a number of studies have found that before children acquire all 11 adult-like color word meanings, they make errors that are systematic in nature (Pitchford & Mullen, 2003; Davies, Corbett, McGurk & MacDermid, 1998; Bartlett, 1978). For example, Pitchford and Mullen found that before mastering the adult-like meanings of the 11 basic color terms, 3-year-olds often use their color words to incorrectly label hues adjacent to the target category (e.g., labeling orange as *red*). On the basis of this evidence, they argued that pre-linguistic perceptual categories strongly constrain early color word meanings. However, these errors are not easily explained by this hypothesis, since it predicts that word boundaries should be determined by non-linguistic perceptual categories, and thus should not extend beyond them. Instead such errors most strongly support the existence of linguistic categories that are *broader* than those used by adults, and thus that are not acquired purely on the basis of pre-defined perceptual categories.

In the current study, we investigated the first meanings that children assign to color words by analyzing the errors they make in both language production and comprehension. Although some past studies have reported error data in early color word use (see above), here we present evidence and analyses not reported in past studies, which directly address the nature of the delay between production and adult-like understanding, and which lead us to draw different conclusions. In Experiment 1, we present data from a color-labeling task, sampled from a large group of children including a subset who have not yet acquired any adult-like meanings. This experiment finds that children make errors that are systematic in nature prior to using any color words in an adult-like manner. These data suggest that children in our study have abstracted color and possess partial knowledge of color words shortly after they begin producing them and possibly before. Furthermore, we show that children acquire color word boundaries via a gradual inductive process: learners begin with overly broad meanings for their earliest color words and gradually narrow these meanings as they add new words to their vocabularies. In Experiment 2, we corroborate these findings using a language comprehension task, and show that children's early overextension of color words reflects overly broad meanings, rather than a communicative strategy.

Experiment 1

Methods

Participants. A total of 141 children (68 girls) participated. Children with a 25% chance or higher of protanopia or deuteranopia color deficiency (based on family history) were eliminated from analysis (n=5). An additional 21 were excluded because they made no errors (mean age=3;5, sd=7.2m), 6 were excluded because they used only one color term during the experiment (mean age=2;6, sd=4.2m), and 17 were excluded because they did not cooperate on half or more than half of the trials (mean=2;7; sd=7.2m). Data from the remaining 98 children (50 girls) were retained for analysis. These children were between the ages of 22 months to 61 months of age (mean= 3;0).

Stimuli. Stimuli were constructed using 11 pieces of colored posterboard, which were chosen by a consensus of five experimenters as being prototypical of the 11 basic color terms in English (i.e., red, orange, yellow, green, blue, purple, pink, brown, black & gray). The CIELAB values of each color were measured using a Photo Research-650 Spectascan under natural sunlight, see Table 1.

The posterboard was cut into a set of 11 fish shapes (Fish Task) and a set of 11 squares (Book Task). For the Fish Task the colored fish were glued to black foam (also cut into fish shapes) and were presented on a black background. For the Book Task, the colored squares were glued onto black pages and covered with white flaps of various shapes.

Procedure

Fish Task. Each child was presented with a black box containing the 11 colored fish, placed color-side down. The experimenter began the task by announcing, “My turn!” and randomly picked up one of the fish randomly asking, “What color is it?” After the child responded, the experimenter placed the labeled fish on the table and told the child, “Your turn!”, indicating that the child should pick up a fish and label it. The experimenter and the child continued taking turns until each fish had been selected and labeled.

Book Task. Following the Fish Task, the experimenter presented the child with a book that contained the colored squares. For each page, the child lifted the flap that covered the color and the experimenter asked, “What color is it?” Colors were presented in the following order, to avoid hue-based groups of items: orange, blue, yellow, pink, white, purple, gray, brown, green, red, black.

When children did not respond on a particular trial, the experimenter repeated the question and gave the child another chance to respond. Trials with no response (103 trials, 4.7%) or with two responses (e.g., the child said both *blue* and *red*, 13 trials, 0.05%) were not analyzed.

Results

Color-Knowledge Groups. To investigate the delay between color word production and adult-like usage, it was necessary to separate children by color-word knowledge rather than biological age, particularly because the age at which children acquire color words is highly variable (e.g., it is accelerated in children who attend preschool, Shatz et al., 1996). Children were separated into four groups based on the number of Basic Color Terms they used in an adult-like manner (e.g., using *red* consistently and exclusively to label red stimuli). We were particularly interested in examining errors in children who had demonstrated adult-like knowledge of no terms, and accordingly placed children with zero color terms in their own group. The remainder of the children had adult knowledge of between 1 and 9 basic color terms. In order to simplify the analysis and increase power, we collapsed the remaining children into three groups with 1-3; 4-6 and 7-9 color terms respectively. Examples of children from each of the 4 levels are presented in Figure 1. Note that in each group the average number of terms produced by children exceeded the average number of terms used with adult-like meanings.

Level 1: Adult-like knowledge of 0 color terms. These children spontaneously produced an average of 3.1 color terms (range 2 to 6) during the experiment and had a mean age of 2;5, sd=4.1m (n=8, 1 girl).

Level 2: Adult-like knowledge of 1-3 color terms (mean=2.0). These children spontaneously produced an average of 6.64 color terms (range 3 to 9) during the experiment and had a mean age of 2;7, sd=5.6m (n=16, 5 girls).

Level 3: Adult-like knowledge of 4-6 color terms (mean=5.1) These children spontaneously produced an average of 9.05 color terms (range 8 to 10) during the experiment and had a mean age of 3;2, sd=7.0m (n=19, 9 girls).

Level 4: Adult-like knowledge of 7-9 color terms (mean=8.2). These children spontaneously produced an average of 10.28 color terms (range 9 to 12) during the experiment and had a mean age of 3;2, sd=7.0m (n=53, 34 girls).

Error Consistency Analysis. This analysis tested the consistency of children's errors: Given that a child used an incorrect label for a particular stimulus color on one task (using *red* to label the orange stimulus on the Fish task), we asked how likely it was for the child to repeat the error on the other task (using *red* to label orange on the book task). For example, participant 3A in Figure 1 consistently labeled gray as *white* twice but was inconsistent in labeling pink, which she labeled as *pink* in one task but as *purple* in the other. Using a binomial test, we asked whether the proportion of consistent trial pairs was greater than expected by chance.

For this analysis, we did not include trial pairs in which the participant labeled the stimulus correctly on both tasks (725 pairs). The remaining 275 pairs were classified as being either consistent (the same incorrect label for a given stimulus color on both tasks, 122 pairs) or inconsistent (153 pairs). The 153 inconsistent pairs were either two different incorrect labels (62 pairs), or one incorrect and one correct label (91 pairs).

The probability of repeating any single label on two separate trials was defined as the square of the child's base rate use of that term, with base rate defined as the proportion of total trials a child used that label. For example, if a child used *red* (either correctly or incorrectly) on 6 out of 22 trials, that child's probability of using *red* incorrectly on both the orange fish trial and the orange book trial was $(6/22)^2$. A child's overall chance probability of consistency was defined as the sum of the probability of repeating each of the different labels that they used. Each color knowledge group's overall chance probability of consistency was defined as the average of the individual participant probabilities, weighted by the number of data points each individual contributed to the analysis. In other words:

$$p(consistency) = \sum_c \sum_j (\frac{l_j}{n})^2 (\frac{i_c}{i})$$

where i is the total number of stimulus pairs in which at least one label (either book or fish) was incorrect, i_c is the number of such incorrect pairs that each child, c, contributed to the analysis, l_j is the number of times a child produced each label j and n is the total number of responses a child produced. Note that by this definition, the chance probability of consistency appropriately decreases as a child adds more color words to his/her lexicon.

Averaged across the different Color-Knowledge groups, the proportion of incorrect trial pairs that was consistent (0.44) was greater than would be expected by chance (0.28), using a binomial test, $p<0.001$ (see Figure 2). When this analysis was conducted separately for the different Color-Knowledge groups, rates of consistency were greater than chance for the Level 2 ($p=0.013$), Level 3 ($p<0.001$) and Level 4 ($p<0.001$). The rate for the Level 1 group was not above chance ($p=0.27$).

The consistency analysis indicates that, with the exception of the Color-Knowledge group that exhibited no adult-like meanings of color words (Level 1), children's color labeling errors were highly consistent, despite the differences in the stimulus shapes between the first and second tasks. This finding is consistent with the idea that children formulate interim meanings before converging on adult-like understanding of color words. However, a test of consistency gives little information regarding the nature of the hypotheses that children are entertaining when they make errors. Also, consistency is only one measure of whether children have formulated partial meanings for their early color words, and may be overly conservative. For example, some children may correctly and consistently use the word *red* to label a red stimulus (which would not enter into the above consistency analysis, because it considered only errors), but overextend the word *red* in an inconsistent way – e.g., labeling red as *red,* but also sometimes labeling orange as *red*. To address these issues we conducted two additional analyses beginning with an overextension analysis.

Overextension Analysis. This analysis asked whether, in some cases, children's color errors were overextensions of adult color categories. For example, a child may correctly know that *red* refers to red objects, yet have a broader meaning for *red* than adults, and therefore overextend it to orange and yellow objects. Given that a child used a label incorrectly for at least one of the twenty-two trials (e.g., *red* to label orange and yellow stimuli on the Fish task), we asked whether they also used that label correctly and consistently for its target color. For example, when a child used the word *red* to label orange and yellow, we asked whether they also used *red* to label red stimuli. Based on these criteria, Participant 1B's use of *blue* to label gray counted as overextension (see Figure 1), since they also used *blue* to label blue, whereas Participant 1A's use of *red* to label pink, orange, yellow, white, gray and brown did not because they did not use *red* to label the red (but instead labeled it as *blue*).

For the overextension analysis, we judged that a child used a term correctly for its target hue only if they did so on both the Fish and Book Task. If the child did not produce responses for the target hue on both tasks, that color word was not included in the analysis (19 incidences). Using a binomial test, we asked whether the proportion of errors that reflect overextensions was greater than chance. As noted in the consistency analysis (above), the probability of repeating any single label on two separate trials is the square of the child's base rate use of that term, with base rate defined as the proportion of total trials a child used that label. In contrast to the consistency analysis in which consistent use of any incorrect label to any color stimulus was sufficient, in order for an incorrectly applied *red* label to be classified as an overextension a child must specifically use *red* (not any other color) in response to the red stimulus in both tasks. To calculate chance for this analysis, we first squared the base rates of every term a child used incorrectly (e.g., using *red* for purple) to calculate the probability that each of these incorrect terms would also be used on both trials containing the correct color stimulus (e.g., red fish and red book trial). We then took the mean of these probabilities to calculate the child's overall probability of overextension. To calculate the overall probability of overextension for each group, we calculated a group mean, weighted by how many labels each child used incorrectly. In other words:

$$p(overextension) = \sum_{c} \frac{i_c}{i} \frac{\sum_j (\frac{i_{cj}}{n})^2}{i_c}$$

where i is the total number of labels that were used incorrectly at least once, i_c is the number of such incorrect labels that each child, *c,* contributed to the analysis, i_{cj} is the number of times a child produced each incorrect label j, and n is the total number of responses a child produced.

A surprisingly large proportion of children's errors – 0.76 – were overextensions, which was significantly greater than expected by chance (chance = 0.054), as measured by a binomial test (p<0.001). This high proportion indicates that if a child produced a color word, they were very likely to use it correctly when presented with its target hue. Thus, it suggests that most of children's errors were overextensions of color terms that were anchored to adult-like focal hues (i.e., hues that are the best examples of English-speaking adults' categories). Critically, rates of overextension were statistically greater than chance and above 0.72 for each Color Knowledge Group (all ps<0.001, see Figure 2), including children who had no adult-like color meanings (Level 1). This is important, because it shows that before children acquire any adult-like meanings, almost all words they produce are assigned a partial meaning. This suggests that there is likely very little lag between children's use of color words as labels and their acquisition of partial knowledge of these words. Consequently, the delay between children's production of color words and their acquisition of adult-like meanings is likely not due to problems with abstraction. Instead, the delay appears to be due to the problem of determining the boundaries of linguistic color categories.

Proximity Analysis. The overextension analysis indicates that before children acquire adult-like color word meanings, they nonetheless use color words correctly for their target hues, consistent with overextension. However, the analyses described so far do not address the nature of children's overextensions, and whether they were made to proximal colors or to distant ones (see Figure 3). Critically, although overextension to distant hues would be consistent with a gradual inductive process (e.g., since children's initial categories may be very large), the inclusion of proximal hues should be significantly more likely even in this case, since any category that includes red and yellow, for example, should also include intervening hues like orange.

To assess this prediction, we conducted a proximity analysis. Specifically, given that a child used a color label incorrectly, we asked whether the label and its referent stimulus were from perceptually proximal color categories (see Figure 3). For example, Participant 1B made a proximal error by labeling black as *brown,* but also labeled gray as *blue,* a non-proximal error. Using a binomial test, we asked whether the proportion of errors that were from proximal color categories was greater than expected by chance. Chance was defined using both the frequency with which children made errors for each stimulus and the frequency with which they used each label incorrectly. It was necessary to account for these base rates because some color words are proximal to a greater number of color categories than others. For example, *red* is considered proximal to *orange, pink, purple,* and *brown*, while *blue* is proximal to *green* and *purple.* To determine chance, we calculated the probability of each label-stimulus error pair (the probability of using *red* to label an orange stimulus) as equal to the product of the base rates. For example, if 20% (0.2) of errors were in response to an orange stimulus and 80% (0.8) of errors involved using the label *red*, then the probability of using *red* to label orange would be 0.2*0.8, or 0.16.

To determine the overall chance probability of proximal errors, we summed across the probability of all label-stimulus pairs that are classified as proximal. In other words:

$$p(proximity) = \sum_i \sum_j p(r|l_j \cap s_i)p(l_j|incorrect)p(s_i|incorrect)$$

where, $p(s_i|incorrect)$ is the probability of a particular stimulus i given an incorrect response; $p(l_j|incorrect)$ is the probability of a particular elicited label j given an incorrect response to stimulus i; and r is the probability of proximity. Note that $p(r|l_j \cap s_i)$ is either 1 or 0 because a given label/stimulus pair is either proximal or not proximal.

No differences between the Fish and Book tasks were found in preliminary analyses, and thus data for the two tasks were combined. The proportion of total errors that were to proximal categories was 0.43. This was significantly greater than the rate predicted by chance (0.24), $p<0.001$. Rates of proximity were statistically greater than chance for all Color-Knowledge groups, including Level 1 children who had no adult-like color word meanings (Level 1 $p=0.019$, all other ps<0.001 see Figure 2). Like the findings of the overextension analysis, this indicates that even before a child acquires the adult meanings of any color terms, they already have partial knowledge of some color words.[2]

Discussion

Experiment 1 examined children's color word production errors in early acquisition. Our results revealed that children made highly systematic errors that were predictable from the color properties of the stimuli. Also, we found that if children used a word in the study, they were very likely to have a systematic meaning for the word, despite the fact that these meanings were often non-adult-like in nature. Together, these results suggests that (1) children have abstracted color as soon as they begin using color words to label stimuli, and thus well before they acquire their adult-like meanings, and (2) they learn color words by making overly broad hypotheses about their meanings, and gradually narrowing these meanings as they acquire additional, contrasting words.

Several pieces of evidence support these conclusions. First, in our Error Consistency Analysis we found that all but the Level 1 children were highly consistent in their errors, demonstrating that these children were able to abstract color across different objects despite their other differences, and use this knowledge to formulate hypotheses about color word meanings. These results suggest that children begin acquiring meanings much earlier than previously supposed, and that color word learning is a gradual inductive process. Also, although the Level 1 children's errors were not consistent, their

---

[2] We also separately analyzed proximity errors that involved achromatic colors (i.e., black, white, & gray) and those that involved chromatic colors (i.e., all remaining colors). These analyses revealed an early focus on the chromatic colors and a later focus on the achromatic colors. Children had the most difficulty with gray, frequently referring to it as *white* or *black* through level 4, a finding that corroborates previous reports (e.g., Pitchford & Mullen, 2005; Bonnardel & Pitchford, 2006).

use of early color words was nonetheless highly systematic, as shown by our two other analyses.

In our Overextension Analysis we found that, at all levels of Color-Knowledge, a significant majority of children's errors were overextensions, e.g., correctly using the label *red* to label a red stimulus, but also using that label for other colors, like orange and yellow. This indicates that children have partial knowledge of the specific color properties denoted by a color word when they first begin producing that color word. Specifically, children appear to know the focal color denoted by the color words they use, though they frequently overextend these words.

Finally, our Proximity Analysis revealed that the errors made by children at all levels of Color-Knowledge were likely to be labels for perceptually similar colors. Our results contrast with reports by Pitchford and Mullen (2003) and Bartlett (1978), who found that color word errors were random in 2-year-olds or children who knew 6 or fewer color words, respectively, and that errors only became proximal in more advanced children. This discrepancy is likely explained by the difference in how each study calculated chance responding. Both studies directly compared the number of proximal errors to the number of non-proximal errors. However, random responding predicts that proximal errors should be much less frequent than non-proximal errors, as there are many more non-proximal color pairs than there are proximal color pairs. Thus, their analysis was overly conservative, and likely underestimated knowledge significantly. In contrast, the current study computed chance using base rates that took account of the probability of proximal errors for each individual color term. In fact, even our method likely underestimated the degree to which children's overextensions reflected coherent, broad categories. This is because if a child used *green* to refer to green, yellow, orange, red and pink, although this is a broad and continuous category, only the most proximal of these errors would be coded as proximal (see participant 1B in Figure 1). Thus, errors that reflect very broad categories were not captured by our analysis, especially for younger children who had fewer adult-like meanings, and who would therefore be more likely to have very broad categories.[3]

In sum, the data from Experiment 1 demonstrate that children with adult-like understanding of no color words have nonetheless abstracted color. These data are consistent with the idea that children begin acquisition of color words by positing overly broad color categories and that these categories are gradually narrowed as children gain experience and acquire other color words that contrast in meaning. We refer to this as the "Broad Color Categories" hypothesis. Another possibility, however, is that overextension errors are in fact not evidence for overly broad meanings, but instead reflect a pragmatic strategy (for a similar discussion in the domain of nouns, see Clark, 1978). On this view, children in Experiment 1 may have possessed adult-like meanings for overextended words, but used them to refer to proximal colors because they lacked knowledge of the correct labels. For example, imagine a child who has an adult-like meaning for *red* but not for orange. When presented with an orange stimulus, the child may recognize that this color is not *red*, but use the word *red* to describe it nonetheless since no better word is available to them. We refer to this as the "Communicative Strategy" hypothesis.

[3] In keeping with this, 25% rate of proximity for Level 1 errors is a surprisingly high number, since these children had no adult-like color meanings, and produced as few as two color words during the task to label all stimuli.

One way of disambiguating between the Broad Color Categories and Communicative Strategy hypotheses is to test children using a comprehension task, where the experimenter selects the label, thereby removing the possibility of overextension as a communicative strategy (e.g., Clark, 1978, Gelman, Croft, Fu, Clausner & Gottfried, 1998). Accordingly, we conducted a comprehension task that asked children to pick the colored stimulus that matched a label produced by the experimenter. We reasoned that, if a child possesses a broad definition for the category *red* (that includes both red and orange), when asked for *red* they should provide either a red stimulus or another stimulus that satisfies their meaning of *red* (e.g., an orange one). By contrast, if the child has adult-like color categories, but only produces the linguistic label *red* and not *orange*, then when asked for *red* they should always prefer to provide a red stimulus over an orange one (even if they use *red* to label orange in language production, as a communicative strategy).

Experiment 2

In Experiment 2 we presented children with stimuli identical to those used in the Fish task in Experiment 1 and asked them to find fish of different colors. To differentiate the Broad Categories hypothesis from the Communicative Strategy hypothesis, we asked whether children made proximity errors. Note that by the Communicative Strategy hypothesis, proximal errors should not occur because responding should either be correct for known color words (above example, *red*) or random for unknown color words (above example, *orange*) because children do not possess broad categories (e.g., that include both red and orange). Proximal errors are only expected under the Broad Categories Hypothesis, since it claims that children possess linguistic categories that include multiple adult categories (e.g. *red* including both red and orange). Thus, in Experiment 2 we conducted a proximity analysis, as we did in Experiment 1.[4]

Methods

Stimuli. Stimuli were identical to those used in the Fish Task in Experiment 1.

Participants. A total of 28 children (14 girls) participated. Participants were screened for color deficiency via a family history questionnaire. Eight children were excluded because they made no errors (mean=3;4; sd=4.6m). Data from the remaining 20 children (8 girls) were analyzed. These children ranged in age from 23 to 48 months (mean= 2;10, sd=6.8m). Unlike in Experiment 1, we did not group children into different Color-Knowledge groups, for two reasons. First, the children in this study were not asked to produce color labels, and we therefore could not determine how many basic color terms they knew. Second, the number of subjects required to test the main hypothesis of this experiment was relatively small, and thus there was insufficient power to analyze subgroups.

Procedure. Children were presented with the fish stimuli placed color-side up and in a random configuration. In succession, the experimenter asked the child to hand her a specific colored fish, "Give me a (red) fish. Can you put a (red) fish in my hand?" After the child handed a fish to the experimenter, it was returned to its place on the table (back with the other fish), and the experimenter requested the next color fish. The experimenter requested the colors in the following order, as in Experiment 1: red, brown, green, orange, white, blue, gray, pink, black, yellow, and purple.

[4] Note that an overextension analysis is not possible for a comprehension task, and also that neither hypothesis predicts consistency in responses (e.g., a child with broad categories may sometimes choose red, sometimes, pink, sometimes purple in their comprehension of *red,* despite consistently labeling all three as *red* in production). Therefore we analyzed only proximity.

If the child did not respond on a particular trial (e.g., they got distracted), the experimenter repeated the question, giving the child an additional opportunity to respond. Trials with no response (n = 3 trials) or on which two or more fish were provided (n = 1 trial) were not included in the analysis.

Results and Discussion

Of the 216 trials collected from the 20 participants, 79 trials (36%) were errors and were included in the analysis. The mean number of correct responses was 6.85 (range 0 to 10).

Like in Experiment 1, we accounted for the base rate of errors that involved each color stimulus. However because Experiment 2 was a comprehension task rather than a production task, our calculation of chance accounted for the base rate of errors that involved each stimulus and the base rate of errors made to a particular request (e.g., *red*) rather than using the base rate of errors that involved each stimulus and the base rates of incorrect labels produced by the child. The analysis was identical to the proximity analysis of Experiment 1 in all other respects.

Consistent with the results from the production task in Experiment 1, the proportion of total errors that were proximal in the comprehension task was 0.58. This was significantly greater than the rate predicted by chance (0.30), $p<0.001$. This pattern was found both in the first six trials (61%; $p<0.01$) and the final five trials (55%; $p<0.01$), confirming that the effect was not driven by children's initial unfamiliarity with the options available to them (e.g., a tendency to choose orange instead of red, because the child failed to notice the red fish). Overall, the results are consistent with the use of broad color categories, rather than a communicative strategy.

It is worth noting that these results are somewhat surprising even on the Broad Color Categories hypothesis. We might have expected, for example, that if children begin with broad categories they should nonetheless treat adult focal hues as central to their broad categories (e.g., treating red as a better exemplar for *red* than orange). What we found, however, is that children often did not exhibit such a preference, raising the possibility that their early linguistic categories either lack a clear focal point, or have a focal point that differs from that of adults (for discussion of this point in the noun literature, see Naigles & Gelman, 1995; Kuczaj, 1982).

General Discussion

We tested the hypothesis that, early in acquisition, the delay between color word production and acquisition of adult-like meanings is due to the gradual construction of linguistic color categories, rather than the process of abstracting color as a domain. Consistent with this idea, we found that if children used color words in our study, they typically used them in a meaningful and consistent way. When children made production errors, the vast majority of these errors (75%) were overextensions of adult-like categories. Also, these overextensions were frequently to proximal hues. This was true for children at all levels of color word competence – even those who had no adult-like meanings. Further, the results of a comprehension task corroborated this hypothesis, and indicated that overextension is not the product of a communicative strategy, but instead reflects broad linguistic color categories.

The results of this study have important implications for our understanding of color word acquisition. First, contrary to previous reports (Bartlett, 1978; Pitchford & Mullen, 2003), the results suggest that children abstract color at very early stages of color word production, and that there is little, if any, lag between children's use of color words as labels and their construction of preliminary meanings for these words. Although abstraction may pose a significant problem to children early in acquisition, this problem is likely resolved by the time children begin using colors to label things in their environment (or shortly thereafter). While children may use some labels prior to constructing preliminary meanings, our overextension analysis indicates that haphazard use accounts for only a small proportion (less then 25%) of children's errors. Second, our results suggest that the observed delay between production and adult-like understanding of color words is likely due to the problem of constructing language-specific category boundaries. Our data indicate that children begin acquisition by making overly broad inductive inferences regarding the scope of their color words, and that they gradually shrink their early categories as they gain experience with the words, and as they acquire other color words that contrast in meaning. Consistent with these findings, research in older children suggests that the refining of linguistic color categories may continue for several years after children are able to correctly label the focal regions of all eleven basic color terms (Raskin, Maital & Bornstein, 1983; Mervis, Catlin & Rosch, 1975).

These data are also consistent with earlier data from Carey and Bartlett (1978), which are most commonly cited as evidence for children's ability to "fast map" color words to their referents. As discussed in the introduction, Bartlett and Carey's fast mapping proposal, unlike some theories of color word learning that followed it, did not assume that learning color word meanings was fast, or that it was a simple mapping problem (see Carey, 2010). Instead, they argued that fast mapping was a first step in the learning process, used to link labels to particular referents, and that acquiring the adult-like meanings of color words likely involved much more additional learning (for similar views on the role of fast mapping in acquisition, see Saji et al., 2011; Swingley, 2010; Clark, 1997). Consistent with this, many of the children in Carey and Bartlett's study used the novel word *chromium*, which was used by experimenters to refer to an olive-colored stimulus, to refer to perceptually similar colors (e.g., green, brown) and often did not converge on the intended narrower meaning for chromium even after many trials. Similarly, in her longitudinal study, Bartlett (1978) found that children often made proximity errors when learning color words. Also, she found that the transition from knowing one adult-like meaning to knowing all 11 took as long as 6 months. Pitchford and Mullen (2003) report similar results, and report a lag of up to 9 months between children's first adult-like color word meanings and their mastery of all 11.

These results, like ours, suggest that color word learning is a gradual inductive process, and that children form interim meanings for their color words well before they attain adult-like understanding. However, because these early studies used statistical methods that underestimated children's knowledge or focused on small samples of children (~20-50 children) who had, for the most part, acquired at least one adult-like color word, their data do not address the nature of the delay between onset of color word production and children's first adult-like color word meaning. In contrast, our study addressed this question using data from a wider range of children (including those who had not acquired any adult-like meanings of color words), and using a novel set of analyses that tested not only proximity errors (as in the Bartlett study, and many other studies) but also consistency and overextension. Consequently, our study arrives at a different conclusion. Our study shows that children possess broad, overextended linguistic color categories early in acquisition, before they have acquired their first adult-like meaning, and perhaps even from the time they first begin producing color words.

The idea that problems with category formation explain the protracted course of color word acquisition would appear, at first pass, to conflict with reports that children have difficulty abstracting color when presented with novel objects (e.g., Sandhofer & Smith, 1999; Kowalki & Zimiles, 2006). For example, in their study, Sandhofer and Smith (1999) showed children different shapes of the same color and asked the children to find other shapes that matched. They found that children were unable to succeed on this task until after they were able to successfully label colors. However, such results are mixed (see Soja, 1994, for a critical discussion) and the tasks used in these studies differ critically from tasks used to specifically probe the meanings of color labels. For example, in the Sandhofer and Smith study, the experimenters specifically avoided reference to color (see also Kowalksi & Zimiles, 2006). In other studies, the tasks involved a memory component (Kowalksi & Zimiles, 2006). While these studies interestingly demonstrate that children do not *preferentially* attend to color early in development, our results suggest that children nonetheless *can* attend to and encode color from an early age, if color is referred to and highlighted linguistically. Although abstracting color as a domain is surely a difficult problem that children must solve before learning color words, the sum of existing evidence suggests that this problem is solved relatively early in acquisition, and probably before children begin using color words as labels.

The view of color word learning proposed here is consistent with findings in other domains of language and conceptual development where children face a similar problem of identifying a domain and then acquiring individual meanings within it (see Carey, 2010). For example, in the case of number, infants also begin acquisition with non-linguistic representations of objects and approximate number and quickly recognize that numerals form a class of words that contrast in meaning (Wynn, 1992; Tare et al., 2008; Brooks, Audet, & Barner, 2012), despite taking years to learn what these meanings are (see Carey, 2009, for review). Similarly, relatively early in acquisition, children recognize that time words like *minute, second,* and *hour* form a lexical class, but take many years to acquire their individual meanings (Shatz et al., 2010; Busby Grant & Suddendorf, 2011). Finally, children produce words that describe emotions from early in development, and understand that they belong to a class of words that describe human sentiment, but nonetheless take years to master their adult-like meanings, and form many interim hypotheses along the way (Widen & Russell, 2003).

Our study suggests that the case study of color is not an exception to this general pattern. In the cases of color, number, space, and time, human infants have access to non-linguistic representations prior to learning language, but nonetheless struggle significantly to learn words that label these representations. This divide – between pre-linguistic competence and early word learning – is a serious puzzle that confronts the study of language learning as a whole (for discussion related specifically to color, see Bonnardel & Pitchford, 2006; Franklin, Clifford, Williamson & Davies, 2005; Dedrick, 1996; Dedrick, 1997). Language allows humans the freedom to go beyond perceptual data, and to select, highlight, and possibly to enrich aspects of the world that are most interesting or important to our interactions. To achieve this expressive power, children must solve the problem of determining the particular linguistic mappings that exist in a language. Our suggestion is that, in the case of color, as in other domains of lexical learning, children solve this problem via a gradual inductive process, aided by prelinguistic categories.

References

Backscheider, A. G. & Shatz, M. (1993). Children's acquisition of the lexical domain of color. In K. Beals *et al.* (eds), *What we think, what we mean, and how we say it,* Papers from the parasession on the correspondence of conceptual, semantic and grammatical representations, CLS 29, v. 2. Chicago: The Chicago Linguistic Society.

Barner, D., & Snedeker, J. (2008). Compositionality and statistics in adjective acquisition: 4-year-olds interpret tall and short based on the size distributions of novel noun referents. *Child Development, 79*, 594-608.

Bartlett, E. J. (1978). The acquisition of the meaning of color terms: A study of lexical development. In R. N. Campbell & P. T. Smith (Eds.), Recent Advances in the Psychology of Language (pp. 89–108). New York: Plenum.

Berlin, B. & Kay, P. (1969). *Basic Color Terms: Their Universality and Evolution*. Berkeley, CA: University of California Press.

Bonnardel, V. & Pitchford, N. J. (2006). Colour categorisation in preschoolers. *Progress in Colour Studies, Vol II: Psychological aspects*. Amesterdam: John Benjamins Ltd.

Bornstein, M., Kessen, W., Weiskopf, S. (1976). Color vision and hue categorization in young human infants. *Journal of Experimental Psychology: Human Perception and Performance, 2,* 115-129.

Brooks, N., Audet, J., & Barner, D. (2012). Pragmatic inference, not semantic competence, guides 3-year-olds' interpretation of unknown number words. *Developmental Psychology.*

Busby Grant, J., & Suddendorf, T. (2011). Production of temporal terms by 3-, 4-, and 5-year-old children. *Early Childhood Research Quarterly*, *26,* 87-95.

Carey, S. (2009). *The Origin of Concepts.* New York: Oxford University Press.

Carey, S. (2010). Beyond Fast Mapping. *Language Learning and Development, 6(3)*, 184-205.

Carey, S. & Bartlett, E. (1978). Acquiring a single new word. *Proceedings of the Stanford Child Language Conference, 15*, 17-29.

Clark, E. (1978). Strategies for communicating. *Child Development, 49,* 953-959.

Clark, E. (1997). Conceptual perspective and lexical choice in acquisition, *Cognition, 64,* 1-37.

Davies, I., Corbett, G., McGurk, H. & MacDermid, C. (1998). A developmental study of the acquisition of Russian colour terms. *Journal of Child Language, 25,* 395-417.

Dedrick, D. (1996). Color language universality and evolution: On the explanation for basic color terms. Philosophical Psychology, 9, 497-524.

Dedrick, D. (1997). Colour categorization and the space between perception and language. Behavioral and Brain Sciences, 20, 187-188.

Franklin, A. (2006) Constraints on children's color term acquisition. *Journal of Experimental Child Psychology, 94,* 322-327.

Franklin, A., Clifford, A., Williamson, E. & Davies, I. (2005). Colour term knowledge does not affect categorical perception of colour in toddlers. *Journal of Experimental Child Psychology, 90,* 114-141.

Franklin, A., Pilling, M. & Davies, I. (2005). The nature of infant color categorization: Evidence from eye movements on a target detection task. *Journal of Experimental Child Psychology, 91,* 227-248.

Franklin, A., Drivonikou, G., Clifford, A, Kay, P., Regier, T. & Davies, I. (2008). Lateralization of categorical perception of color changes with color term acquisition. *Proceedings of the National Academy of Science, 105,* 18222-18225.

Gelman, S., Croft, W., Fu, P., Clausner, T. & Gottfried, G. (1998). Why is a pomegranate an *apple*? The role of shape, taxonomic relatedness, and prior lexical knowledge in children's overextensions of *apple* and *dog*, *Journal of Child Language, 25,* 267-291.

Kay, P., Berlin, B., Maffi, L., Merrifield, W. R. & Cook, R (2009). *The World Color Survey*. Palo Alto, CA: CSLI Press.

Kowalski & Zimiles (2006). The relation between children's conceptual functioning with color and color term acquisition. *Journal of Experimental Child Psychology, 94,* 301-321.

Kuczaj, S. (1982). Young children's overextension of object words in comprehension and/or production: support for a prototype theory of early object word meaning. *First Language, 3,* 93-105.

Long, J. & Luke, J. (2001). *The New Munsell Student Color Set, Second Edition*. Fairchild Books: New York.

Mervis, C., Catlin, J. & Rosch, E. (1975). Development of the structure of color categories. *Developmental Psychology, 11,* 54-60.

Naigles, L. & Gelman, S. (1995). Overextensions in comprehension and production revisited: preferential-looking in a study of 'dog', 'cat', and 'cow'. *Journal of Child Language, 22,* 19-46.

O'Hanlon, C. & Roberson, D. (2006). Learning in context: Linguistic and attentional constraints on children's color term learning. *Journal of Experimental Child Psychology, 94,* 275-300.

Pitchford, N. & Mullen, K. (2005). The role of perception, language, and preference in the development of basic color terms. *Journal of Experimental Child Psychology, 90,* 275-302.

Pitchford, N. & Mullen, K. (2003). The development of conceptual colour categories in pre-school children: Influence of perception on categorization. *Visual Cognition, 10,* 51-77.

Quine, W.V.O. (1960). *Word and object.* MIT Press: Cambridge. MN.

Raskin, L., Maital, S., and Bornstein, M. (1983). Perceptual categorization of color: A life-span study. Psychological Research, 45, 135–145.

Roberson, D., Hanley, J.R. & Pak, H. (2009). Thresholds for color discrimination in English and Korean speakers. *Cognition*, *112*, 482-487.

Roberson, D., Davidoff, J., Davies, I., & Shapiro, L. (2005). Color categories: Evidence for the cultural relativity hypothesis. *Cognitive Psychology, 50,* 378-411.

Saji, N., Imai, M., Saalback, H., Zhang, Y., Shu, H., Okada, H. (2011). Word learning does not end at fast mapping: Evolution of verb meanings through reorganization of an entire semantic domain. *Cognitiion, 118,* 45-61.

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

Shatz, M., Behrend, D., Gelman,S., & Ebeling, K. (1996). Colour term knowledge in two-year-olds: evidence for early competence. *Journal of Child Language, 23,*177-199.

Shatz, M. Tare, M., Nguyen, S.P., Young, T. (2010). Acquiring non-object terms: The case for time words. *Journal of Cognition and Development, 11,* 16–36.

Soja, N. (1994). Young children's concept of color and its relation to the acquisition of color words. *Child Development, 65,* 918-937.

Swingley, D. (2010). Fast Mapping and Slow Mapping in Children's Word Learning. *Language Learning and Development, 6,* 179-183.

Tare, M., Shatz, M., & Gilbertson, L. (2008). Maternal uses of non-object terms in child-directed speech: Color, number and time. *First Language, 28,* 87–100.

Widen, S. & Russell, J. (2003). A closer look at preschoolers' freely produced labels for facial expressions, *Developmental Psychology, 39,* 114-128.

Winawer, J., Witthoft, N., Frank, M., Wu, L., & Boroditsky, L. (2007). Russian blues reveal effects of language on color discrimination. *Proceedings of the National Academy of Sciences, 104,* 7780-5.

Wynn, K. (1990). Children's understanding of counting. *Cognition,* 36*,* 155-193.

Wynn, K. (1992). Children's acquisition of the number words and the counting system. *Cognitive Psychology, 24,* 220-251.

Table 2.1: CIE L*a*b* values for experimental stimuli

Color	L*	a*	b*
Red	22.18	34.93	16.26
Orange	36.94	33.72	38.72
Yellow	51.62	8.63	57.87
Green	31.02	-27.33	24.34
Blue	27.80	-3.33	-22.94
Purple	19.97	13.10	-12.74
Pink	34.92	32.99	5.24
Brown	25.13	11.94	16.79
Black	15.23	1.42	2.62
White	54.91	1.51	9.73
Gray	38.84	2.64	8.12

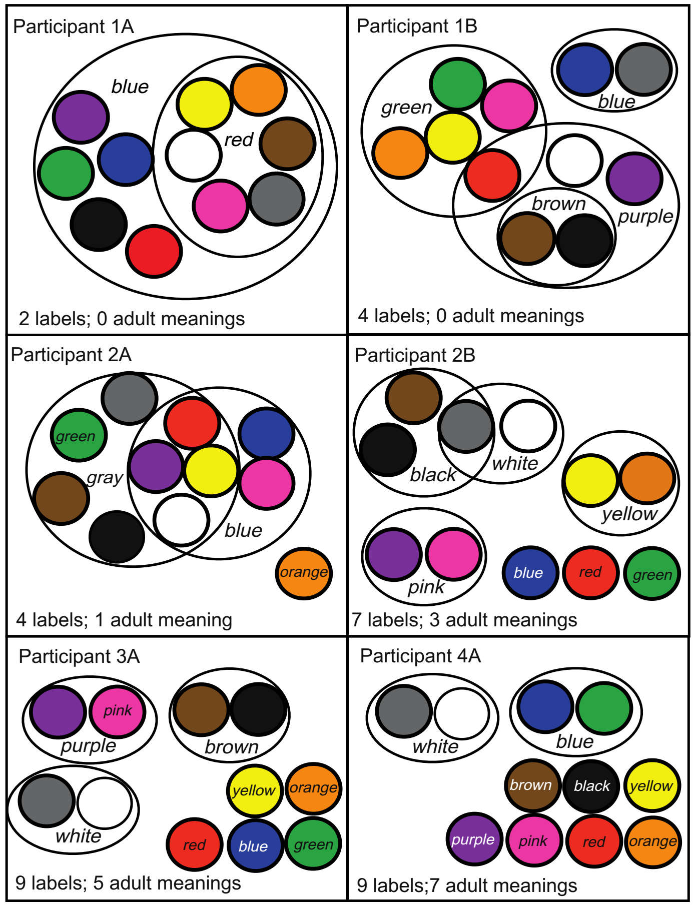

Figure 2.1: Examples of data from six children. Each child is labeled with a number and a letter. The number indicates the child's level and the letter refers to a particular child that is an example of the number level. Labels that are presented within a colored circle were used exclusively for that color (e.g., green for 2A). Each large circle represents a group of colors that were given the same label by a child (e.g., blue and gray as blue by 1B). Overlapping circles or concentric circles indicate that the colors which fall within both circles were given different labels across the two tasks (e.g., brown and black as brown and purple in 1B). Colored circles labeled by correct words indicate adult-like meanings. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

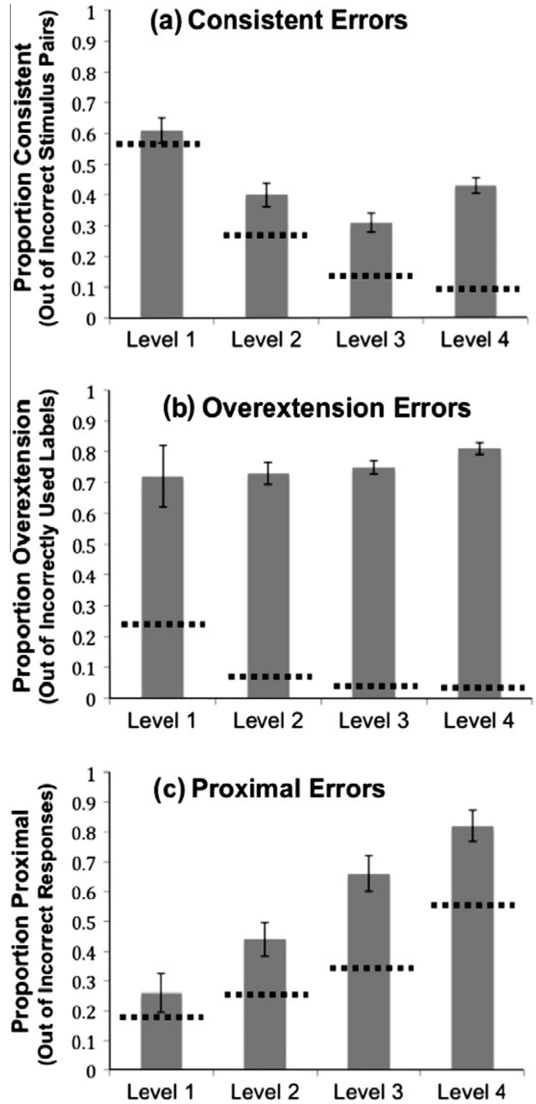

Figure 2.2: Proportion of errors that were consistent, overextended and proximal. Chance rates are marked with a dotted line. Error bars are standard errors of the proportions.

Chapter 2, in full, is a reprint of the material as it appears in Slow mapping: Color word learning as a gradual inductive process. *Cognition,* 22, *127,* 307-317. Wagner, K., Dobkins, K. & Barner, D. (2013). The dissertation author was the primary investigator and author of this paper. Permissions for use of this material have been obtained from Elsevier Limited.

Chapter 3

Partial Comprehension of Color Words Precedes Production

Katherine Wagner, Jill, Jergens, and David Barner

University of California, San Diego

Corresponding Author

Katherine Wagner

Department of Psychology, 0109

University of California, San Diego

La Jolla, CA 92093

kgwagner@ucsd.edu

Abstract

Previous studies report that children use color words in a seemingly haphazard manner before mastering adult meanings. The most common explanation for this is that children struggle to abstract color properties as a domain of linguistic meaning, and this results in a stage in which children produce but do not comprehend color words. However, recent evidence suggests that children's early usage of color words is not random, and that children acquire partial but systematic meanings prior to acquiring adult-like meanings. Here we provide evidence that infants acquire color word meanings even before beginning to produce them. Parent report, a color word production task, and an eye-tracking comprehension task provide converging evidence that partial color word meanings precede production. We conclude that on average color word comprehension precedes production, and this rules out the idea that children's late acquisition of adult-like meanings is due to a failure to abstract color as a domain of linguistic meaning.

Color words pose a particularly difficult problem for children learning language (e.g., Backscheider & Shatz, 1993; Sandhofer & Smith, 1999; Kowalski & Zimiles, 2006; O'Hanlon and Roberson, 2006). As noted in a number of previous reports, children produce color words for many months before converging on adult-like meanings (e.g. Pitchford & Mullen, 2003; Sandhofer & Smith, 1999; Soja, 1994), a pattern also found in other domains of word learning, such as time and number words (Brooks, Audet, & Barner, 2012; Busby Grant & Suddendorf, 2011; Shatz, Tare, Nguyen, & Young, 2010; Wynn, 1992). On the basis of this observation, most previous studies have concluded that the production of color words precedes comprehension. Challenging this conclusion, the present study shows that, on average, children acquire partial meanings of color words prior to production despite the fact that adult-like meanings come much later. We argue that these data are inconsistent with the view that the delay between production and adult-like comprehension is due to a failure to identify color as the appropriate dimension of meaning (Franklin, 2006; Kowalksi & Zimiles, 2006; O'Hanlon & Roberson, 2006; Sandhofer & Smith, 1999). Instead, we suggest that the delay between the onset of production and the acquisition of adult-like meanings is best explained by the problem of determining the boundaries of individual color word meanings (Wagner, Barner & Dobkins, 2013).

In other domains of vocabulary acquisition, researchers have noted that children often acquire some basic comprehension of a word before they being to produce it in speech, such that in infancy and early childhood, the number of words that children comprehend far exceeds the number of words that they produce. For example, Goldin-Meadow, Seligman and Gelman (1976) found that 14- to 27-month-olds understood on average 2.5 times as many words as they produced. Similarly, Harris, Yeeles, Chasin and Oakley (1995) found that there was a consistent lag in the first year of life between when a child first comprehended a word and when they produced that same word. Finally, parent report data on infants between 8 and 16 months that was collected as part of the Macarthur Communicative Development Inventory project revealed that infants comprehend around four times as many words as they produce (Fenson et al., 1994).

However, there appear to be some notable exceptions to this pattern of comprehension before production. For example, children learn to count, i.e. produce number words, many months before they exhibit understanding of the words and can use them correctly to enumerate sets. Similar observations have been made in the domains of time (Shatz et al., 2010), emotion (Widen & Russell, 2003), and of course color (Sandhofer & Smith, 1999). In each of these lexical domains, children respond with incorrect but domain appropriate terms when asked a question (e.g., responding *red* to *What color is it?* when asked about a purple object) before they eventually learn to respond with a correct term (e.g. responding *purple* to previous question; for discussion see Shatz et al., 2010). Typically, this pattern of production of before adult-like comprehension is taken as evidence that the meanings of these words are particularly difficult to master.

In the case of color words, the most common explanation for this difficulty is that children struggle to abstract color as the relevant domain of linguistic meaning. In other words, although children quickly learn to produce and form a category of color words that are associated with one another, they struggle to identify color as the dimension of experience that this category of color words encodes (e.g., Franklin, 2006; Kowalksi &

Zimiles, 2006; Sandhofer & Smith, 1999). Critically, on this account, children's difficulty is specific to abstracting color, per se, rather than with mapping individual color words to particular hues. As evidence for this, proponents of this view note that preverbal infants possess perceptual color categories that are similar to that of English-speaking adults (e.g. Franklin, Pilling & Davies, 2005; Bornstein, Kessen & Weiskopf, 1976). For example, according to Shatz, Behrend, Gelman, and Ebeling (1996), ‘‘on perceptual tasks, infants treat the continuous dimension of hue categorically much as adults do. . . Thus, the apparent difficulty children have with color term acquisition cannot be primarily because the perceptual domain is continuous whereas the lexical domain is discontinuous’’ (p. 178). Accordingly, these accounts argue that once children identify color as the relevant dimension of meaning, color word acquisition can proceed quickly because children can easily map new color words onto pre-linguistic perceptual color categories rather than having to form the color categories as they acquire each word.

While these accounts offer a parsimonious account both of children's difficulty with color words and the origin of color word meanings (i.e., as rooted in perceptual categories), they ultimately cannot explain how children converge on language-specific color categories. This is because children must be able to learn the color boundaries of any of the world's languages. And critically, languages vary both with respect to the number of categories they encode and the precise location of the color category boundaries (Kay, Berlin, Maffi, Merrifield & Cook, 2009). For example, Berinmo, a tribal language with five basic color categories spoken in Papa New Guinea, has the colors *nol* (green, blue and purple) and *wor* (green, yellow, orange and brown) which mark a color boundary that is missing in English, while failing to mark others that are found in English (Roberson, Davidoff, Davies, & Shapiro, 2005). These differences in color word boundaries cannot be entirely explained by the number of basic color words a language has: Korean, English and Russian all divide the blue-green region of color space differently despite having around 11-12 basic color words (e.g Berlin & Kay, 1969; Roberson, Hanley, & Pak, 2009; Winawer, Witthoft, Frank, Wu, & Boroditsky, 2007).

Also, recent evidence suggests that among children who produce color words but do not yet have adult-like meanings, many appear to have partial meanings for their color words nevertheless. As evidence for this, the errors that children make in color word production before they acquire adult-like meanings are both consistent from trial to trial, and systematically related to the target color being tested. In a study of 98 2- to 4-year-olds acquiring English, Wagner, Dobkins and Barner (2013) found that when children incorrectly label a color – e.g., use the word *blue* for a color other than blue – their errors are highly likely to be to a proximal color, like green, rather than to a non-proximal color, like pink. Also, they found that children's errors were typically overextensions of adult categories. For example, when children used *blue* to label green, they almost always also used it to label blue. These data suggest that children acquire partial definitions of color words prior to achieving full adult-like meanings of these words and that the errors that occur during the delay between color word production and comprehension are in part due to a gradual process of determining color word boundaries. Consequently, these data also show that children must identify color as the relevant dimension of meaning earlier than previously thought.

While this study shows that many children have meanings for color words as soon as they produce them, it remains unknown whether partial meanings are acquired shortly after children begin producing color words or instead *prior* to production, as is the case with object labels (Goldin-Meadow, Seligman & Gelman, 1976; Harris, Yeeles, Chasin & Oakley,1995). Evidence of color word meanings – even if partial – prior to production would suggest that children can map color words onto specific regions of color space very early in development, and that identifying the relevant domain of meaning for color words may be no more difficult than it is for other domains of meaning. To investigate this possibility, we first collected parent report data on color word comprehension and production in a sample of children from 18- to 33- months old to identify if, according to parent report, there is a subset of children that comprehend color words without producing any and vice versa, a subset of children that produce color words without yet comprehending any. Then we conducted an eye-tracking task as a more sensitive measure of color word comprehension and an in-lab color word production task to verify that children do not produce any more color words that indicated by the parent report. We identified a group of children that according to parent report, produced color words without comprehending them and a group of children that comprehended color words without producing them. However, both of these groups showed evidence of comprehending color words in the eye-tracking task, allowing us to conclude that on average color word comprehension precedes production.

Methods

Participants

A total of 55 18- to 33-month-olds (24 girls; mean age = 1;11, SD=3.2 mo) participated. An additional 6 children were exclude due to a 50% chance of protanopia or deuteranopia color deficiency based on family history (n=1), failure to complete the task (n=2), less than 50% successful tracking (n=1) and full knowledge of color terms demonstrated during the production task (n=2). For comparison, 23 adults (6 women; mean age = 21;8, SD=1;6) also participated.

Procedure

Parent Report. Parents were asked to complete a questionnaire that asked separately whether children understand and spontaneously say each of the eleven English basic color words (*red*, *orange*, *yellow*, *green*, *blue*, *purple*, *pink*, *black*, *brown*, *gray,* and *white*) as well as the twelve nouns used in the Comprehension Task (see below).

In-lab Production Task. We created eleven cards with pictures of fish in each of the eleven basic English colors. The fish were placed color side down. The experimenter flipped over each card one at a time. After each colored fish was presented to the child the experimenter asked, "What color is this?" The child's response was recorded for each color and the fish card was placed out of view. This was repeated for all eleven colors. See Table 1 for CIE L*a*b* values of stimuli values as measured by a Photo Research-650 SpectraScan.

Eye-tracking Comprehension Task. We utilized Tobii Studio 3.1.6 in combination with a Tobii eye tracker to track children's eye movements. Children were calibrated using Tobii Studio's standard 5-point calibration.

Children viewed 24 scenes, each containing four pictures of the same object, each in a different color. The objects presented were obtained from the UCSD International Picture Naming Project (Szekely et al., 2004) and included socks, chairs, balloons, purses, boxes, cups, cars, kites, stars, boats, books and bows. Images were presented on a white background and the color *white* was not included among the test colors for this reason. The objects were digitally colored with prototypical exemplars of the remaining ten of the eleven basic English colors (*red*, *orange*, *yellow*, *green*, *blue*, *purple*, *pink*, *black*, *brown*, and *gray*) using Photoshop. The stimuli were presented on a Dell P2211H liquid crystal display monitor running off of an ATI Radeon HD 2600HT graphics card, see Table 1 for CIE L*a*b* values as measured by a Photo Research-650 SpectraScan. The four objects in each scene included two pairs of colors, where the colors of each pair were perceptually adjacent to each other, but distant from the members of the other pair (e.g. red and orange vs. blue and green). On each of the 24 trials, there was a target hue (e.g., blue) that was labeled by a color word (*blue*). The color which served as the target was counterbalanced between children such that for every scene, each of the pictures served as the target color for 25% of participants, the close distractor for 25% of participants and one of the two far distractors for 50% of participants. The clips were also presented in a counterbalanced order. Each of the ten colors was targeted two to three times for each participant.

Prior to each trial, an attractor – centered and equidistant between all four objects – was presented to direct the child's gaze to the center of the screen. During each trial, a voice first directed the participant's attention to all of the objects (e.g., *Look at the socks*). Then the voice gave additional information about one of the objects – identifying it by color (e.g., *Look, the orange sock is my sister's*; see Table 2 for list of colors and objects contained in each scene and the audio sentences presented with each). Each scene was presented for approximately six seconds and the color word was spoken at the three-second mark.

If a participant comprehends the target color word, we expect that after the color word is spoken, the proportion of fixations on the target color should be greater than the mean proportion of fixations on the two far distractors. If a participant has an overly broad meaning of the target color word, the proportion of fixations on the close distractor may also be greater than the mean proportion of fixations on the two far distractors.

Results

Parent report

According to parent report, children's comprehension of color words exceeded their production of color words, on average. Out of the 11 basic color terms, parents reported that their children understood a mean of 4.1 words (sd=4.5) and said a mean of 2.6 words (sd = 3.9). On average, parent report exceeded children's performance during the in-lab production task: children produced an average of 1.7 color words in lab (out of the 11 tested). Of the 12 common nouns used in the eye-tracking task, on average parents reported that their children understood 7.8 (sd=2.7) and produced 4.7 (sd=3.9).

Eye Tracking Analyses

For these analyses, children were divided into four groups based both on parent report and the in-lab production task. This allowed us to separately test for color word comprehension in a subset of children reported to comprehend but not produce any color words, and vice versa - a subset of children reported to produce but not comprehend any color words. The remaining two groups were children thought to have no knowledge of color words and children thought to have both production and comprehension knowledge.

Children that did not comprehend or produce any color words according to parent report and who also did not produce color words in lab were included in the No-

Knowledge group (n=17). Children that comprehended but did not produce any color words according to parent report and who did not produce any color words in lab were included in the Comprehension-Only group (n=11). Children who produced but did not comprehend any color words according to parent report were included in the Production-Only group (n=4). Two additional children who had no knowledge of color words according to parent report but produced 1-2 color words in lab were also included in this group for a total of six in the Production-Only group. Children who produced and comprehended color words according to parent report were included in the Comprehension-and-Production group (n=19). One additional child who only comprehended color words according to parent report but did produce color words in lab was also included in this group for a total of 20 in the Comprehension-and-Production group. Only the trials in which there was evidence either from parent report or the lab that the child comprehended the target color word (for the Comprehension-Only group), produced the target color word (for the Production-Only group), or both produced and comprehended the color word (for the Comprehension-and-Production group) were included in the analysis. All trials were included for the Adult and No-Knowledge groups. See Table 3 for details on how many words children in each of these groups know according to parent report and performance on the in-lab comprehension task.

Eye movements were successfully tracked 81% (sd=15.4) of the time for children and 92% (sd=4.4) of the time for adults. The data from the eye-tracking task were binned into 4 time periods. The first time period was used as a baseline and it ranged from the beginning of each trial to 250ms after the color word was spoken to allow time for participants to plan and execute eye-movements. The second time period ranged from

251ms after the color word was spoken to 1000ms, the third from 1001ms to 2000ms and the fourth from 2001ms to 3000ms. For each time period, the proportion of time spent fixating the target and close distractor was calculated and compared to the mean of the time spent fixating both far distractors.

For each participant group (including adults and the three groups of children), we performed a 4 (time: baseline, 251-1000ms, 1001-2000ms, 2001-3000ms) x 3 (fixated image: target, close distractor, far distractor) mixed design ANOVA with time as a within subjects factor and fixated image as a between subjects factor to determine whether participants' looking behavior changed after the target color word was spoken. Fixated image was not analyzed as a within-subjects variable because the amounts of time spent fixating each image type (target, close distractor, far distractor) are necessarily inversely related to each other: as proportion of time fixating the target increases, proportion fixating the distractors consequently decreases (analyzing a variable within-subjects removes the variability due to subject from the error term and would be an inappropriate analysis for this case because fixations to the different image types are inversely, not positively, related within each subject).

A significant interaction between time and fixated image was found for the Adult group ($F(6,198)=157.2$, $p<0.001$), the Production-Only group ($F(6,45)=3.8$, $p=0.004$), the Comprehension-Only group ($F(6,90)=3.1$, $p=0.009$) and the Comprehension-and-Production group ($F(6,171)=2.7$, $p=0.02$). These interactions indicate that the looking behavior of these groups changed over time in response to the spoken color word. No interaction between time and fixated image was found in the No-Knowledge group ($F(6,153)=1.99$, $p=0.07$). These interactions show that all of the participants (except the

No-Knowledge group) responded to the color word. The analyses below are planned comparisons that we ran to determine what is driving the interactions. Specifically they 1) compare fixations to the target to fixations to the far distractors in order to determine if participants comprehend the target color words, and 2) compare fixations to the close distractor to fixations to the far distractors in order to determine if participants have an overly broad understanding of the target color word.

Fixations to the target. Participants who comprehend a target color word should spend more time fixating the target color then far distractors after but not before the color word is spoken. In order to compare looks to the target to looks to the far distractor, we carried out planned one-tailed, independent samples t-tests in all groups. In all but the No-Knowledge group, participants looked more towards the target than to the far distractor after the color word was spoken (but not before), though the timing of the effect varied between groups. Adults looked more towards the target during all three time windows (all ts(44)>14, all ps<0.001) after the color word was spoken. The Comprehension-Only group looked more towards the target than the far distractor during the time window 1000-2000ms after the target color word was spoken ($t(20)=2.7$, $p=0.007$), the Production-Only group looked more towards the target during the 250 to 1000ms time window ($t(10)=2.0$, $p=0.04$), and the Production-and-Comprehension group looked more towards the target during both the 1000-2000ms ($t(38)=4.6$, $p<0.001$) and 2000-3000ms ($t(38)=2.8$, $p=0.005$) time windows. See Figure 1.

Fixations to close distractors. Participants who possess an overextended meaning of a target color word may also show increased fixations to the close distractor as well as the target after the color word is spoken. In order to test the possibility that some children's categories may be overextended, we compared the proportion of time spent fixating close distractors to far distractors in all groups. The Production-Only group looked more to the close distractor than far distractors during the 2000-3000ms time window ($t(10)=2.3$, $p=0.02$). The Comprehension and Production group looked more to the close distractor than far distractors during both the 1000 to 2000ms time window ($t(38)=1.8$, $p=0.04$) and during the 2000 to 3000ms time window ($t(38)=2.6$, $p=0.006$). For both of these groups, early fixation to the target was followed by later fixation to the close distractor, see Figure 1. Neither Adults nor the Comprehension-Only group looked statistically more towards the close distractor than far distractors during any of the time windows after the color word was spoken, though Adults approached significance during the 250 to 1000 ms time window ($t(44)=1.5$, p=0.07). In many studies, adults look more towards semantic and phonetic competitors during spoken language comprehension as measured by the visual world paradigm (e.g. Allopenna, Magnuson & Tanenhaus, 1998; Canseco-Gonzalez et al., 2010; Huettig & Altmann, 2005). Similarly, we find that our Adult group looked slightly more to the close distractor than the far distractors. However, this effect does not reach statistical significance and is much less pronounced then the fixations to the close distractor that occurred towards the ends of the trials in the Production-Only and Production-and-Comprehension groups. These fixations to the close distractor may indicate that our Production-Only and Production-and-Comprehension groups possess overextended color categories, or alternatively, they could also indicate that the children simply have a poorer ability to inhibit looks to incorrect but close competitors.

Analysis of production errors.

We next analyzed the errors made on the in-lab production task by the Production-Only and Production-and-Comprehension groups to determine if the errors indicate that these children possess overextended color categories. We conducted two analyses that determined whether errors were (1) overextensions of adult-like meanings and 2) proximal to the targets in color space. In doing so, we replicated two of the analyses reported in Wagner, Dobkins and Barner (2013).

Based on parent report and performance on the in-lab production task (see method), 20 children were included in the Comprehension-and-Production group and 6 children were included in the Production Only group. Only 18 children of these children (15 and 3, from the respective groups) used color words incorrectly in lab. Of the remaining 8 children, 1 child from the Comprehension and Production group did not respond on all trials but produced only correct color words when she did respond, and 7 were reported by parents as producing color words but did not produce any color words during the in lab production task. Because of the small amount of analyzable data in the Production Only group (3), we combined the data from both groups for the analyses.

First we asked whether children's errors reflected overextensions of adult categories. For example, a child may correctly know that *red* refers to red objects, yet have a broader meaning for *red* than adults, and therefore overextend it to orange and yellow objects. Given that a child used a label incorrectly for at least one of the eleven trials (e.g., *red* to label orange and yellow stimuli), we asked whether they also used that label correctly for its target color. For example, when a child used the word *red* to label orange and yellow, we asked whether they also used *red* to label red stimuli. Of the 41 total words used incorrectly, 30 fit this definition of overextension. This rate (73%) is greater than would be expected from chance (0.33, p<0.001, where chance was calculated similarly to that in Wagner, Barner & Dobkins, 2013). To calculate chance for this analysis, we first calculated the base rates of every term a child used incorrectly to calculate the probability that each of these incorrect terms would also be used on the trials containing the correct color stimulus (e.g., using *red* to label the red fish). We then took the mean of these probabilities to calculate the child's overall probability of overextension. To calculate the overall probability of overextension for each group, we calculated a group mean, weighted by how many labels each child used incorrectly.

Next we asked, given that a child used a color label incorrectly, whether the label and its target stimulus were from perceptually proximal color categories where proximity was defined primarily using the Munsell system. For all colors except gray[5], an error was considered proximal if the stimulus color and the correct referent color for the misused label were from adjacent categories in this space. For example, if a child labeled orange as *red,* this would be considered proximal, but if a child labeled yellow or blue as *red* this would be considered a non-proximal error. The children produced a total number of 111 errors, 41% of which were proximal, a number greater than chance (30%; p=0.006, where chance was calculated similarly to that in Wagner, Barner & Dobkins, 2013). To determine chance, we calculated the probability of each label-stimulus error pair (the probability of using red to label an orange stimulus) as equal to the product of the base

[5] In Munsell space, gray borders every chromatic and achromatic color. However, the focal regions of some colors are much more proximal to gray than those of others. Performing an analysis based solely on adjacency for gray would yield no useful distinctions and would obscure differences in proximity, i.e. would provide no distinction between children that label gray as black and those that label gray as red. Thus, we counted as proximal to gray only those colors that like gray have focal regions with low Chroma -similar to saturation in other spaces (i.e., white, black, and brown).

rates. For example, if 20% (0.2) of errors were in response to an orange stimulus and 80% (0.8) of errors involved using the label *red*, then the probability of using *red* to label orange would be 0.2 x 0.8, or 0.16. To determine the overall chance probability of proximal errors, we summed across the probability of all label stimulus pairs that are classified as proximal.

Discussion

We investigated color word comprehension in a group of young children, including a subset who had yet to produce any color words. We found that many children construct preliminary meanings for color words very early in the acquisition process, sometimes even before they begin producing them in speech. This finding is consistent with recent reports that children have partial meanings for color words early in acquisition (Wagner, Dobkins & Barner, 2013; Pitchford & Mullen, 2003). However, it does not support the idea that the delay between color word production and the acquisition of adult-like meanings is due to a failure to abstract color as a domain of meaning (Franklin, 2006; Kowalksi & Zimiles, 2006; O'Hanlon & Roberson, 2006; Sandhofer & Smith, 1999).

Replicating Wagner, Dobkins and Barner (2013), we found that children's color word errors are highly systematic and perceptual in nature. This suggests that the delay between color word production and acquisition of adult-like meanings is primarily the result of a process whereby children gradually converge on adult-like category boundaries. The current study also provides the first evidence that children often construct partial meanings for color words *prior* to the onset of production. In other words, color word comprehension exceeds production, just as it does for object labels (e.g. Goldin-Meadow, Seligman & Gelman, 1976; Benedict, 1977; Harris, Yeeles, Chasin & Oakley, 1995). Parent reports indicated that children comprehended more color words then they produced. Also, our sample included children who, despite failing to produce color words, nevertheless exhibited comprehension both according to parent report and data from our eye-tracking task. While data from this group indicates that some children comprehend color words prior to production, with data from this group alone, it is difficult to draw any conclusions about a general pattern of development because there could individual differences such that some children produce color words prior to comprehension and others comprehend color words prior to production. However, despite identifying a group of children that according to parent report produced color words without comprehension, these children were found to have at least partial comprehension of the color words they produce according to our eye-tracking task. Taken together, this pattern of results suggests that on average, partial color word comprehension precedes production.

The current results also have important implications for color word learning because contrary to most viewpoints (Franklin, 2006; Kowalksi & Zimiles, 2006; O'Hanlon & Roberson, 2006; Sandhofer & Smith, 1999), they necessitate that children abstract color properties prior to color word production and that the frequent errors children make when using color words cannot be attributed to a failure to identify color as the correct domain of meaning.

This may initially seem at odds with a number of reports that children at this stage have difficulty abstracting color (e.g., Kowalski & Zimiles, 2006; Sandhofer & Smith, 1999). For example, Sandhofer and Smith (1999) found that children were unable to match based on color until after they could successfully label colors. However, such studies often specifically avoid referencing color (Sandhofer & Smith, 1999; Kowalski & Zimiles, 2006; for further critical discussion see Soja, 1994). While these studies interestingly demonstrate that children do not preferentially attend to color early in development, our results among others suggest that children nonetheless can abstract color from an early age, if color is highlighted linguistically or otherwise. In fact, in studies where color properties are highlighted, evidence of color abstraction has been found in infants as young as five to twelve months of age (Catherwood, Crassini & Freiberg, 1989; Waxman, 2007; Wilcox, Woods & Chapa, 2008; Wilcox, 1999).

We suggest that children's errors are not caused by a failure to abstract color but instead that children's color word errors reflect a process by which children begin with approximate definitions and gradually converge on language-specific adult definitions. Our previous account proposed that early approximate definitions begin broad and are narrowed as children acquire additional color terms (Wagner, Dobkins & Barner, 2013). However, the current data suggests that the developmental trajectory of children's definitions may be more complex. Children that were already producing color words looked to objects of both the target color and the closer distractor color, suggesting that their definitions may be broad and include both. However, no such evidence of broad categories was found in the children who comprehended some color words but have yet to produce any. This suggests that children's very first definitions may not be overextended and that overextensions do not occur until later, around the time children being producing color words. One alternative is that children's very first definitions are conservative, perhaps even underextended. Later, before converging on adult definitions, children's definitions may become overextended to accommodate the wide range of hues receiving the same label in their input, e.g. *blue* sky; *blue* jeans; *blue* berries (see Reich, 1976; Barrett, 1978 for precedence of underextension followed by overextension in the domain of object label acquisition).

Evidence of comprehension of color words prior to production also has important implications because it suggests that the acquisition processes for color words and concrete nouns may be more similar than previously thought. Children typically achieve some comprehension of words before they begin to produce them and color words are no exception. Furthermore current and prior evidence (Wagner, Dobkins & Barner, 2013) suggests that the abundance of color word errors frequently noted in the literature are likely overextension errors, an error type also commonly found in the domain of concrete nouns.

Imprecise definitions during the interim between production and adult-like comprehension are not unique to color words. Evidence that children acquire partial meanings prior to converging on adult-like meanings have been found for object labels (Ameel, Malt & Storm, 2008), time words (Shatz et al., 2010; Tillman & Barner, 2013) and emotions (Widen & Russell, 2003). Additionally, whether or not children acquire partial meanings of number words prior to converging on adult-like meanings is currently an issue of active debate (Brooks, Audet, & Barner, 2012; Condry & Spelke, 2008; Sarneck & Gelman, 2004). While children may be able to fast map words to approximate meanings, they require time and experience to converge on the precise adult meanings that are marked in their native language (Carey & Bartlett, 1978).

References

Allopenna, P. D., Magnuson, J. S., & Tanenhaus, M. K. (1998). Tracking the time course of spoken word recognition using eye movements: Evidence for continuous mapping models. *Journal of Memory and Language*, 38, 419-439.

Ameel, E., Barbara, M., & Storms, G. (2008). Object naming and later lexical development: From baby bottle to beer bottle, *Journal of Memory and Language*, 58, 262-285.

Backscheider, A. G., & Shatz, M. (1993). Children's acquisition of the lexical domain of color. In K. Beals et al. (Eds.). *What we think, what we mean, and how we say it, papers from the parasession on the correspondence of conceptual, semantic and grammatical representations, CLS 29 (Vol. 2)*. Chicago: The Chicago Linguistic Society.

Barrett, M. (1978). Lexical development and overextension in child language. *Journal of Child Language, 5*, 205-219.

Bendict, H. (1979). Early Lexical development: comprehension and production. *Journal of Child Language, 6,* 183-200.

Berlin, B. & Kay, P. (1969). *Basic Color Terms: Their Universality and Evolution.* Berkeley, CA: University of California Press.

Bornstein, M., Kessen, W., Weiskopf, S. (1976). Color vision and hue categorization in young human infants. *Journal of Experimental Psychology: Human Perception and Performance, 2,* 115-129.

Brooks, N., Audet, J., & Barner, D. (2012). Pragmatic inference, not semantic competence, guides 3-year-olds' interpretation of unknown number words. *Developmental Psychology, 49*, 1066-1075.

Busby Grant, J., & Suddendorf, T. (2011). Production of temporal terms by 3-, 4-, and 5-year-old children. *Early Childhood Research Quarterly, 26*, 87–95.

Canseco-Gonzalez, E., Brehm, L., Brick, C., Brown-Schmidt, S., Fischer, K. & Wagner, K. (2010). Carpet or Cárcel: The effect of age of acquisition and language mode on bilingual lexical access. *Language and Cognitive Processes, 25*, 669-705.

Carey, S. & Bartlett, E. (1978). Acquiring a single new word. *Proceedings of the Stanford Child Language Conference, 15*, 17-29.

Catherwood, D. Crassini, B. & Freiberg, K. (1989) Infant response to stimuli of similar hue and dissimilar shape: Tracing the origins of the categorization of objects by hue. *Child Development, 60*, 752-762.

Condry, K. & Spelke, E. (2008). The development of language and abstract concepts: the case of natural number. *Journal of Experimental Psychology: General, 137*, 22-38.

Fenson, L., Dale, P., Reznik, S., Bates, E., Thal, D., Pethick, S., Tomasello, M., Mervis, C. & Stiles, J. (1994). Variability in Early Communicative Development. *Monographs of the Society for Research in Child Development, 59*, 5.

Franklin, A. (2006). Constraints on children's color term acquisition. *Journal of Experimental Child Psychology, 94*, 322–327.

Franklin, A., Pilling, M. & Davies, I. (2005). The nature of infant color categorization: Evidence from eye movements on a target detection task. Journal of Experimental *Child Psychology, 91*, 227-248.

Goldin-Meadow, S., Seligman, M., & Gelman, R. (1976). Language in the two-year old. *Cognition, 4*, 189-202.

Harris, M., Yeeles, C., Chasin, J. & Oakley, Y. (1995). Symmetries and asymmetries in early lexical comprehension and production. *Journal of Child Language, 22*, 1-18.

Huettig, F. & Altmann, G., (2005). Word meaning and the control of eye fixation: semantic competitor effects and the visual world paradigm. *Cognition, 96*, B23-B32.

Kay, P., Berlin, B., Maffi, L., Merrfield, W. R. & Cook, R (2009). *The World Color Survey.* Palo Alto, CA: CSLI Press.

Kowalski & Zimiles (2006). The relation between children's conceptual functioning with color and color term acquisition. *Journal of Experimental Child Psychology, 94,* 301–321.

O'Hanlon, C., & Roberson, D. (2006). Learning in context: Linguistic and attentional constraints on children's color term learning*. Journal of Experimental Child Psychology, 94*, 275–300.

Pitchford, N., & Mullen, K. (2003). The development of conceptual colour categories in pre-school children: Influence of perception on categorization. *Visual Cognition, 10*, 51–77.

Reich, P. (1976). The early acquisition of word meaning. *Journal of Child Language, 3,* 117-123.

Roberson, D., Hanley, J.R. & Pak, H. (2009). Thresholds for color discrimination in English and Korean speakers. *Cognition*, *112*, 482-487.

Roberson, D., Davidoff, J., Davies, I., & Shapiro, L. (2005). Color categories: Evidence for the cultural relativity hypothesis. *Cognitive Psychology*, *50*, 378-411.

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

Sarnecka, B. & Gelman, S. (2004). Six does not just mean a lot: preschoolers see number as specific. *Cognition, 92*, 329-352.

Shatz, M., Behrend, D., Gelman,S., & Ebeling, K. (1996). Colour term knowledge in two-year-olds: evidence for early competence. *Journal of Child Language*, *23*,177-199.

Shatz, M., Tare, M., Nguyen, S. P., & Young, T. (2010). Acquiring non-object terms: The case for time words. *Journal of Cognition and Development, 11*, 16–36.

Soja, N. (1994). Young children's concept of color and its relation to the acquisition of color words. *Child Development*, *65*, 918-937.

Szekely, A., Jacobsen, T., D'Amico, S., Devescovi, A., Andonova, E., Herron, D., Lu, C. C., Pechmann, T., Pleh, C., Wicha, N., Federmeier, K., Gerdjikova, I., Gutierrez, G., Hung, D., Hsu, J., Iyer, G., Kohnert, K., Mehotcheva, T., Orozco-Figueroa, A., Tzeng, A., Tzeng, O., Arevalo, A., Vargha, A., Butler, A. C., Buffington, R., & Bates, E. (2004). A new on-line resource for psycholinguistic studies, *Journal of Memory and Language, 51,* 247-250.

Tillman & Barner, D. (2013). Learning the language of time: Children's acquisition of duration words. *Cognitive Sciences Society*, Berlin, Germany.

Wagner, K., Dobkins, K & Barner, D. (2013). Slow mapping: Color word learning as a gradual inductive process. *Cognition, 127,* 307-317.

Waxman, S. R. (2007) Early Word-Learning and Conceptual Development: Everything had a Name, and Each Name Gave Birth to a New Thought, in *Blackwell Handbook of Childhood Cognitive Development* (ed U. Goswami), Blackwell Publishers Ltd, Malden, MA, USA.

Widen, S. & Russell, J. (2003). A closer look at preschoolers' freely produced labels for facial expressions, *Developmental Psychology, 39*, 114-128.

Wilcox, T. (1999). Object individuation: infants' use of shape, size, pattern, and color. *Cognition*, *72,* 125-166.

Wilcox, T., Woods, R. & Chapa, C. (2008). Color-function categories that prime infants to use color information in an object individuation task. *Cognitive Psychology, 57*, 220-61

Winawer, J., Witthoft, N., Frank, M., Wu, L., & Boroditsky, L. (2007). Russian blues reveal effects of language on color discrimination. *Proceedings of the National Academy of Sciences, 104,* 7780-5.

Wynn, K. (1992). Children's acquisition of the number words and the counting system. *Cognitive Psychology, 24*, 220–251.

Table 3.1: CIE L*a*b* values for production and comprehension tasks. Fish stimuli used for the in-lab production task are on the left and for the stimuli presented on screen during the in-lab eyetracking comprehension task are on the right.

	Production Task			Comprehension Task		
Color	L*	a*	b*	L*	a*	b*
Red	58.5	60.4	43.4	59.6	67.3	59.5
Orange	68.2	48.1	54.9	77.3	41	49.7
Yellow	96.7	-2.1	76.7	88.2	-6.2	67.9
Green	65.0	-33.3	24.0	50.1	-38.5	30
Blue	40.9	-4.8	-42.5	28.9	2.4	-107.5
Purple	44.8	20.6	-23.2	37.5	41.5	-56.5
Pink	70.4	48.0	3.9	68.1	48.0	-8.5
Brown	55.5	24.8	34.8	43.7	15.7	27.3
Black	37.6	1.1	4.8	*		
White	99.2	0.2	1.0	**		
Gray	67.4	2.8	4.1	68.7	-0.9	-3.2

* The black color that was presented on screen was too dim to be picked up by the PR-650 Spectrascan.

** White was not included on the eye-tracking comprehension task, see methods.

Table 3.2: Audio and visual stimuli used in the eye-tracking task. All participants saw all 24 scenes, 4 scenes for each of the 6 color groups below. Each scene contained 4 differently colored but otherwise identical objects. The 4 colors included 2 pairs of colors, where the colors of each pair were perceptually adjacent to each other, but distant from the members of the other pair (e.g. red and orange vs. blue and green). Which color served as the target for each scene was counterbalanced between participants.

Colors in scene		
Pair 1	Pair 2	Audio Sentence
yellow orange	gray black	“Look at the cups, look the *target color* cup has water inside." “Look at the socks, look the *target color* sock is my sister's." “Look at the boxes, look the *target color* box has a surprise inside." “Look at the chairs, look the *target color* chair is comfy."
red orange	green blue	“Look at the balloons, look the *target color* balloon is my favorite." “Look at the bows, look the *target color* bow is pretty." “Look at the boxes, look the *target color* box has a toy inside." “Look at the purses, look the *target color* purse is my friend's."
yellow green	red pink	“Look at the balloons, look the *target color* balloon is from the zoo." “Look at the purses, look the *target color* purse is my mom's." “Look at the socks, look the *target color* sock is my brother's." “Look at the kites, look the *target color* kite is awesome."
blue purple	orange brown	“Look at the cars, look the *target color* car is going to the store." “Look at the stars, look the *target color* star is pretty." “Look at the boats, look the *target color* boat is fancy." “Look at the chairs, look the *target color* chair is old."
brown black	green blue	“Look at the bows, look the *target color* bow is new." “Look at the books, look the *target color* book is funny." “Look at the stars, look the *target color* star is awesome." “Look at the boats, look the *target color* boat is my dad's."
gray brown	pink purple	“Look at the kites, look the *target color* kite is new." “Look at the cups, look the *target color* cup has soup inside." “Look at the books, look the *target color* book is exciting." “Look at the cars, look the *target color* car is going to school."

Table 3.3: Summary data for the four groups of children. Children were considered having an adult meaning of a word if they used it to label the correct color (e.g. using *red* to label the red fish) on the in-lab production task and did not use it to label any incorrect stimuli.

	No Knowledge	Comprehension Only	Production Only	Comprehension and Production
N	18	11	6	20
Age in mos	22.3 (0..25)	23.6 (0.30)	21.2 (0.25)	22.8 (0.27)
Parents: Words Comprehended	0	7.0 (3.21) range: 2-11	0	7.5 (3.8) range: 1-11
Parents: Words Produced	0	0	2.1 (1.8) range: 0 to 5	6.7 (3.8) range: 0-10
In-lab: Words Produced	0	0	0.83 (1.2) range: 0 to 3	4.3 (3.4) range: 0 to 10
In lab: Adult Meanings	0	0	0	2.4 (3.0) range: 0 to10

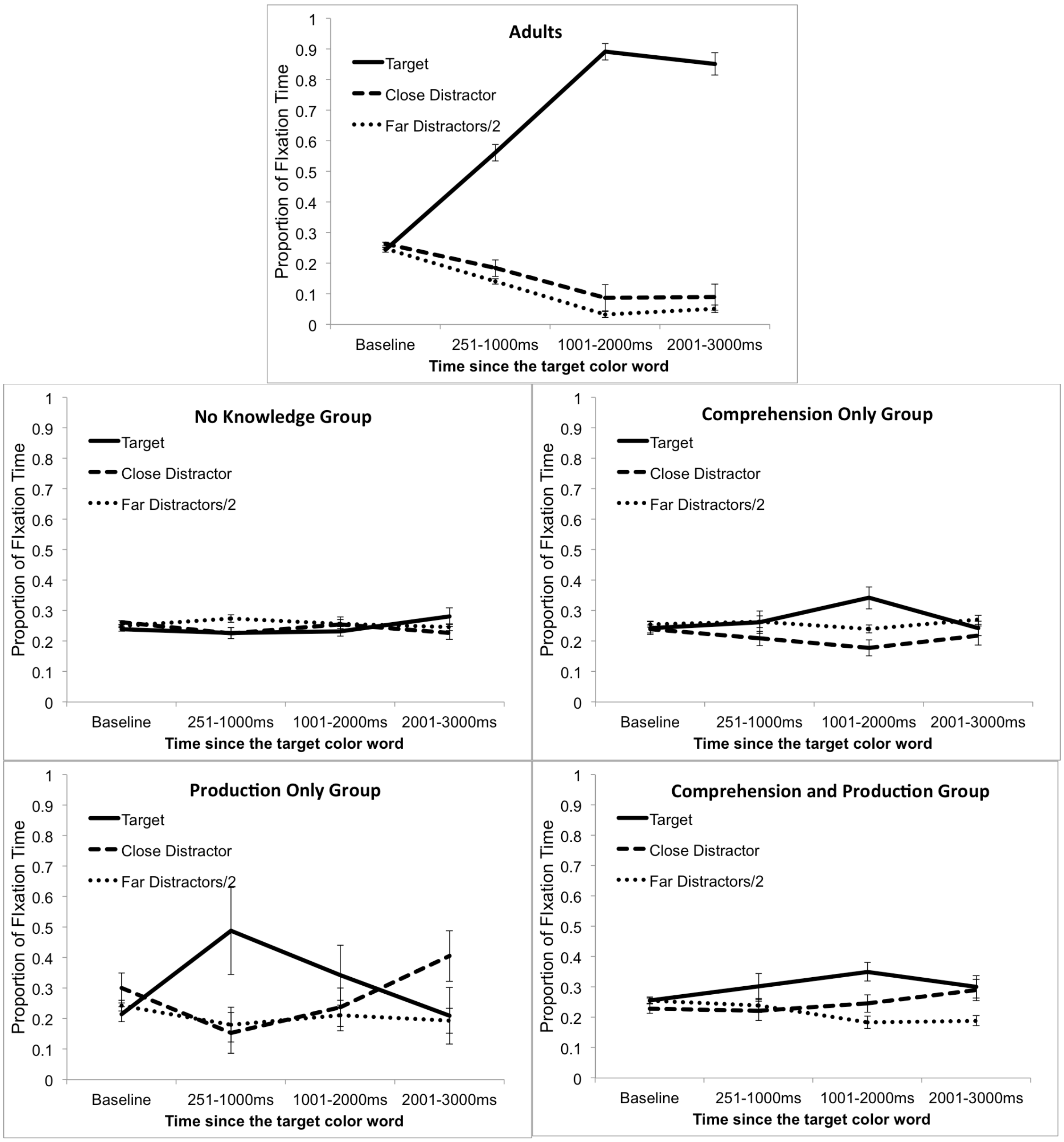

Figure 3.1: The proportion of fixation time spent on the target, close distractor and far distractor plotted over time for each group. Note that since there are two far distractors, the mean of the two far distractors is plotted to allow them to be directly compared to the close distractor and target. The baseline time period includes the period of time from the start of the trial to 250ms after the target color word was spoken.

Chapter 3, in full, is currently being prepared for submission for publication of the material. Wagner, K, Jergens, J. & Barner, D. The dissertation author was the primary investigator and author of this paper.

GENERAL DISCUSSION

Cognitive development begins in infancy when infants start to organize and categorize their perceptual experiences. Towards the end of infancy, human cognition surpasses that of other species with the unique ability to supplement perceptual experiences and cognitive processes with language. Here, we employed color as a case study. At birth, infants have only a weak ability to visually discriminate colors. Yet, within about two years children have developed the ability to categorize colors, form abstract representations of color and assign labels to these representations.

This dissertation investigated early perceptual representations of color and then later, children's ability to supplement and further organize these perceptual representations with language.

In Chapter 1, we investigated the long-standing hypothesis that infants may experience greater intermingling of their senses due to the developmental period of excess connectivity that in visual cortex occurs during the first year of life. We measured looking preferences in response to combinations of colors and shapes. Stimuli consisted of a field of black shapes (triangles or circles) on a colored background, the left and right halves being red and green, respectively (or vice versa). We predicted that if an infant perceives triangles as one color and circles as another, these shapes will interact with the background colors, such that preference for the red versus green background differs between the triangle and circle conditions. Blue/yellow backgrounds were also tested.

We found evidence for an effect of shape on red/green preference at 2-months but not at older ages. We also observed an effect of shape on blue/yellow preference at 3-months that was not present at 2-months and declined with age after 3-months. These results are consistent with infants experiencing synesthetic-like associations between colors and shapes for a brief developmental period.

Additionally, the differences in timing of shape influence on red/green and blue/yellow preference aligns nicely with current knowledge on the developmental trajectory of the parvocellular and koniocellular subcortical pathways that mediate red/green and blue/yellow perception respectively. Red/green chromatic sensitivity develops very early in life, and the decline in synesthetic associations observed between 2 and 3 months is consistent with prior evidence that the red-green pathway may retract from other cortical areas (specifically, motion areas) around the same time, between 2 and 4 months (Dobkins & Anderson, 2002). In contrast to red/green sensitivity, blue/yellow sensitivity develops much later. In fact several researchers have failed to find evidence that infants can discriminate yellow and blue hues until around two months of age (Crognale, Kelly, Weiss & Teller, 1998; Suttle, Banks & Graf, 2002), thus it is likely that the period of exuberant connectivity of the blue/yellow pathways begins and retracts later.

While Chapter 1 investigated the presence of unnecessary mappings between color and shapes that are removed in infancy during the course of normal development, Chapters 2 and 3 were concerned with the formation of highly useful mappings between color concepts and words that allow humans to reference and share their color experiences with one another.

Color words are thought to be particularly difficult to acquire because children use color words incorrectly for several months before they converge on adult-like definitions. Most current explanations exclude the possibility that these errors may result from children's struggle to determine the correct color word boundaries (Franklin, 2006; Kowalksi & Zimiles, 2006; O'Hanlon & Roberson, 2006; Sandhofer & Smith, 1999) primarily because of evidence that infants possess pre-linguistic perceptual categories that are similar to the linguistic categories of English-speaking adults (Franklin, Pilling & Davies, 2005; Bornstein, Kessen & Weiskopf, 1976). Instead, they propose that the delay between production and adult-like comprehension of color words is due to children's failure to abstract color properties as the correct domain of meaning for color words (Kowalksi & Zimiles, 2006; O'Hanlon & Roberson, 2006; Sandhofer & Smith, 1999). Chapters 2 and 3 challenge this assumption.

In order to understand the root of children's errors with color words, we conducted thorough analyses of them in Chapter 2. We asked children to label each of the 11 English Basic Color Terms on two separate tasks and found that their errors are highly systematic and perceptual in nature. We found that when children made labeling errors, they were highly likely to be made in both tasks. Furthermore, most errors appeared to be overextensions of adult color categories. 75% of the labels that were applied to incorrect referents were also used to label the correct referent on both tasks. Finally, a proximity analysis showed that when children made errors, the label they chose often belonged to a perceptually adjacent category. Critically, we found overextension and proximity errors even in children who had not yet acquired any adult-like color word meanings, and for the majority of words that were used by children in the experiment. This suggests that almost all children in the study had abstracted color as a domain and had begun acquiring meanings, but that these meanings were not yet adult-like.

Although these data are consistent with the idea that children begin acquisition by positing overly broad color meanings that are gradually narrowed, it is also possible that early overextension errors reflect a pragmatic strategy (see Clark, 1978). For example, a child who has an adult-like meaning for *red* but not for *orange* might recognize that an orange stimulus is not *red*, but use the word *red* to describe it if they have no better word is available to them. To test this, we used a comprehension task in which the experimenter selected the labels, thereby removing the possibility of communicative overextensions (Clark, 1978; Gelman, Croft, Fu, Clausner & Gottfried, 1998). As in the first production experiment, proximity errors were highly frequent, a finding consistent with the existence of overly broad color meanings, rather than the use of a pragmatic strategy.

While the results discussed in Chapter 2 provide evidence that children construct preliminary meanings for color word much earlier than previously thought, they do not clarify whether partial comprehension is achieved shortly after children begin producing color words or if partial comprehension occurs *prior* to production, as is the case with object labels (Goldin-Meadow, Seligman & Gelman, 1976; Harris, Yeeles, Chasin & Oakley, 1995). Chapter 3 approached this question by examining color word comprehension in younger children, some of who had not yet begun producing color words. Convergent evidence from parental report, an in-lab production task and an eye-tracking comprehension task showed that on average color word comprehension exceeds color word production, just as it does for concrete nouns.

The results from Chapters 2 and 3 have important implications for our understanding of children's ability to represent color and map labels to these representations. First, evidence that two-year-olds form preliminary meanings of color words prior to color production demonstrates that children can abstract color and identify it as a potential domain of linguistic meaning much earlier than previously thought. Consequently, the frequent color word errors that have been noted in the literature (e.g. Kowalksi & Zimiles, 2006; O'Hanlon & Roberson, 2006; Sandhofer & Smith, 1999) cannot be attributed to a failure to abstract color properties. Instead, we showed that these errors are highly systematic and perceptual in nature. This implies that these errors are not caused by a failure to abstract color but rather that they reflect a gradual, inductive process through which children eventually converge on adult like color word boundaries. Furthermore, our account, in contrast to previous accounts, also provides an explanation for how children learning different languages can converge on different color categories.

What remains to be reconciled with the current account is evidence that pre-linguistic infants possess color category boundaries that are highly similar to those of adults (Bornstein, 1976; Franklin et al., 2008; Franklin, Pilling & Davies, 2005)). However, it is still early to be certain that infant's color categories are as similar to adults' as recent evidence suggests. Languages vary both in the number of basic color categories they mark and the location of the boundaries between these color categories. Therefore it is impossible that all infants possess the exact color categories highlighted by the particular language that they will eventually acquire. So far only the category boundaries highlighted in the English language have been tested. The first studies examining phonetic categorical perception in infants concluded that the categories of infants were similar to that of adults (Eimas, Siqueland, Jusczyk & Vigorito (1971) until follow-up studies tested non-native category boundaries in infants and found that unlike adults, infants are sensitive to changes that cross phonetic categorical boundaries of all of the words languages (Werker & Tees, 1984). Further work examining non-native color boundaries in infants may yield similar results, such that while English-learning infants may possess all of the category boundaries that English-speaking adults do, they may also possess categorical boundaries that are not present in English speaking adults but are highlighted in other languages. Additionally, to date these studies have not yet rigorously controlled for perceptual distance when equating within color category changes to between color category changes, though discrepancies in perceptual distance cannot account for all of the results observed (personal communication with Anna Franklin).

Finally, pre-linguistic competence in a domain does not imply an easy word-learning process. The divide between pre-linguistic competence and early word learning is not unique to the domain of color words but is a serious puzzle that confronts the study of language learning as a whole (for discussion related specifically to color, see Bonnardel & Pitchford, 2006; Dedrick, 1996, 1997; Franklin, Clifford, Williamson, & Davies, 2005). Infants possess non-linguistic representations of color, number and time (e.g. Franklin, Pillies & Davies, 2005; Feigenson, Dehaene, & Spelke (2004); Columbo & Richman, 2002) yet in each of these cases struggle to learn words that label these representations (e.g. Busby Grant & Suddendorf, 2011; Shatz, Tare, Nguyen &Young 2010; Brooks, Audet, & Barner, 2012; Sandhofer & Smith, 1999). Language is powerful in part because it can highlight the most relevant aspects of a specific perceptual experience and cultural context but to achieve this expressive power, children must have the flexibility to acquire any of the wide variety of potential linguistic mappings that exist in the world's languages.

REFERENCES

Bonnardel, V. & Pitchford, N. J. (2006). Colour categorisation in preschoolers. *Progress in Colour Studies, Vol II: Psychological aspects*. Amesterdam: John Benjamins Ltd.

Bornstein, M., Kessen, W., Weiskopf, S. (1976). Color vision and hue categorization in young human infants. *Journal of Experimental Psychology: Human Perception and Performance, 2,* 115-129.

Busby Grant, J., & Suddendorf, T. (2011). Production of temporal terms by 3-, 4-, and 5-year-old children. *Early Childhood Research Quarterly, 26*, 87–95.

Catherwood, D. Crassini, B. & Freiberg, K. (1989) Infant response to stimuli of similar hue and dissimilar shape: Tracing the origins of the categorization of objects by hue. *Child Development, 60,* 752-762.

Clark, E. (1978). Strategies for communicating. *Child Development, 49,* 953-959.

Colombo, J., & Richman, W. A. (2002). Infant timekeeping: Attention and temporal estimation in 4-month-olds. *Psychological Science*, 13, 475-479.

Crognale, M., Kelly, J., Weiss, A., & Teller, D. (1998). Development of the spatiochromatic visual evoked potential (VEP): A longitudinal study. *Vision Research, 38,* 3283–3292.

Dedrick, D. (1996). Color language universality and evolution: On the explanation for basic color terms. Philosophical Psychology, 9, 497-524.

Dedrick, D. (1997). Colour categorization and the space between perception and language. Behavioral and Brain Sciences, 20, 187-188.

Dobkins, K. R. (2009). Does visual modularity increase over the course of development? *Optometry and Vision Science, 86,* 583–588.

Dobkins, K. R., & Anderson, C. (2002). Color-based motion processing is stronger in infants than in adults. *Psychological Science, 13,* 76–80.

Dobkins, K., & Teller, D. (1996). Infant motion: detection (M:D) ratios for chromatically defined and luminance-defined moving stimuli. *Vision Research , 36* (20), 3293-310.

Eimas, P., Siqueland, E., Jusczyk, P., & Vigorito, J. (1971). Speech perception in infants. *Science , 171* (968), 303-6.

Feigenson, Dehaene, & Spelke (2004). Origins and endpoints of the core systems of number: Reply to Fias and Verguts. *Trends in Cognitive Sciences, 8*(10), 448-449.

Franklin, A. (2006) Constraints on children's color term acquisition. *Journal of Experimental Child Psychology, 94,* 322-327.

Franklin, A., Clifford, A., Williamson, E. & Davies, I. (2005). Colour term knowledge does not affect categorical perception of colour in toddlers. *Journal of Experimental Child Psychology, 90,* 114-141.

Franklin, A., Drivonikou, G., Clifford, A, Kay, P., Regier, T. & Davies, I. (2008). Lateralization of categorical perception of color changes with color term acquisition. *Proceedings of the National Academy of Science, 105,* 18222-18225.

Franklin, A., Pilling, M. & Davies, I. (2005). The nature of infant color categorization: Evidence from eye movements on a target detection task. *Journal of Experimental Child Psychology, 91,* 227-248.

Gelman, S., Croft, W., Fu, P., Clausner, T. & Gottfried, G. (1998). Why is a pomegranate an *apple*? The role of shape, taxonomic relatedness, and prior lexical knowledge in children's overextensions of *apple* and *dog*, *Journal of Child Language, 25,* 267-291.

Goldin-Meadow, S., Seligman, M., & Gelman, R. (1976). Language in the two-year old. *Cognition, 4*, 189-202.

Goldstein, J., Davidoff, J. & Roberson, D. (2009). Knowing color terms enhances recognition: Further evidence from English and Himba. *Journal of Experimental Child Psychology, 102,* 219-238.

Harris, M., Yeeles, C., Chasin, J. & Oakley, Y. (1995). Symmetries and asymmetries in early lexical comprehension and production. *Journal of Child Language, 22*, 1-18.

Huttenlocher, P., de Courten, C., Garey, L., & Van der Loos, H. (1982). Synaptogenesis in human visual cortex: Evidence for synapse elimination during normal development. *Neuroscience Letters, 33,* 247–252.

James, W. (1890). *The principles of psychology. Volume 1* New York, NY: Henry Holt & Co.

Kelly, D., Quinn P., Slater, M., Lee, K., Ge, L., & Pascalis, O. (2007). The other-race effect develops during infancy: Evidence of perceptual narrowing. *Psychological Science, 18,* 1084–1089.

Kowalski & Zimiles (2006). The relation between children's conceptual functioning with color and color term acquisition. *Journal of Experimental Child Psychology, 94,* 301-321.

Lindsey, D., & Teller, D. (1990). Motion at isoluminance: discrimination/detection ratios for moving isoluminant gratings. *Vision Research , 30* (11), 1751-61.

Maurer, D. (1993). Neonatal synesthesia: Implications for the processing of speech and faces. *Developmental neurocognition: Speech and face processing in the first year of life* , 109-124.

Mills, D., Coffey-Corina, S., & Neville, H. (1997). Language comprehension and cerebral specialization from 13 to 20 months. *Developmental Neuropsychology, 13,* 397–445.

Neville, H. J. (1995). Developmental specificity in neurocognitive development in humans. In M. Gazzaniga (Ed.), *The cognitive neuroscience* (pp. 219–234). Cambridge, MA: MIT Press.

O'Hanlon, C. & Roberson, D. (2006). Learning in context: Linguistic and attentional constraints on children's color term learning. *Journal of Experimental Child Psychology, 94,* 275-300.

Pascalis, O., de Haan, M., & Nelson, C. (2002). Is face processing species-specific during the first year of life? *Science, 296,* 1321–1323.

Rodman, H., & Consuelos, M. (1994). Cortical projections to anterior inferior temporal cortex in infant macaque monkeys. *Visual Neuroscience, 11,* 119–133.

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

Shatz, M., Tare, M., Nguyen, S. P., & Young, T. (2010). Acquiring non-object terms: The case for time words. *Journal of Cognition and Development, 11*, 16–36.

Soja, N. (1994). Young children's concept of color and its relation to the acquisition of color words. *Child Development, 65,* 918-937.

Spector, F., & Maurer, D. (2008). The colour of Os: Naturally biased associations between shape and colour. *Perception-London , 37* (6), 841-847.

Suttle, C., Banks, M., & Graf, E. (2002). FPL and sweep VEP to tritan stimuli in young human infants. *Vision Research, 42,* 2879–2891.

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

Wilcox, T. (1999). Object individuation: infants' use of shape, size, pattern, and color. *Cognition, 72,* 125-166.

Wilcox, T., Woods, R. & Chapa, C. (2008). Color-function categories that prime infants to use color information in an object individuation task. *Cognitive Psychology*, *57*, 220-61