*Color constancy and color term knowledge are positively related during early childhood*.

---

**Article:**

Rogers, MR, Witzel, C, Rhodes, P et al. (1 more author) (2020) Color constancy and color term knowledge are positively related during early childhood. Journal of Experimental Child Psychology, 196. 104825. ISSN 0022-0965

https://doi.org/10.1016/j.jecp.2020.104825

---

# Color Constancy and Color Term Knowledge are Positively Related in Early Childhood

Rogers, M. R.[1], Witzel, C.[2], Rhodes, P.[3], & Franklin, A.[4]

1 The Royal College of Paediatrics and Child Health, London, UK
2 Abteilung Allgemeine Psychologie, Justus-Liebig-Universität, Gießen, Germany
3 School of Design, University of Leeds, UK
4 The Sussex Color Group, School of Psychology, University of Sussex, UK

Word count: 9060

# Color Constancy and Color Term Knowledge are Positively Related in Early Childhood

## 1. Abstract

The ability to keep perception constant despite environmental changes of illumination, viewing angle or distance is a key feature of perception. Here, we investigate how ‘perceptual constancy’ relates to language learning by investigating the relationship between color constancy and color term knowledge in three to four-year-olds. We used a novel method to test color constancy where children are required to match colored stimuli under different illuminations. We found a positive relationship between color constancy and color term knowledge: children who knew more color words also had better color constancy. The relationship remained even when accounting for the effect of age and ability to discriminate colors. The findings have implications for understanding the development of perceptual constancy, language learning and the link between perceptual processing and cognitive development.

## 2. Introduction

We can recognize an apple under many different types of illumination and at varying viewing angles and distances, even though the image reaching the retina varies greatly between those contexts. This process, known as perceptual constancy, gives us a backdrop to detect important changes (Walsh & Kulikowski, 1998) and also enables us to label, and therefore communicate, about objects independent of contextual variation (e.g. “apple”). In childhood, we are faced with the challenge of acquiring words and applying them consistently, and a stable perceptual environment is likely to be key in helping children with the challenge of language learning (Garrigan & Kellman, 2008).

Color is an elementary feature of our visual environment. Some argue that children find it more difficult to learn color terms than other types of words (e.g. Bornstein, 1985; Soja, 1994), although the rate of color-term acquisition also varies greatly between children (Kowalski & Zimiles, 2006). It is unclear why color, which is a seemingly simple perceptual attribute, would be a challenge for children to name. One potential factor is that color naming requires dividing up millions of discriminable colors into relatively few categories. For example, there are many variations of color that would all be termed “green”. To acquire color terms, children must learn how to map terms onto these color categories and identify the boundaries of color categories in their language: where “green” becomes “blue” (Wagner, Dobkins, & Barner, 2013).

A further challenge is that the color of a surface varies due to illumination changes. The light reflected from the surface of an object changes depending on the illumination of the object, leading to different tristimulus color signals when the light reaches the retina. The ability to identify the colors of objects and surfaces despite changes in illumination is called color constancy (Foster, 2011; Smithson, 2005). For example, if you are drinking from a white cup and you move from outside (illuminated by the sun), to indoors (illuminated by incandescent light bulbs), the light reflected from that cup would change. However, our brain discounts the illumination and keeps a constant perceptual experience of a white cup. The adult visual system usually achieves color constancy remarkably well (Abrams, Hillis, & Brainard, 2007). Some of the effects of illumination are already compensated at an early, sensory stage of color processing, such as adaptation or local contrast (Hansen, Walter, & Gegenfurtner, 2007; Kraft & Brainard, 1999). Therefore, it is unsurprising that evidence for rudimentary color constancy, adaptation and local contrast are found to be already present during infancy (Dannemiller, 1989; Dannemiller & Hanko, 1987; Pereverzeva & Teller, 2009; Yang, Kanazawa, Yamaguchi, & Kuriki, 2013). However, low-level sensory mechanisms cannot fully explain color constancy (Foster, 2011; Smithson, 2005), therefore higher-level processes must be involved in disentangling changes of color that are inherent to the object, from changes of color that are due to the illumination. For example, scene interpretation, such as knowing that certain objects have a canonical color, plays an important role in color constancy (Mollon, Bosten, Peterzell, & Webster, 2017; Winkler, Spillmann, Werner, & Webster, 2015; Witzel, Racey, & O'Regan, 2017).

Additionally, if we consider all theoretically possible surfaces and illuminations, there are infinite possible shifts in the color of a surface under illumination changes – there are an infinite number of ways in which surface colors can change (Logvinenko, 2009; Logvinenko & Tokunaga, 2011), but very few of them occur in our visual environment (Foster, Amano, & Nascimento, 2006). Hence, color constancy may also require the observer to learn which color shifts can be expected in the visual environment (Witzel, van Alphen, Godau, & O'Regan, 2016). To be able to reliably name the colors of objects, a child has to distinguish both the variability of colors within and across color categories (e.g., Wagner, Dobkins & Barner, 2013), and the variability of colors due to objects and illumination changes.

Therefore, the challenge of color constancy may contribute to the difficulty of color-term learning. If so, children with advanced color constancy could also be advanced in color naming as a result. Whilst good color constancy could support color term learning in this way, color term learning could also aid color constancy: having terms for the colors of objects may help anchor perception under environmental change. The first step in distinguishing between these theoretically important possibilities, is to establish whether the development of color constancy and the learning of color terms are related.

In adults, colors that are named most consistently across observers are also named most consistently across illumination changes, demonstrating a link between color naming and color constancy changes (Olkkonen, Hansen, & Gegenfurtner, 2009; Olkkonen, Witzel, Hansen, & Gegenfurtner, 2010). Three- to four-year-olds who can group colors well with color terms, are also better at keeping these color groupings constant under different illuminations (Witzel, Sanchez-Walker, & Franklin, 2013). However, these studies are concerned with the consistency or constancy of naming and sorting multiple colors into groups across illuminations, rather than explicitly matching surfaces across illuminations. They do not demonstrate that someone who is better able to name colors is also better able to keep a color constant under changes in illumination. The current study aims to establish whether color constancy and color naming are related in 3-4 year old children who are still learning color terms. Are children with more advanced color naming also better at keeping colors perceptually constant across illumination changes?

To measure color constancy in the current study, children are required to match a physical colored target under one illumination to one of four surfaces viewed under a different illumination. The task is made into a developmentally-appropriate game about matching the colors of trousers for two bears (adapted from Franklin, Clifford, Williamson, & Davies, 2005). One stimulus was created to be a “light match” to the target: it reflected light with the same sensory color signal (CIE xyY; u`v`) as the target when under a different illumination. Another stimulus was the “color constant match”: it consisted of the same surface as the target and cut from the same colored card. The other two stimuli were colorimetrically midway between the light match and the color constant match. This means that one stimulus under red light has the same color measurements as the other stimulus under white light.

A child without color constancy would see the light match stimulus as identical to the target under the two different illuminations. On the other hand, a child with adult-like color constancy would not be ‘fooled’ by this, and would recognize the stimuli cut from the same piece of paper under the two lighting conditions. The midpoint stimuli allow us to test graduations of color constancy. Each trial was scored from 0 to 3: the higher the score, the closer the child’s choices were to the color constant match. The experimental setup is such that the child would not adapt fully to one illumination or the other, as they are simultaneously visible, thus likely resulting in adaptation to a midpoint. However, this does not present a problem to the method as this would make the task more challenging and prevent a ceiling effect in the results. To control for general task ability such as children’s ability to match and discriminate colors, there were also control trials were children were required to match stimuli with no difference in illumination (i.e. the task required only color discrimination). Color-term knowledge was assessed by testing production and comprehension of the eight basic chromatic color terms. If the development of color constancy and color-term learning are linked, we predict that children who are better able to match colors correctly despite changes in illumination will also be better at naming and comprehending color terms. Such a finding would provide the first clear evidence for a link between the development of perceptual constancy and language learning.

## 3. Methods

### *Participants*

Forty-nine children (21 girls) with no family history of color vision deficiency took part.[1] Seven children were excluded due to completion of less than 50% of the trials (mean age = 34 months, SD = 3.5 months). Data from the remaining 42 children (17 girls) were retained for analysis. They were between 33 months and 45 months of age (mean age = 37 months, SD = 3.1), 86% went to nursery (full time or part time) and they had an average birth weight of 3.34kg (SD = 0.51). All children were presented with a small gift (a book) at the end of the study as a thank you for participating. This study received ethical approval from the Science and Technology Cross-Schools Ethics Committee (C-REC) at the University of Sussex and the European Research Council Executive Agency Ethics Committee (ref 230685).

---

[1] [Family history of color vision deficiency was taken but children were not tested on formal color vision deficiency tests due to both the unreliability of many currently available tests with this age range and the added demands on testing time and children's attention during the testing sessions. Based on prevalence rates of color vision deficiency we expect no more than two participants to be affected and likely less given the screening using family history. Therefore, we do not anticipate any undetected color vision deficiency to have a meaningful effect on our results]

### ***Color constancy task***

#### Apparatus and stimuli

Two stimuli were custom printed using a Hewlett-Packard Designjet Z3200 large-format printer on HP Premium Matte photo paper. This printer uses a 12-ink system to achieve a wide color gamut and its pigmented inks are claimed to be lightfast for 200 years ("Datasheet: HP DesignJet Z3200 Photo Printer series," 2017). The stimuli were matched to target values precisely (0.3 $\Delta E_{00}$) and were designed such that card 1, "*purple*", under D65 illumination (which simulates daylight) was a metameric light match of card 2, "*blue*", under red Lee filtered light. This means that when card 2 is measured under red light, it is the same color as card 1 under natural light: they are a "light match". This was confirmed by measuring the card under the two illuminations with a SpectraScan PR-655 spectroradiometer (Photo Research Inc., Chattsworth, CA; see Table 1 and Figure 1 for measurements).

Two additional stimuli were then selected in matte Munsell card to fall roughly between the measured chromaticity and luminance values (CIE u'v') of the two custom printed cards (see Figure 1). CIE u' and v' are the red-green and blue-yellow axes of a perceptual color space. The additional midpoint stimuli had a Munsell value of 5 PB 5/6 and 7.5 PB 4/6, we have labelled these as "*blue-purple*" (*b-p)* and "*purple-blue*" (*p-b*), respectively. The midpoint stimuli were indistinguishable from the printed test card in all ways except for the color (e.g., identical in texture), therefore type of card could not be a cue in the task. The training stimuli were matte Munsell card, in clearly discriminable colors, but with a similar luminance (Munsell values: 2.5B 7/8; 7.5B 3/6; 5Y 8/12; 5R 5/12).

All stimuli were cut into trouser shapes, of the same dimensions (9cm x 4.5cm). We also created white cardboard cut-outs of bears for the task (18cm x 9.5cm). We used a customized lightbox in this task to control the illumination. The box was separated into two compartments: one side was lit by unfiltered D65 simulator bulb (VeriVide fluorescent D65, length 600mm, wattage 18, diameter 38mm) and the other side had a red Lee filter (number 035) covering the aperture resulting in red-filtered light from the same source. We refer to these as “white illumination” and “red illumination” respectively. See Figure 2 for photos of the setup.

Table 1. The CIE xyY and u`v` values of the test stimuli under white illumination and red filtered illumination. The rows in bold show that the purple stimulus under white light is a metameric match of the blue stimulus under red light.

| | White illumination | | | | | Red filtered illumination | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | x | y | Y | u` | v’ | x | y | Y | u’ | v’ |
| *blue* | 0.242 | 0.266 | 26.56 | 0.170 | 0.419 | **0.261** | **0.244** | **15.33** | **0.193** | **0.403** |
| *blue-purple (b-p)* | 0.246 | 0.253 | 22.09 | 0.177 | 0.410 | 0.266 | 0.234 | 12.98 | 0.202 | 0.399 |
| *purple-blue (p-b)* | 0.242 | 0.227 | 18.86 | 0.185 | 0.390 | 0.263 | 0.213 | 8.95 | 0.210 | 0.381 |
| *purple* | **0.261** | **0.236** | **14.95** | **0.197** | **0.400** | 0.285 | 0.223 | 7.39 | 0.223 | 0.393 |

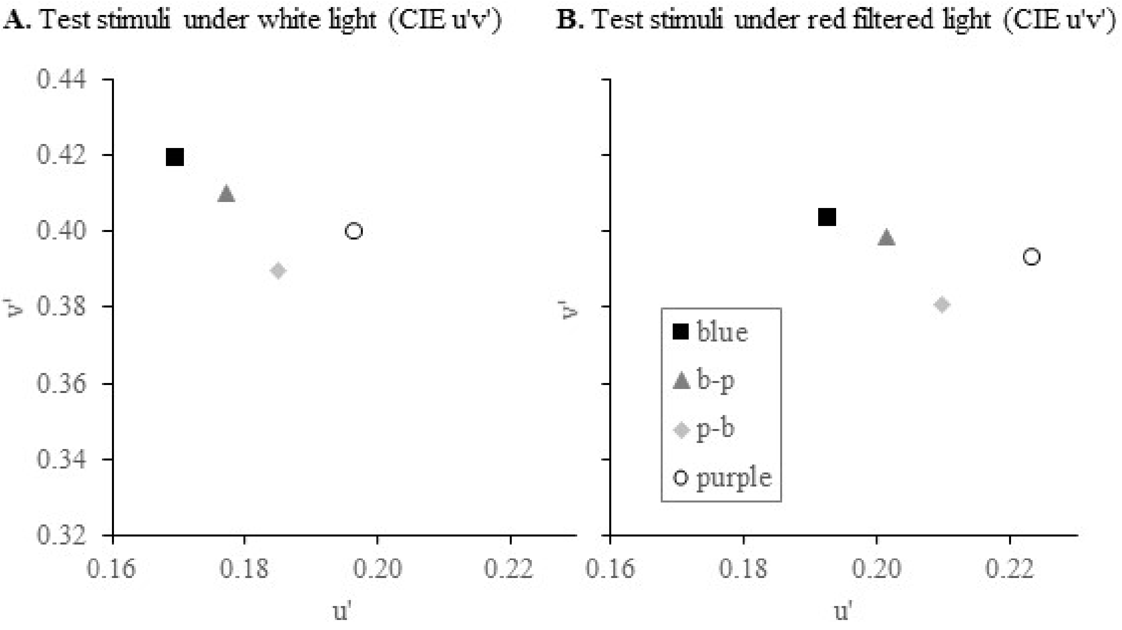

Figure 1. Colorimetric measurements of the four test stimuli in CIE u'v', under A) white light and B) red light. These figures show that blue stimulus (black square) under white light measures colorimetrically the same as the purple stimulus (unfilled circle) under red light. Stimuli b-p and p-b fall roughly midway between the blue and purple stimuli. CIE u'v' is a perceptual color space where the chromaticity of color is represented numerically as u' and v' values. The axis u' represents how red-green the stimulus is (high numbers indicating redder colors), and v' how blue-yellow the stimulus is (high numbers indicate yellower colors).

### Procedure

Color constancy was measured with a simultaneous matching task, where target and response stimuli had different illumination. We also conducted control trials with all stimuli under the same illumination. The child sat in a chair facing the lightbox, Experimenter 1 sat next to the child, the child's parent or carer sat close by, and Experimenter 2 sat behind them. Experimenter 1 read the child a story book introducing the concept of matching colored trousers on bears. The story consisted of photos of the bear cut-outs and stimuli used in the experiment, and showed two bears wearing the same colored trousers set in different everyday scenes (e.g., at the park).

Experimenter 1 then turned on the lightbox whilst Experimenter 2 simultaneously turned off the main light illuminating the room. Thus, the room was completely dark apart from the light from the lightbox. Experimenter 1 explained the task whilst the child adapted to the illumination. The training trials then began. Experimenter 1 put the cardboard cut-out bears (see Figure 2) in the white light section of the lightbox with the four training stimuli the child could choose from. Experimenter 1 also held four of the same color stimuli hidden in her hand. The child was reminded that these two bears always liked to wear the same color, then placed one stimulus from her hand on a bear. "Remember that these bears always like to wear the same trousers. Can you make them match?" Experimenter 1 placed one trouser-shaped stimulus from her hand onto the bear and asked the child to place a stimulus from the four options in the lightbox on the other bear to make them match. When the child had made a match, the procedure was repeated. The stimuli placement was shuffled between trials. The participants did not receive any feedback on the selection. The training phase finished when the child got three in a row correct.

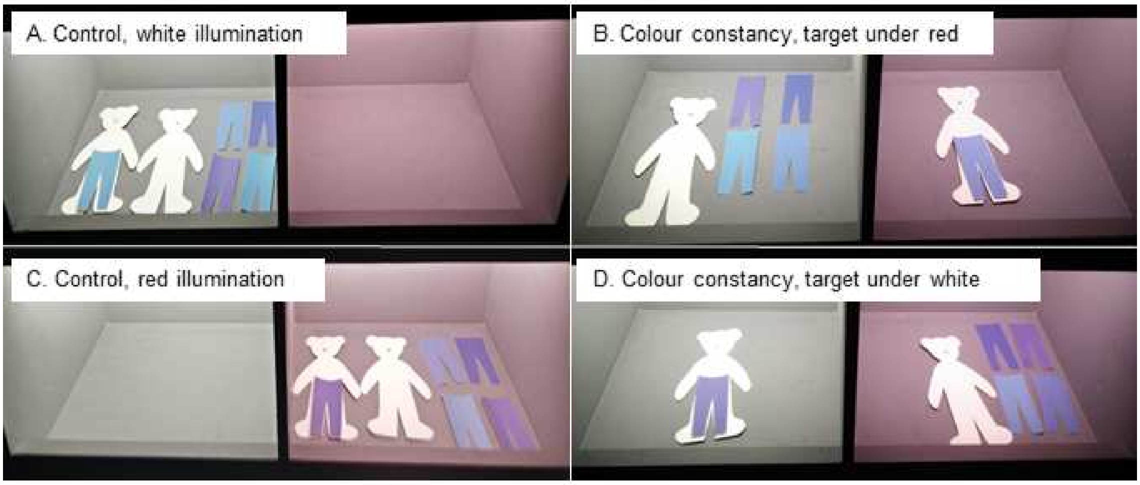

Figure 2. Representation of the layout of the lightbox and the position of the stimuli in the four test conditions: A) Match target stimulus to options under white illumination B) Match target stimulus under red illumination to options under white illumination; C) Match target stimulus to options under red illumination and D) Match target stimulus under white illumination to options under red illumination.

For the experimental phase, Experimenter 1 replaced the training stimuli with the four experimental stimuli in the lightbox (see Figure 2). The procedure was the same as the training trials. In the control trials, all stimuli were presented under the same illumination (i.e. all stimuli were in one compartment of the box).

Experimenter 1 placed a target stimulus on one bear, and then the child was asked to make them match by selecting the corresponding stimulus from the four options laid out. For the purple color constancy trials, the purple target and one of the bears were placed in the white illuminated side of the box, whilst the other bear and four options were placed in the red illuminated side. For the blue color constancy trials, the blue target and one of the bears were placed in the red illuminated side of the box, whilst the other stimuli were in the white illuminated side. The target was only ever the ‘blue’ or ‘purple’ stimulus in the discrimination and color constancy trials. Experimenter 2 noted down the child’s selection.

The red filter was counterbalanced between left and right position across participants. The task order was either A, B, C, D or C, D, A, B, so that trials alternated between control and color constancy (see Table 2 and Figure 2 for trial types). Each type of trial was conducted three times. The trial order was not fully counterbalanced as this would require moving the selection stimuli between compartments for each trial. A pilot study showed that children lost attention and completed fewer trials when the stimuli had to be moved between each trial.

Table 2. Four different types of trial were conducted, that assessed the child's ability to match the target stimuli to the selection stimuli, under the same or different illumination.

| Trial type | Task | Illumination for selection stimuli | Target stimulus | Illumination for target stimulus |
|---|---|---|---|---|
| A | Control (match under same illumination) | Neutral illumination | Blue | Neutral illumination |
| B | Color constancy | Neutral illumination | Blue | Red illumination |
| C | Control (match under same illumination) | Red illumination | Purple | Red illumination |
| D | Color constancy | Red illumination | Purple | Neutral illumination |

### Coding

For each trial in the color constancy task, the child's target stimulus selection was appointed a score from 0 to 3. When the target was *purple*, zero points were given for a selection of *blue*, one point for a selection of *blue-purple,* two points for a selection of *purple-blue* and three points for a selection *purple* (i.e. target match). When the target was *blue*, the awarding of points were reversed. See Table 3 for a full list of scores.

If the child chooses a target colorimetrically further away from the target under the same illumination, it indicates that the ability to discriminate that color from the target is low; hence, the further the chosen color, the lower the discrimination performance. By calculating points for each response, this allowed us to calculate a graded color constancy score and also a color discrimination score from the control trials.

An overall *color constancy score* and a *discrimination score* were calculated for each participant, by summing their points for each trial (i.e. a number out of 3) within a condition, and then calculating this as a percentage of the highest number of points possible for the number of trials they attempted. The equation to calculate the color constancy score is:

$$Colour\ constancy\ score = \frac{\sum Scores\ from\ all\ CC\ trials}{Highest\ score\ possible\ from\ CC\ trials\ attempted} \times 100$$

Equation 0.1

The discrimination score calculation is the same, except using the control trial scores. Not all children completed the whole set of trials; 88.76% of all possible trials were completed. Scores were calculated out of number of trials that the child attempted rather than number of all possible trials. This measurement best represents children's comprehension, and reflects their true color constancy capacity rather than their fatigue with the task. The following analyses find the same results if measures are calculated out of all possible trials.

Table 3. Points for color constancy and control tasks

| Selection | Points when target blue | Points when target purple |
|---|---|---|
| blue | 3 | 0 |
| blue-purple (b-p) | 2 | 1 |
| purple-blue (p-b) | 1 | 2 |
| purple | 0 | 3 |

### ***Color naming task***

#### Stimuli

For the color term comprehension task, we used an image of a rabbit surrounded by eight different colored clothes on laminated card. These colors were chosen to be good examples of the eight basic chromatic color terms: blue, green, red, yellow, pink, purple, orange and brown. For the color term production task, we used images of a rabbit wearing each of the different colored clothes on separate pieces of laminated card (as in Pitchford & Mullen, 2002).

#### Procedure

Following the color constancy task, the main light was switched back on, and the child was presented with the card described above. The child was asked to point to the color as Experimenter 1 said the color term ("Where is the *red* jumper?"). This was done for each color term. For the color term production task, the child was shown each of the eight color naming cards and asked to say the name of the colors ("What color jumper is the rabbit wearing here?"); the color naming cards were shuffled between participants.

## 4. Results

Table 4 gives the percentage of children who could accurately comprehend and produce each of the color terms.

Table 4. Percent accuracy and SD for comprehension and production of each of the color terms.

| Color Term | Comprehension percent accuracy | Comprehension SD | Production percent accuracy | Production SD |
|---|---|---|---|---|
| red | 85.7% | 0.35 | 88.1% | 0.33 |
| orange | 85.7% | 0.35 | 95.2% | 0.22 |
| pink | 83.3% | 0.38 | 90.5% | 0.30 |
| purple | 85.7% | 0.35 | 76.2% | 0.43 |
| blue | 81.0% | 0.40 | 88.1% | 0.33 |
| green | 90.5% | 0.30 | 78.6% | 0.42 |
| yellow | 85.7% | 0.35 | 78.6% | 0.42 |
| brown | 92.9% | 0.26 | 73.8% | 0.45 |

The average number of color terms participants could comprehend was 6.90 out of 8 (SD = 1.56), and the mean number of color terms participants could produce was 6.69 out of 8 (SD = 1.51). Pearson's correlation showed that color term comprehension and production were positively correlated ($r = 0.705$, $p < .0001$). Therefore, to investigate the relationship between color naming and color constancy, we averaged together color term comprehension and production to produce a color naming score for each participant.

The color constancy scores were non-normally distributed, as indicated by the Shapiro-Wilk Test ($p = .004$) with skewness of -.581 (*SE* = 0.365) and kurtosis of -0.818 (*SE* = 0.717). Therefore, Spearman's rank correlations were carried out on the data. Spearman's rank correlation indicated a positive relationship between color naming and color constancy (*rho* = .521, $p < .001$).[2] Further, more stringent, correlations were carried out using the robust correlation toolbox (Pernet, Wilcox, & Rousselet, 2013), to ensure outliers do not contribute to the correlation. The bend correlation and the skipped correlation protect against outliers, and we still find a relationship using these measures (bend correlation: $r = 0.500$, $p < .001$; skipped correlation: $r = 0.521$, $p < .05$). This suggests that children who know more color terms have a higher level of color constancy. See Figure 3 for a scatterplot showing the relationship between color naming and color constancy score. Piloting revealed that adult participants achieved 100% scores on this task.

In order to control for color discrimination, we also calculated a "color constancy ratio", in which we divided the raw color discrimination score by the raw color constancy score. Thus, individual discrimination abilities are used as a baseline and taken into account. A ratio of 1 indicates equal performance on both conditions, and the lower the ratio the poorer color constancy score is relative to the color discrimination score. The color constancy ratios were also non-normally distributed according to the Shapiro-Wilk Test ($p = .004$; skewness = -.437, *SE* = 0.365; kurtosis = -0.822, *SE* = 0.717). Spearman's rank correlation indicated a positive relationship between color naming and color constancy ratio (*rho* = .553, $p < .001$).

---

[2] Note that correlations retain significance if conducted on color production (*rho* = 0.610, $p < .001$) or color comprehension (*rho* = 0.371, $p = .016$) separately rather than the averaged color naming measure

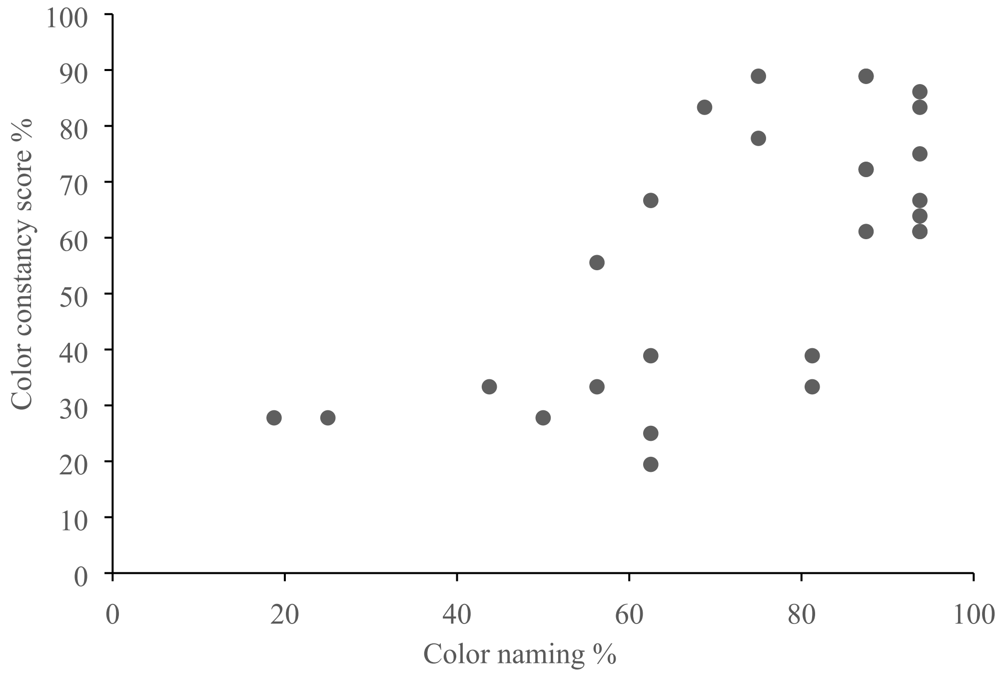

Figure 3. Color constancy score and color naming ability for each participant. Color naming is calculated by averaging the color term comprehension and color term production scores, converted to a percentage. There is a positive relationship between the two variables. N = 42.

When color discrimination is controlled for in a partial Spearman's correlation, the relationship between color constancy and color term knowledge remains (*rho* = 0.549, $p < .001$), which suggests that poor discrimination performance on the control trials cannot account for the relationship between color constancy and color term knowledge. Furthermore, the relationship remains when controlling for both age and color discrimination score (*rho* = 0.587, $p < .001$). This suggests that it is not simply the case that older children are better at both color naming and the color constancy task, thus driving the relationship between color naming and color constancy score. See Figure 4 for a scatterplot of color naming and color discrimination score. There was also no relationship between color naming and age (*rho* = 0.085, $p = .591$), suggesting that age is not the driving factor explaining individual differences in color naming.

We also conducted Bayesian analyses on the relationships of interest. Bayes factors allow interpretation of the strength of the evidence for either the null or alternative hypothesis (Dienes, 2014). A Bayes Factor (*B*) of 0.33 or lower indicates evidence for the null hypothesis, and a *B* of 3 or above indicates substantial support for the alternative hypothesis. Values between 0.33 and 3 (closer to 1) suggest that the data is not sensitive enough to concretely support either the null or alternative hypothesis (Wetzels & Wagenmakers, 2012). As there has been little prior work in this field, and due to the non-parametric nature of the data, it is difficult to predict priors (i.e. the pattern of data we expect). Therefore we conducted Bayesian non-parametric correlations (Kendall's Tau) and used the default beta width prior of 1, which assigns equal prior probability to all correlation values between -1 and 1 (van Doorn, Ly, Marsman, & Wagenmakers, 2016). These analyses provided support for the experimental hypothesis in the relationship between color naming and color constancy score, $\tau = 0.390$, $BF_{10} = 126.8$; support for the null hypothesis in the relationship between color naming and discrimination score, $\tau = 0.073$ $BF_{10} = 0.251$ and support for the null hypothesis in the relationship between age and color naming $\tau = 0.072$ $BF_{10} = 0.249$. The data were not sensitive for the relationship between age and color constancy score, $\tau = -0.185$ $BF_{10} = 0.857$.

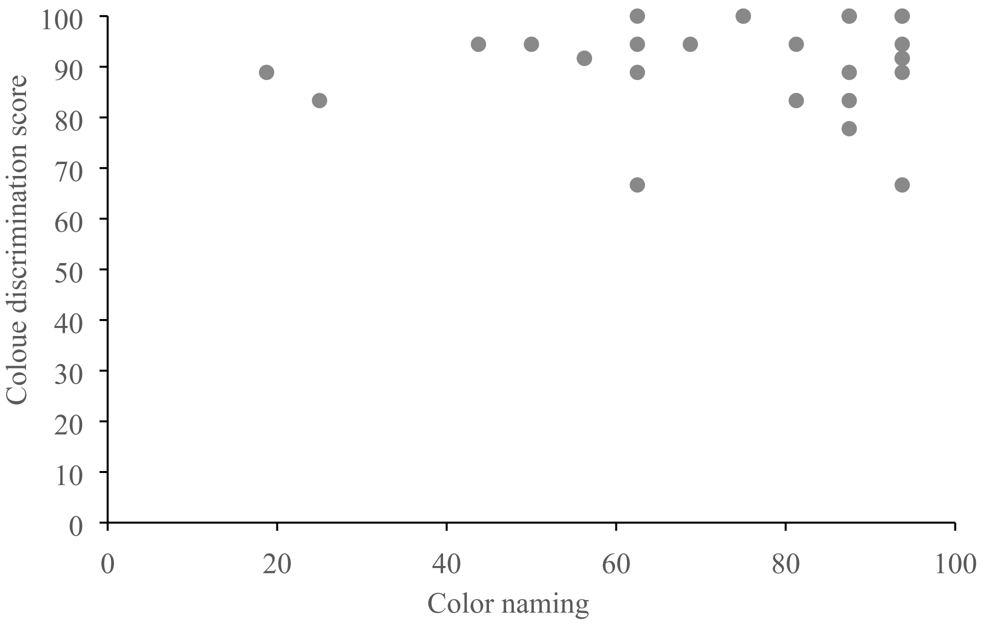

Figure 4. Color discrimination score and color naming ability for each participant. There is no relationship between the two variables. N = 42.

## 5. Discussion

This study investigated the relationship between color constancy maturity and color-term knowledge, in three to four year olds. We found a positive relationship between color-term knowledge and color constancy. We found no relationship between color-term knowledge and color discrimination, nor between age and color constancy. We can be confident that the relationship between color constancy and color-term learning is not due to children's ability to understand the task, general task demands, or their age, since children's color-term knowledge does not relate to their ability to do the task when there is no illumination change and older children were not more color constant. These findings suggest that children who have more mature color constancy also tend to know more color words and vice versa. This may be because immature color constancy makes it challenging to learn color terms by mapping them to colored objects (e.g., yellow banana), as the color of the object would change with an illumination shift. Furthermore, color-term knowledge may help to 'anchor' the representation of color during illuminant changes, via top-down influence of color naming (Witzel, Maule, & Franklin, 2013; Witzel, Sanchez-Walker, et al., 2013).

This study demonstrates that two to four year olds do not yet have fully mature color constancy, and that individual variation in constancy may impact the development of other cognitive domains, such as language learning. Previous studies have found that colors that are named most consistently across observers are also named most consistently across illumination changes, when participants were asked to sort colors into categories under different illuminations in adults (Olkkonen et al., 2009, 2010), and in three to four year olds (Witzel, Sanchez-Walker, et al., 2013). Our findings provide the first clear support for a relationship between perceptual constancy and language learning, when constancy is measured with a task that is independent of naming.

In line with previous literature on color-term acquisition (Bornstein, 1985; Franklin, 2006; Kowalski & Zimiles, 2006; Soja, 1994), our findings demonstrate that even at three to four years old, children are still acquiring mature color-term usage. In the current study, some children could name 100% of the eight basic chromatic color terms whereas, some children could name fewer than 20%. We did not find a correlation between color-term knowledge and age, which suggests that other factors are driving the individual differences in color-term knowledge. This finding may seem surprising given that children acquire language as they develop, and therefore it might be expected that older children have greater color-term knowledge than younger children. However, the absence of a correlation between age and color-term knowledge likely reflects the wide range of individual differences in rate of color-term acquisition, as found in previous studies (e.g. Kowalski & Zimiles, 2006).

Color constancy has been documented in a wide range of animals, including non-human primates, honeybees and goldfish (Olsson, Wilby & Kelber, 2016). The current study potentially suggests that color constancy in humans takes years to reach adult levels. This might be considered surprising given color constancy in non-human animals. However, it is possible that color constancy in non-human animals is based on more rudimentary sensory mechanisms (e.g., low level adaptation) than those that underpin mature color constancy in human adults, where object color names can also contribute. Prior studies have also revealed rudimentary constancy abilities in young infants using preferential looking techniques (Dannemiller, 1989; Dannemiller & Hanko, 1987; Yang et al., 2013). However, with these methods it is difficult to interpret whether the infant is showing a preference for the novel or familiar stimulus (Houston-Price & Nakai, 2004), which means that inferences about constancy can be difficult to make. Furthermore, some of these studies used monitor-rendered simulations of illumination changes rather than real-world stimuli, and it is unclear whether infants would be able to interpret these simulations as an illumination change. In the current study, we adapted a color constancy task from the adult literature (i.e. simultaneous matching; Arend & Reeves, 1986) to be suitable for young children, and we used real-world illumination changes rather than monitor-based simulations. This task can now be used to further chart the development of color constancy and to understand the processes that underlie the development of this important perceptual skill.

The equipment restraints in the experiment meant that the control trials had a different physical setup to the color constancy trials. That is, the color constancy trials were conducted across two sections of the lightbox whereas the control trials were conducted in one section. This may have meant that the color constancy task was more challenging than the illumination task, beyond the change in illumination challenge. However, this was the same across all participants and thus would not affect the finding that children with greater color term knowledge had higher color constancy.

Perceptual constancy is a key feature of our visual system, and investigation into constancy is informative for engineering solutions, such as in computer vision (Vazquez-Corral, Vanrell, Baldrich, & Tous, 2012) and machine learning (Agarwal, Gribok, & Abidi, 2007; Gouko & Kobayashi, 2010). Therefore, further work into the development of constancy would be highly informative for many areas. We know that there is individual variability in the constancy abilities of adults (Allen, Beilock, & Shevell, 2012; Allen & Shevell, 2012), further research into individual differences during development could shed light on what makes a “constant brain”. Future work could also adopt methodological innovations from the adult constancy literature. For example, adult color constancy studies have used tunable LED lighting systems to illuminate a room and test constancy over changes in illumination (Pearce, 2015; Radonjic et al., 2016). This is far more realistic than screen-based simulations of illumination change, yet still precisely controlled. This method could be used adapted into a realistic, child-appropriate measure of color constancy, for example a task involving finding and retrieving objects in the illuminated room.

In conclusion, there is a positive relationship between the maturity of color constancy and color-term knowledge in young children. This demonstrate the impact of perceptual constancy on the maturation of other cognitive domains during childhood, such as language learning. It also raises questions about the development of color constancy through children's interaction with their visual environment.

## Competing interests

The authors declare no competing interests.

## Acknowledgements

We thank the participants and their parents. This work was supported by a European Research Council Grant project, "CATEGORIES", 283605.

Abrams, A. B., Hillis, J. M., & Brainard, D. H. (2007). The relation between color discrimination and color constancy: when is optimal adaptation task dependent? *Neural Computation*, *19*(10), 2610–2637. http://doi.org/10.1162/neco.2007.19.10.2610

Agarwal, V., Gribok, A. V., & Abidi, M. A. (2007). Machine learning approach to color constancy. *Neural Networks*, *20*(5), 559–563. http://doi.org/10.1016/j.neunet.2007.02.004

Allen, E. C., Beilock, S. L., & Shevell, S. K. (2012). Individual differences in color constancy are related to working memory, *29*(2), 1–15.

Allen, E. C., & Shevell, S. K. (2012). Working memory is related to perceptual processing: A case from color perception, 1–14. http://doi.org/10.1037/a0023257.Working

Arend, L., & Reeves, A. (1986). Simultaneous color constancy. *Journal of the Optical Society of America. A, Optics and Image Science*, *3*(10), 1743–51.

Bornstein, M. H. (1985). On the development of color naming in young children: Data and theory. *Brain and Language*, *26*(1), 72–93. http://doi.org/10.1016/0093-934X(85)90029-X

Dannemiller, J. L. (1989). A test of color constancy in 9- and 20-week-old human infants following simulated illuminant changes. *Developmental Psychology*, *25*(2), 171–184. http://doi.org/10.1037//0012-1649.25.2.171

Dannemiller, J. L., & Hanko, S. A. (1987). A test of color constancy in 4-month-old human infants. *Journal of Experimental Child Psychology*, *44*(2), 255–67.

Datasheet: HP DesignJet Z3200 Photo Printer series. (2017). Retrieved from http://www8.hp.com/h20195/V2/GetPDF.aspx/4aa4-4150eee

Dienes, Z. (2014). Using Bayes to get the most out of non-significant results. *Frontiers in Psychology*, *5*(July), 781. http://doi.org/10.3389/fpsyg.2014.00781

Foster, D. H. (2011). Color constancy. *Vision Research*, *51*(7), 674–700. http://doi.org/10.1016/j.visres.2010.09.006

Foster, D. H., Amano, K., & Nascimento, S. M. C. (2006). Color constancy in natural scenes explained by global image statistics. *Visual Neuroscience*, *23*(3–4), 341–349. http://doi.org/10.1017/S0952523806233455.Color

Franklin, A. (2006). Constraints on children's color term acquisition. *Journal of Experimental Child Psychology*, *94*(4), 322–327. http://doi.org/10.1016/j.jecp.2006.02.003

Franklin, A., Clifford, A., Williamson, E., & Davies, I. (2005). Color term knowledge does not affect categorical perception of color in toddlers. *Journal of Experimental Child Psychology*, *90*(2), 114–141. http://doi.org/10.1016/j.jecp.2004.10.001

Garrigan, P., & Kellman, P. J. (2008). Perceptual learning depends on perceptual constancy. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(6), 2248–53. http://doi.org/10.1073/pnas.0711878105

Gouko, M., & Kobayashi, Y. (2010). State Representation with Perceptual Constancy Based on Active Motion (pp. 100–109). http://doi.org/10.1007/978-3-642-17248-9_11

Hansen, T., Walter, S., & Gegenfurtner, K. R. (2007). Effects of spatial and temporal context on color categories and color constancy. *Journal of Vision*, *7*(4), 1–15. http://doi.org/10.1167/7.4.2.Introduction

Houston-Price, C., & Nakai, S. (2004). Distinguishing novelty and familiarity effects in infant preference procedures. *Infant and Child Development*, *13*(4), 341–348. http://doi.org/10.1002/icd.364

Kowalski, K., & Zimiles, H. (2006). The relation between children's conceptual functioning with color and color term acquisition. *Journal of Experimental Child Psychology*, *94*(4), 301–21. http://doi.org/10.1016/j.jecp.2005.12.001

Kraft, J. M., & Brainard, D. H. (1999). Mechanisms of color constancy under nearly natural viewing. *Proceedings of the National Academy of Sciences of the United States of America*, *96*(1), 307–12.

Logvinenko, A. D. (2009). An object-color space. *Journal of Vision*, *9*(11), 5–5. http://doi.org/10.1167/9.11.5

Logvinenko, A. D., & Tokunaga, R. (2011). Color Constancy as Measured by Least Dissimilar Matching. *Seeing and Perceiving*, *24*, 407–452. http://doi.org/10.1163/187847511X588746

Mollon, J. D., Bosten, J. M., Peterzell, D. H., & Webster, M. A. (2017). Individual differences in visual science: What can be learned and what is good experimental practice? *Vision Research*. http://doi.org/10.1016/j.visres.2017.11.001

Olkkonen, M., Hansen, T., & Gegenfurtner, K. R. (2009). Categorical color constancy for simulated surfaces. *Journal of Vision*, *9*(12), 6–6. http://doi.org/10.1167/9.12.6

Olkkonen, M., Witzel, C., Hansen, T., & Gegenfurtner, K. R. (2010). Categorical color constancy for real surfaces. *Journal of Vision*, *10*(9), 16–16. http://doi.org/10.1167/10.9.16

Olsson, P., Wilby, D., & Kelber, A. (2016). Quantitative studies of animal colour constancy:

using the chicken as model. *Proceedings of the Royal Society B: Biological Sciences*, *283*(1830), 20160411.

Pearce, B. (2015). *The asymmetries of color constancy as determined through illumination discrimination using tuneable LED light sources*. Newcastle University.

Pereverzeva, M., & Teller, D. Y. (2009). Simultaneous Color Contrast in 4-Month-Old Infants. *Perception*, *38*(1), 30–43. http://doi.org/10.1068/p6098

Pernet, C. R., Wilcox, R., & Rousselet, G. A. (2013). Robust Correlation Analyses: False Positive and Power Validation Using a New Open Source Matlab Toolbox. *Frontiers in Psychology*, *3*(January), 1–18. http://doi.org/10.3389/fpsyg.2012.00606

Pitchford, N. J., & Mullen, K. T. (2002). Is the Acquisition of Basic-Color Terms in Young Children Constrained? *Perception*, *31*(11), 1349–1370. http://doi.org/10.1068/p3405

Radonjic, A., Aston, S., Krieger, A., Cottaris, N. P., Brainard, D. H., & Hurlbert, A. C. (2016). Illumination discrimination in real and simulated scenes ′, *16*, 1–18. http://doi.org/10.1167/16.11.2.doi

Smithson, H. E. (2005). Sensory, computational and cognitive components of human color constancy. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *360*(1458), 1329–1346. http://doi.org/10.1098/rstb.2005.1633

Soja, N. N. (1994). Young Children's Concept of Color and Its Relation to the Acquisition of Color Words. *Child Development*, *65*(3), 918–937.

van Doorn, J., Ly, A., Marsman, M., & Wagenmakers, E.-J. (2016). Bayesian Inference for Kendall's Rank Correlation Coefficient. *The American Statistician*, 0–0. http://doi.org/10.1080/00031305.2016.1264998

Vazquez-Corral, J., Vanrell, M., Baldrich, R., & Tous, F. (2012). Color Constancy by Category Correlation. *IEEE Transactions on Image Processing*, *21*(4), 1997–2007. http://doi.org/10.1109/TIP.2011.2171353

Wagner, K., Dobkins, K., & Barner, D. (2013). Slow mapping: color word learning as a gradual inductive process. *Cognition*, *127*(3), 307–17. http://doi.org/10.1016/j.cognition.2013.01.010

Walsh, V., & Kulikowski, J. (Eds.). (1998). *Perceptual constancy: Why things look as they do.* Cambridge University Press.

Winkler, A. D., Spillmann, L., Werner, J. S., & Webster, M. A. (2015). Asymmetries in blue–yellow color perception and in the color of 'the dress.' *Current Biology*, *25*(13), R547–R548. http://doi.org/10.1016/j.cub.2015.05.004

Witzel, C., Maule, J., & Franklin, A. (2013). Focal colors as perceptual anchors of color categories. *Journal of Vision*, VSS2013 abstracts.

Witzel, C., Racey, C., & O'Regan, J. K. (2017). The most reasonable explanation of "the dress": Implicit assumptions about illumination. *Journal of Vision*, *17*(2), 1. http://doi.org/10.1167/17.2.1

Witzel, C., Sanchez-Walker, E., & Franklin, A. (2013). The development of categorical color constancy. *Perception*, *42*(ECVP Abstract Supplement), 19.

Witzel, C., van Alphen, C., Godau, C., & O'Regan, J. K. (2016). Uncertainty of sensory signal explains variation of color constancy. *Journal of Vision*, *16*(15), 8. http://doi.org/10.1167/16.15.8

Yang, J., Kanazawa, S., Yamaguchi, M. K., & Kuriki, I. (2013). Investigation of color constancy in 4.5-month-old infants under a strict control of luminance contrast for individual participants. *Journal of Experimental Child Psychology*, *115*(1), 126–36. http://doi.org/10.1016/j.jecp.2012.11.013