



Slow mapping: Color word learning as a gradual inductive process



Katie Wagner*, Karen Dobkins, David Barner

Department of Psychology, 0109, University of California, San Diego, La Jolla, CA 92093, United States

ARTICLE INFO

Article history:

Received 17 April 2012
Revised 8 October 2012
Accepted 31 January 2013
Available online 28 March 2013

Keywords:

Categorization
Color categories
Word learning
Perception
Development

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.

© 2013 Elsevier B.V. All rights reserved.

1. Introduction

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 is 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 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 iden-

* Corresponding author.

E-mail address: kgwagner@ucsd.edu (K. Wagner).

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).

tifying 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 and Roberson, 2006). Others have argued that color may be salient to children, but still more difficult to abstract than other types of content (Kowalski & 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 Behrend, Gelman, & 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). Similarly, according to Pitchford and Mullen (2003), "Developmental studies have shown young children's perceptual color space is organized in a similar manner to that of the adult. . . Thus, when children engage in the learning of color terms, they already possess color percepts on which color 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, Berlin, Maffi, Merrfield, & Cook, 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 Papua 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 pre-linguistic 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](#); [O’Hanlon and Roberson, 2006](#); [Soja, 1994](#)). 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 ([Bartlett, 1978](#); [Davies, Corbett, McGurk, & MacDermid, 1998](#); [Pitchford & Mullen, 2003](#)). 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 chil-

dren 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.

2. Experiment 1

2.1. Methods

2.1.1. 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, $sd = 7.2$ m), 6 were excluded because they used only one color term during the experiment (mean age = 2;6, $sd = 4.2$ m), and 17 were excluded because they did not cooperate on half or more than half of the trials (mean = 2;7; $sd = 7.2$ m). 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 and gray). The CIELAB values of each color were measured using a Photo Research-650 spectoscan under natural sunlight, see [Appendix A](#).

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.

2.1.2. Procedure

2.1.2.1. 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.

2.1.2.2. 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.

2.2. Results

2.2.1. 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, Behrend, Gelman, & Ebeling, 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 Fig. 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 2–6) during the experiment and had a mean age of 2;5, *sd* = 4.1 m (*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 3–9) during the experiment and had a mean age of 2;7, *sd* = 5.6 m (*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 8–10) during the experiment and had a mean age of 3;2, *sd* = 7.0 m (*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 9–12) during the experiment and had a mean age of 3;2, *sd* = 7.0 m (*n* = 53, 34 girls).

2.2.2. 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 Fig. 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:

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

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 Fig. 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

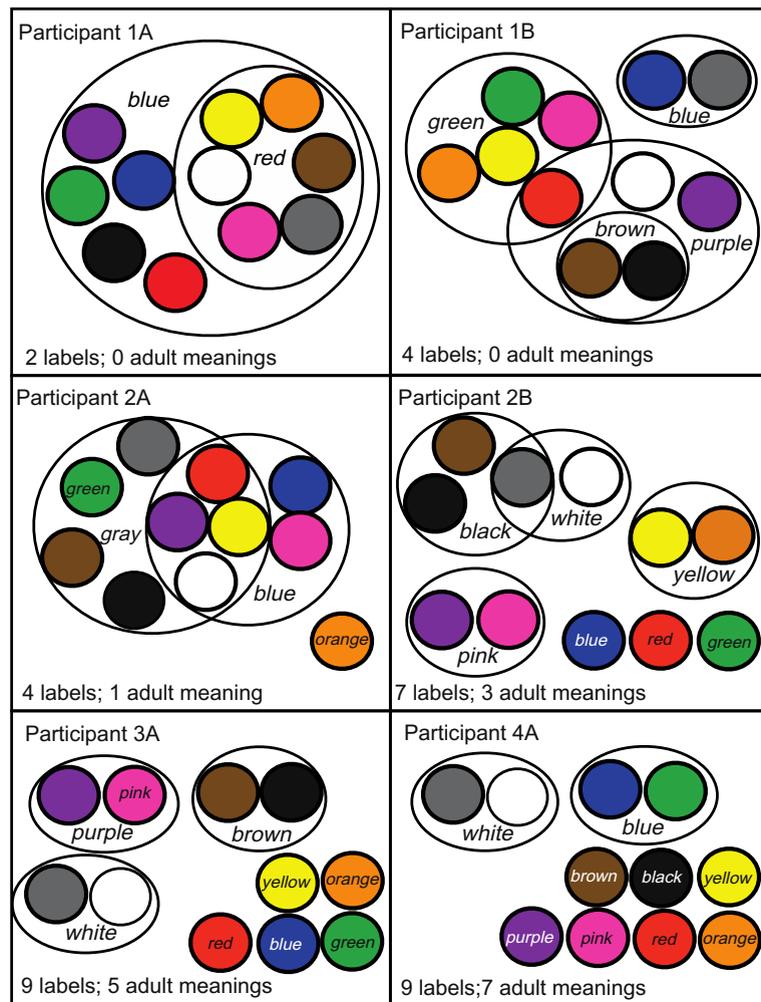


Fig. 1. Examples of data from six children. Each child is labeled with a number and a letter. The number indicates the child's level and the letter refers to a particular child that is an example of the number level. Labels that are presented within a colored circle were used exclusively for that color (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.)

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.

2.2.3. 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 Fig. 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

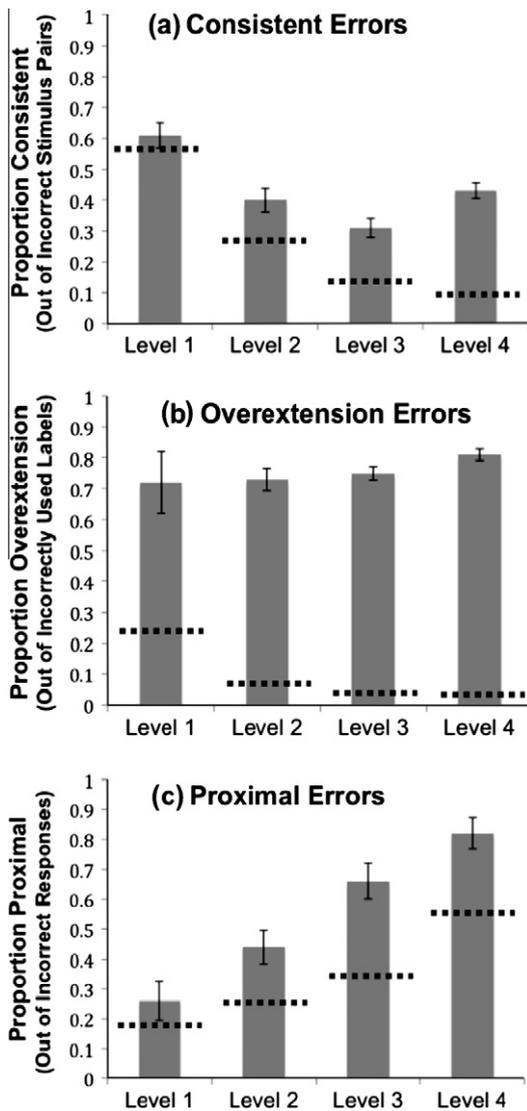


Fig. 2. Proportion of errors that were consistent, overextended and proximal. Chance rates are marked with a dotted line. Error bars are standard errors of the proportions.

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 prob-

ability 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:

$$p(\text{overextension}) = \sum_c \frac{i_c}{i} \frac{\sum_j \left(\frac{i_{cj}}{n}\right)^2}{i_c}$$

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 Fig. 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.

2.2.4. 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 Fig. 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 Fig. 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:

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

where $p(s_i|\text{incorrect})$ is the probability of a particular stimulus i given an incorrect response; $p(l_j|\text{incorrect})$ is the

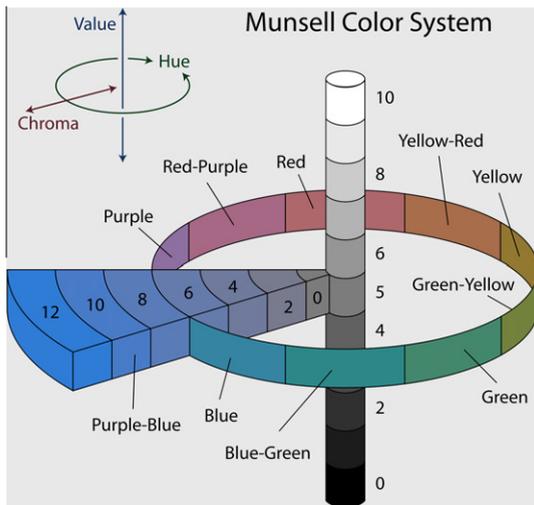


Fig. 3. Representation of Munsell color space (Long & Luke, 2001). Hue is represented by position around the ring, Chroma (similar to saturation in other color spaces) is represented by positions on each spoke and Value (similar to brightness in other color spaces) is represented along the vertical axis. For all colors but gray (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)), 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 interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

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 p s < 0.001 see Fig. 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

2.3. 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 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

² We also separately analyzed proximity errors that involved achromatic colors (i.e., black, white, and 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, 2006).

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 Fig. 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.

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).

3. 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

3.1. Methods

3.1.1. Stimuli

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

3.1.2. 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 = 4.6 m). Data from the remaining 20 children (8 girls) were analyzed. These children ranged in age from 23 to 48 months (mean = 2;10, sd = 6.8 m). 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

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.

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.

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.

3.1.3. 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.

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.

3.2. 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 0–10).

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 Kuczaj, 1982; Naigles & Gelman, 1995).

3.3. 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 than 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 (Mervis, Catlin, & Rosch, 1975; Raskin, Maital, & Bornstein, 1983).

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 Clark,

1997; Saji et al., 2011; Swingley, 2010). 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 to 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., Kowalski & Zimiles, 2006; Sandhofer & Smith, 1999). For example, in their study, Sandhofer and Smith (1999) showed children different shapes of the same color and ask the children to find other shapes that match. 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 Kowalski & Zimiles, 2006). In other studies, the tasks involved a memory component (Kowalski & 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 (Brooks, Audet, & Barner, 2012; Tare, Shatz, & Gilbertson, 2008; Wynn, 1992), 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 (Busby Grant & Suddendorf, 2011; Shatz et al., 2010). 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, 2006; Dedrick, 1996, 1997; Franklin, Clifford, Williamson, & Davies, 2005). 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 pre-linguistic categories.

Appendix A.

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

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* (Vol. 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*. Amsterdam: 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., 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.
- 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.
- 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* (2nd ed.). 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. (2003). The development of conceptual colour categories in pre-school children: Influence of perception on categorization. *Visual Cognition, 10*, 51–77.
- 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.
- Quine, W. V. O. (1960). *Word and object*. Cambridge, MN: MIT Press.
- Raskin, L., Maital, S., & Bornstein, M. (1983). Perceptual categorization of color: A life-span study. *Psychological Research, 45*, 135–145.
- Roberson, D., Davidoff, J., Davies, I., & Shapiro, L. (2005). Color categories: Evidence for the cultural relativity hypothesis. *Cognitive Psychology, 50*, 378–411.
- Roberson, D., Hanley, J. R., & Pak, H. (2009). Thresholds for color discrimination in English and Korean speakers. *Cognition, 112*, 482–487.
- 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. *Cognition, 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.
- Swingle, 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*, 7780–7785.
- 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.