Received: 2 March 2021 | Revised: 15 May 2022 | Accepted: 12 July 2022

DOI: 10.1111/desc.13306

Developmental Science WILEY

PAPER

Developmental changes in colour constancy in a naturalistic object selection task

Rebecca Wedge-Roberts[1] | Stacey Aston[1] | Ulrik Beierholm[1] | Robert Kentridge[1,2] | Anya Hurlbert[3] | Marko Nardini[1] | Maria Olkkonen[1,4]

[1]Department of Psychology, Durham University, Durham, UK

[2]Azrieli Programme in Brain, Mind & Consciousnesses, Canadian Institute for Advanced Research, Toronto, Canada

[3]Neuroscience, Institute of Biosciences, Newcastle University, Newcastle upon Tyne, UK

[4]Department of Psychology and Logopedics, Faculty of Medicine, University of Helsinki, Helsinki, Finland

Correspondence
Rebecca Wedge-Roberts, Department of Psychology, Durham University, Durham, UK.
Email: rebecca.j.wedge-roberts@durham.ac.uk

Funding information
Leverhulme Trust, Grant/Award Number: RPG-2017-097

Abstract

When the illumination falling on a surface change, so does the reflected light. Despite this, adult observers are good at perceiving surfaces as relatively unchanging—an ability termed colour constancy. Very few studies have investigated colour constancy in infants, and even fewer in children. Here we asked whether there is a difference in colour constancy between children and adults; what the developmental trajectory is between six and 11 years; and whether the pattern of constancy across illuminations and reflectances differs between adults and children. To this end, we developed a novel, child-friendly computer-based object selection task. In this, observers saw a dragon's favourite sweet under a neutral illumination and picked the matching sweet from an array of eight seen under a different illumination (blue, yellow, red, or green). This set contained a reflectance match (colour constant; perfect performance) and a tristimulus match (colour inconstant). We ran two experiments, with two-dimensional scenes in one and three-dimensional renderings in the other. Twenty-six adults and 33 children took part in the first experiment; 26 adults and 40 children took part in the second. Children performed better than adults on this task, and their performance decreased with age in both experiments. We found differences across illuminations and sweets, but a similar pattern across both age groups. This unexpected finding might reflect a real decrease in colour constancy from childhood to adulthood, explained by developmental changes in the perceptual and cognitive mechanisms underpinning colour constancy, or differences in task strategies between children and adults.

KEYWORDS
colour constancy, daylight prior, development, perception, psychophysics

Highlights

- Six- to 11-year-old children demonstrated better performance than adults on a colour constancy object selection task.
- Performance decreased with age over childhood.

- These findings may indicate development of cognitive strategies used to overcome automatic colour constancy mechanisms.

1 | INTRODUCTION

The light reflected from a surface depends on both the surface reflectance and incident illumination. Colour constancy is the ability to judge surfaces as relatively invariant under different illuminations. This is crucial for recognising material properties of objects—without it, objects would appear to change colour radically in different environments. Empirical measurements of colour constancy find varying levels depending on the method of measurement (including instructions; L. E. Arend et al., 1991), surfaces and illuminations used, and the individuals tested (see Foster, 2011; H. E. Smithson, 2005, for a review). Typical methods of measuring colour constancy include achromatic setting (D. Brainard, 1998) in which observers adjust a patch to appear achromatic; asymmetric matching (D. H. Brainard et al., 1997) in which observers adjust the colour of a test patch to match a target patch under a different illumination; and object selection (Radonjić et al., 2015a) in which observers select one object from several to match a target seen under a different illuminant. Object selection tasks are more naturalistic, and generally elicit the highest levels of colour constancy (Radonjić & Brainard, 2016).

An almost infinite variety of illumination and surface reflectance combinations may give rise to the same light signal at the eye. Bayesian models (D. H. Brainard & Freeman, 1997; D. H. Brainard et al., 2006; Olkkonen et al., 2016) propose that the visual system narrows down the possibilities and estimates the actual physical stimulus by learning the most likely combinations. Prior experience, acquired during development, is arguably necessary to learn these mappings (Beau Lotto, 2004). Developmental studies into colour constancy are therefore needed to empirically determine the relationship between experience and perception. Whilst animal studies have shown that experience with broadband illuminations is essential for colour constancy (Sugita, 2004), little is yet known about the development of colour constancy in children.

Many low-level systems necessary for colour vision develop during the first months of life (see Brown, 1990, for a review), including macular pigment density (Bone et al., 1988), cone contrast sensitivity and acuity (Concetta Morrone et al., 1990). Two-month-old infants can discriminate chromatic from achromatic surfaces (Peeples & Teller, 1975), and 4-month-old infants have colour categories similar to those of adults (Skelton et al., 2017). Additionally, young infants exhibit other complex aspects of perception including transparency perception (Johnson & Aslin, 2000) and certain visual illusions (Yang et al., 2010). However, other aspects continue to develop, including detecting colour-defined form, which is not adult-like until teenage years (Hollants-Gilhuijs et al., 1998). This ability is likely to depend on processes independent of those giving rise to the experience of colour.

Few studies have investigated colour constancy in children. Dannemiller and Hanko (1987), Dannemiller (1989), and Yang et al. (2013) and (2015) used a preferential-looking paradigm to study colour constancy in 3- to 7-month-old infants and found rudimentary colour constancy by 4.5 months. Similarly, Hui Lin Chien et al. (2006) and Kavšek (2011) found looking behaviour consistent with lightness constancy in 4- and 6-month-old infants, respectively. Using an object selection task with a limited set of targets and a coarse measure of colour constancy, Rogers et al. (2020) found no relationship between age and performance in 2- to 4-year-old toddlers. Although there were large individual differences, no toddler performed as well as the adults.

On the other hand, Katz (2013) anecdotally reports that children aged 8–15 years may have equal or superior colour constancy to that of adults. However, this report is not supported by explicit data, so should be interpreted with caution. Similarly, Beck (1966) found no difference in lightness constancy between 5-year-old children and adults when judging a single chip's lightness.

The development of size and shape constancy has received more experimental attention. Some studies report that size constancy improves with age up to 7–9 years (Brislin & Leibowitz, 1970; Granrud & Schmechel, 2006; Kavšek & Granrud, 2012), although others find adult-like levels in 3-year-olds (Tronick & Hershenson, 1979). This discrepancy may be explained by differences in instructions and strategies (Granrud, 2009; Rapoport, 1967). Kaess et al. (1974) found increasing shape constancy with age up to 19 years, whereas others found no effect of age (Field & Collins, 1977), or *decreasing* constancy with age, but only for small viewing distances (3 vs. 15 feet; Meneghini & Leibowitz, 1967). In shape and size constancy, the viewing distance matters as different mechanisms are involved in judging the depth and distance of near (e.g., stereopsis) versus far objects.

According to a popular hypothesis, adults' colour constancy may be optimised for daylight illuminations, which vary in appearance from yellowish to blueish (Hernández-Andrés et al., 1999; Judd et al., 1964; Spitschan et al., 2017). Through experience, observers may have developed a "daylight prior," in line with Bayesian models (D. H. Brainard et al., 2006). The finding of higher constancy for scenes illuminated by blueish daylights than non-daylights (Delahunt & Brainard, 2004; Pearce et al., 2014; Weiss et al., 2017) partially supports this hypothesis. Whilst the development of a daylight prior has not been investigated, studies have found other perceptual priors developing between four and 12 years (Chambers et al., 2018; Stone, 2011; Thomas et al., 2010; Yonas & Hagen, 1973).

In summary, there is limited research into the development of colour constancy, with no studies investigating the developmental trajectory of colour constancy with age in children over 4 years, and few using a comparable task for different age groups. This gap may be because many methods of measuring colour constancy in adults (achromatic matching, asymmetric matching) are inappropriate for young children, either requiring observers to remember a colour, and/or perform fine-tuned matches which require a long attention span and fine motor skills. Therefore, we have adapted an object selection task from Radonjić et al. (2015b), related to developmental work from Rogers et al. (2020), to develop a novel, child-friendly measure of colour constancy with no memory demands, and no need to make explicit judgements about colours. We used this task to better understand the role of development in perceptual constancies. Specifically, we aimed to answer three research questions: (1) is there an overall difference between 6- to 11-year-old children and adult observers' colour constancy? (2) Is there a developmental trajectory in colour constancy from 6 to 11 years? and (3) Does the pattern of colour constancy across surfaces and illuminations, such as effects driven by a daylight prior, differ between adults and children? A developing daylight prior would be demonstrated by a larger asymmetry between daylights and non-daylights for adults and older children than in younger children. We focussed on the age range (6–11 years) in which developmental changes in size and shape constancy occur. In the first experiment, we aimed to measure colour constancy in children and adults using simple two-dimensional stimuli, comparable to those used in much previous research. In the second experiment, we measured colour constancy with more realistic three-dimensional rendered stimuli, to ask whether the findings apply to scenes which more closely resemble the real world.

2 | EXPERIMENT 1

We developed a novel computer-based object selection task, which children would find engaging, based on the materials and methods of Radonjić et al. (2015a), using simple two-dimensional scenes. This involved finding a dragon's favourite "sweet" from a set seen under one illumination to match a target sweet seen under a neutral illumination. This experiment allowed us to measure baseline performance with minimal cues, and to confirm whether the task was appropriate for children. The use of a computer allowed precise manipulation of reflectances and illuminants. Six- to 11-year-old children, and adults participated, with scenes simulated under either daylight or non-daylight illuminants to determine whether overall colour constancy, or the pattern across illuminations, differs between children and adults.

2.1 | Method

2.1.1 | Observers

Twenty-six adults (mean age = 22.57 years, SD = 6.51; five male, 21 female) and 33 children aged between 6 and 11 years (11 male, 22 female) participated. This age range was chosen based on previous findings for constancy in other visual domains, and because pilot experiments found that younger children were unable to use the equipment. Informed consent was given by adults and parents of children. The children assented to take part and were repeatedly asked during the experimental session if they were happy to continue. All observers were screened for colour vision deficiencies using Ishihara plates (38 plates edition) (Ishihara, 2006). Two children had scores outside the normal range (more than two errors), so their data were not included in analyses. Adults were either psychology undergraduate students who took part for course credits, or paid volunteers. Children were rewarded with a small prize at the session's end. All observers had normal or corrected to normal visual acuity.

2.1.2 | Materials and apparatus

Scenes were presented on an ASUS PA382Q 23″ monitor with 10 bits per channel, controlled by an Nvidia quadro k600 graphics card. The monitor was characterized with a Konica Minolta CS-2000 spectroradiometer and display linearization and colour conversions were achieved with standard methods (D. H. Brainard et al., 2002). The computer presenting the stimuli and the testing room were changed after the first 20 participants (five adults and 15 children) due to availability of laboratory space and occasional computer malfunction. The monitor was recalibrated after moving, to maintain stimulus properties. To check this did not affect results, we ran a linear mixed effects (lme) model including computer setup as a predictor of colour constancy indices and found no significant main effect of computer setup, and no significant interaction with age group. For all conversions to CIE L*u*v*, the white point was defined as the chromaticity of the neutral illumination (D57), at a luminance of 60 cd/m^2 (Yxy = 60, 0.328, 0.344). Observers sat in the dark, approximately 60 cm from the monitor, with free head movement. Experiments were run on MATLAB (The Math Works, Inc., 2018), using functions from the Psychophysics Toolbox (D. H. Brainard, 1997; Kleiner et al., 2007; Pelli, 1997).

2.1.3 | Stimuli

Each experimental scene filled the monitor, subtending 46 × 27 degrees of visual angle. The scenes consisted of two halves (each 23 × 27 degrees); a square target "sweet" was shown against a neutral background on the left-hand side, and eight competitors were shown against a "chromatic" background on the right-hand side (Figure 1). All nine sweets were identical in shape and size (129 pixels^2; 3.3 degrees of visual angle). The target sweet was centred in the left half. The eight competitor sweets were aligned in two equally spaced rows of four, symmetric above and below a horizontal line at the vertical middle of the screen. Numerals were added above the top row and below the bottom row of competitors so that children unable to use the mouse could say which number to pick.

The backgrounds were designed to simulate different illuminations, so will hereafter be referred to as illuminations. The neutral illumination had a CCT of 5698K (D57), close to the peak chromaticity of the measured daylight distribution from 30 natural scenes (Nascimento et al., 2016). Four chromatic illuminations were used: two along the

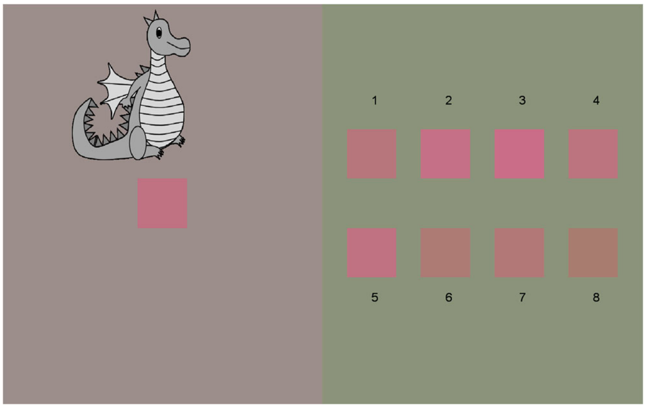

FIGURE 1 Example of a scene with a rose sweet under green illumination. Note that Derek the dragon was not present in scenes in the main experiment. Sweet 5 is the tristimulus match and sweet 6 is the reflectance match

TABLE 1 Co-ordinates of the illuminations used in CIEYxy and CIEL'u'v' colour spaces

Illumination	X	Y	u'	v'
Neutral (D57)	0.328	0.344	0.203	0.478
Blue	0.297	0.314	0.193	0.458
Yellow	0.364	0.372	0.216	0.497
Red	0.329	0.310	0.217	0.460
Green	0.327	0.382	0.189	0.496

daylight locus in the blueish and yellowish directions, and two perpendicular to the daylight locus at CCT 5698K (in u*v*), which appear reddish and greenish. In all conditions the background was scaled to have a luminance of 60 cd/m^2. At this luminance, all chromatic illuminations were $30\Delta E_{u*v*}$ from D57. Illumination chromaticities are given in Table 1.

We used four target reflectances which appeared grey, green, rose, and teal under neutral illumination (Radonjić et al., 2015b). For each target and illumination, we generated eight competitors, equally spaced on a line in u'v', which contained a tristimulus match (T; this has the same chromaticity that the target sweet would have under the neutral illumination, therefore indicating no constancy) and a reflectance match (R; this has the chromaticity that the target sweet would have under the different illumination, therefore indicating perfect constancy). An example of the competitors' chromaticities is shown in Figure 2. These range from an over-constancy match, which observers would pick if they over-corrected for the illumination (with a chromaticity beyond R), to under-constancy matches (with chromaticities beyond T). Full details of how these were generated, and a description of tristimulus and reflectance matches, is given in the Supplementary Material. Although the competitors should have slightly different luminances due to the interaction between the illumination and reflectance, for simplicity we fixed the luminance of each competitor at 50 cd/m^2 (The difference in luminance between competitors before fixing was small, with a maximum discrepancy between T and R of 1.8 cd/m^2).

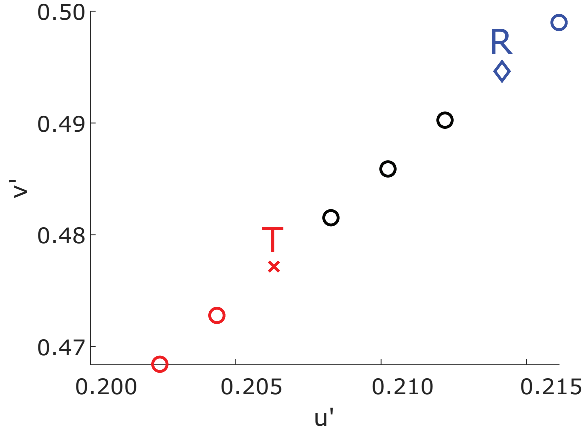

FIGURE 2 Example chromaticity of eight competitors in u'v'. Red X represents the tristimulus match, T; blue diamond represents the reflectance match, R; the black Os are the competitors in between T and R; red Os are underconstant and blue O is overconstant

In addition to these sixteen experimental scenes (four illuminations × four sweet colours), there were two control conditions for each target sweet, in which both sides of the screen were illuminated by D57, to test for any internal biases—such as a preference to pick more saturated sweets. In these scenes, the background was uniform across the screen. Full details of how the eight competitors in these scenes were generated are given in the Supplementary Material.

Overall, there were 24 conditions (six illuminations × four sweet colours). Each observer was presented with each condition 10 times, for a total of 240 trials, with the eight competitors positioned in a random order on each trial.

2.1.4 | Task

The observers' task was to feed "Derek the dragon" his favourite sweets. They saw his "favourite sweet"—the target sweet—on the left-hand side of the screen under neutral illumination (D57). From the eight competitors on the right-hand side of the screen, observers used a mouse to click on the sweet they thought was his favourite. The instructions (see Supplementary Material) did not mention colour, or explain that the backgrounds simulated illumination changes. Therefore, no specific strategy was encouraged. Children who could not use a mouse spoke the number of the sweet to select, and the experimenter clicked on it. The selected sweet was then indicated by a small, uniform increase in size. Observers were allowed to change their selection as often as required, with unlimited time. To feed Derek the sweet, observers pressed the space bar.

2.1.5 | Procedure

All adults completed all 24 conditions in a single session, except the first five who did not complete the red-green control condition, as

TABLE 2 Mean and SD number of blocks completed by children of each age (out of five for first 15 children and out of six for remaining children)

Age (years)	Mean number of blocks completed	SD blocks completed
6	3.89	1.90
7	5	1.10
8	5.17	1.17
9	5	0.82
10	4.4	1.52
11	5.67	0.58

this was added later. Fifteen children were also tested before the red-green control condition was added. Children completed as many trials as they could in a single session before becoming fatigued. Generally, the older children were more likely to complete all conditions than the younger children. Over all 33 children, 28 blocks (out of a possible 183) were not completed due to fatigue. The breakdown of number of blocks completed by age is shown in Table 2. The testing sessions lasted approximately an hour, with large variability between observers.

The full sessions were split into six blocks—one for each illumination. Within a block, each sweet was shown 10 times, resulting in 40 trials per block. Trial and block orders were randomised for each observer.

Each block began with two minutes of dark adaptation. In the first block immediately following the adaptation period, the instructions were presented on the screen by an animation of Derek the Dragon. The illumination of the first block was used in the example scene in the instructions. For children who struggled to read, the experimenter read out the instructions, and pointed to the relevant parts of the scene. Following the instructions, the first trial was presented.

Children were given a star chart at the beginning of the session and were rewarded with a star sticker after a set of six, seven, or eight trials, with the set length randomised each time. Between every trial, observers saw a black screen containing an energy bar, alongside text telling them to continue. The energy bar was filled incrementally, either 1/6, 1/7 or 1/8 after each trial depending on the set length. When the bar was full a black screen containing a large silver star was shown, alongside an animation of Derek talking with the text "Thank you for feeding me, you have earned a star!." This reward was designed to keep children motivated, without giving any meaningful feedback.

At the end of each block, an animation of Derek appeared alongside text saying "Thank you for feeding me so many sweets!." Between each block, observers were asked if they wanted to play the next round.

2.1.6 | Data analysis

We developed two criteria to exclude random responses. This was done for each illumination condition, but not separated by sweet colour, as sufficient responses were needed to test for potential randomness. There were, therefore, 40 responses (each a number between one and eight indicating the competitor selected on that trial) to assess. A uniform and a Gaussian distribution were fitted to these responses. Sets of responses where the Bayesian information criterion (BIC) of the Gaussian fit was larger than, or within six units from, the BIC for the uniform fit were deemed as equally or better fit by the uniform than the Gaussian distribution, and were excluded from analysis. This excluded three conditions, all from children. Additionally, any set of responses with SD greater than 2.32 (the SD of 40 responses uniformly distributed amongst eight competitors) were excluded. This excluded five conditions, all from children. Due to overlap between the exclusion criteria, five conditions in total were excluded, from four children, representing 3.23% of conditions completed by children. No data were excluded from adults.

To analyse the remaining data, all responses were first converted to the associated chromaticities of the competitors in CIE u′v′. T had the same chromaticity as the target sweet, despite the change in illumination, and therefore indicated no colour constancy, while R simulated the target sweet under the chromatic illumination, and therefore indicated perfect constancy. The observer's colour constancy index (CCI) was calculated for each trial as:

$$CCI = 1 - \frac{a}{b} \tag{1}$$

where a is the signed Euclidean distance, in u′v′, from R to the competitor selected, which was negative for over-constancy matches, and b is the Euclidean distance from R to T. Possible CCIs ranged from −0.25 for under-constancy matches to 1.5 for over-constancy matches, with CCIs of one indicating perfect constancy.

Figure 3 shows an example of how the chromaticity of the sweets selected corresponds to the CCIs.

In order to test our three research questions, we ran lme models in R (R Core Team, 2020) using the lme4 package (Bates et al., 2015). As many children did not complete all 16 conditions and some conditions were excluded, it was not possible to conduct an ANOVA. We created two models: one to investigate main effects of age group, illumination, and sweet colour:

$$CCI \sim ageGroup + block + illumination + (1|observer) \tag{2}$$

and a second to test for interactions between these variables

$$CCI \sim ageGroup * block * illumination + (1|observer) \tag{3}$$

2.2 | Results

The mean chromaticities of the sweets selected by adults and children under each of the 16 conditions are shown in Figure 4. These graphs are in the same format as Figure 3a but show all four sweet colours, with the corresponding target sweet shown above. The pattern across

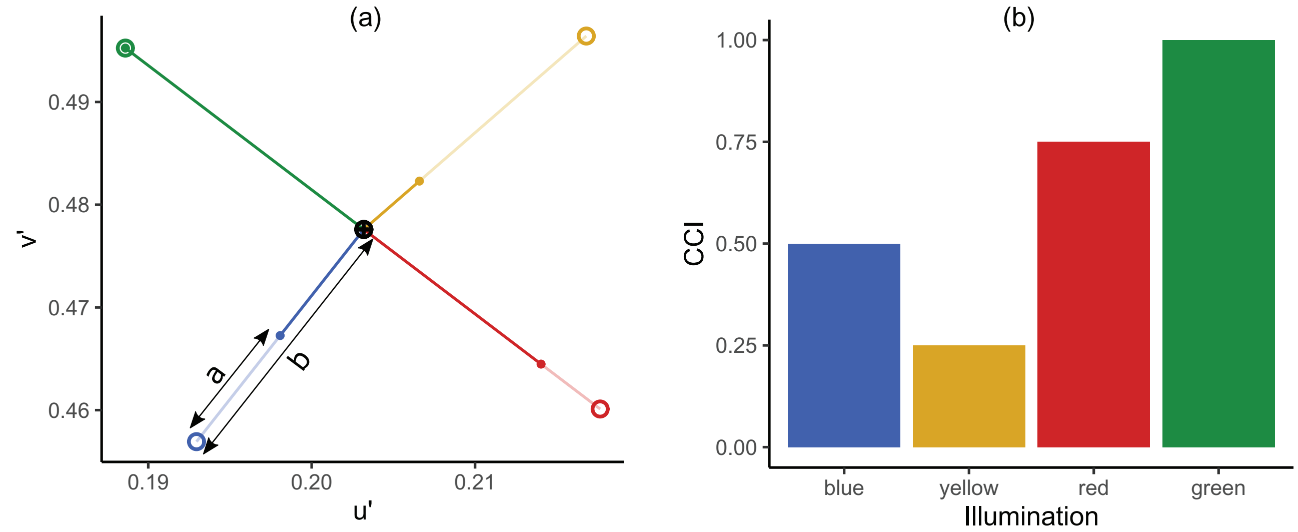

FIGURE 3 (a): Example of an observer's matches for grey sweets in u'v'. The black filled O is the tristimulus match (T). The open coloured Os are the reflectance matches (R) under each of the four illuminants, with the colour of the O corresponding to the colour of the illumination. The filled coloured Os are the observer's matches under each illuminant, with the colour corresponding to the illumination. Pale lines show the whole distance from T to R (b in Equation 1); darker lines show the distance from T to the match. (b): the colour constancy indices associated with the chromaticities in (a). As this observer selected R under green, they had perfect colour constancy under green, but they chose a competitor close to T under yellow, resulting in poorer constancy

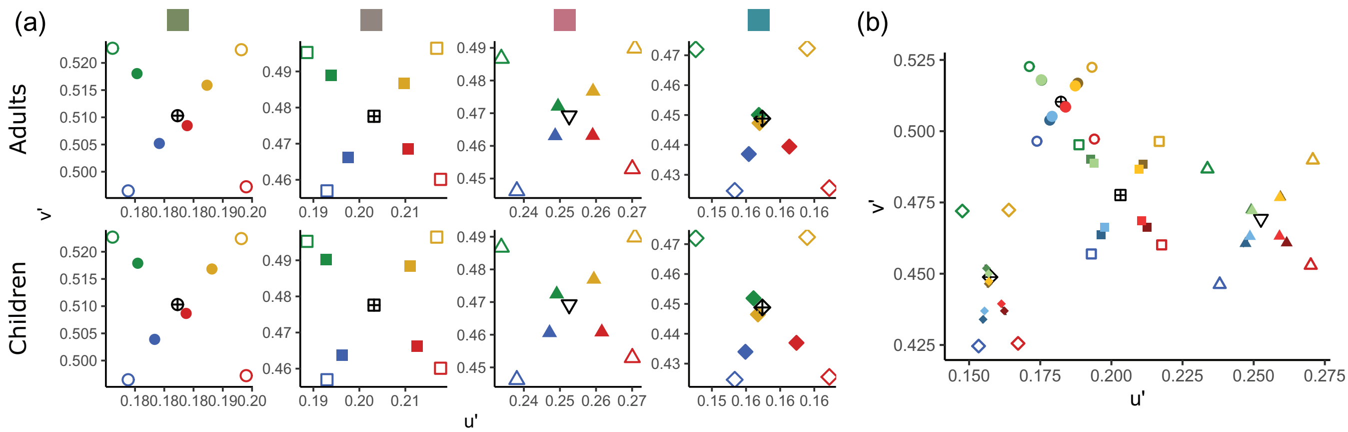

FIGURE 4 (a) Mean chromaticities of sweets selected in u'v', formatted as in Figure 3a for each sweet colour separately, with the shape families representing different sweet colours: Os are green; squares are grey; triangles are rose; diamonds are teal. Adults' results are in the top row and children's are in the bottom row. (b) shows the mean chromaticities selected by both adults (pale symbols) and children (dark symbols) for all sweet colours. In all plots, black symbols are the targets/tristimulus matches, open coloured symbols are reflectance matches, and filled coloured symbols are mean matches across observers. The colour of the symbols indicates the colour of the illumination. Error bars (too small to be visible in many cases) show ± 1 SEM

adults and children is similar, with large differences between sweet colours and illuminations.

In order to meaningfully analyse colour constancy performance, we first needed to ensure that observers could discriminate between the competitors, and accurately match without colour constancy demands. To this end, we analysed the neutral control conditions. Observers performed generally well in these, with a mean $\Delta E_{u'v'}$ (Euclidean distance) from the target of 0.00242, corresponding to a just-noticeable-difference (L*u*v*) of 3.74. In a 2(condition) × 4(sweet) × 2(age group) ANOVA, deviations were significantly larger in the blue-yellow control than the red-green control condition ($p = 0.0152$), with a bias towards yellower competitors for Green, Grey, and Rose sweets, and a bias towards bluer competitors for Teal sweets. Observers had a small bias towards green for Green and Grey sweets, and towards red for Rose and Teal sweets. There was no significant difference between adults and children ($p = 0.603$), and both followed the same pattern.

Our first research question was whether CCIs differed between adults and children. Figure 5a shows that children (dark violins) had,

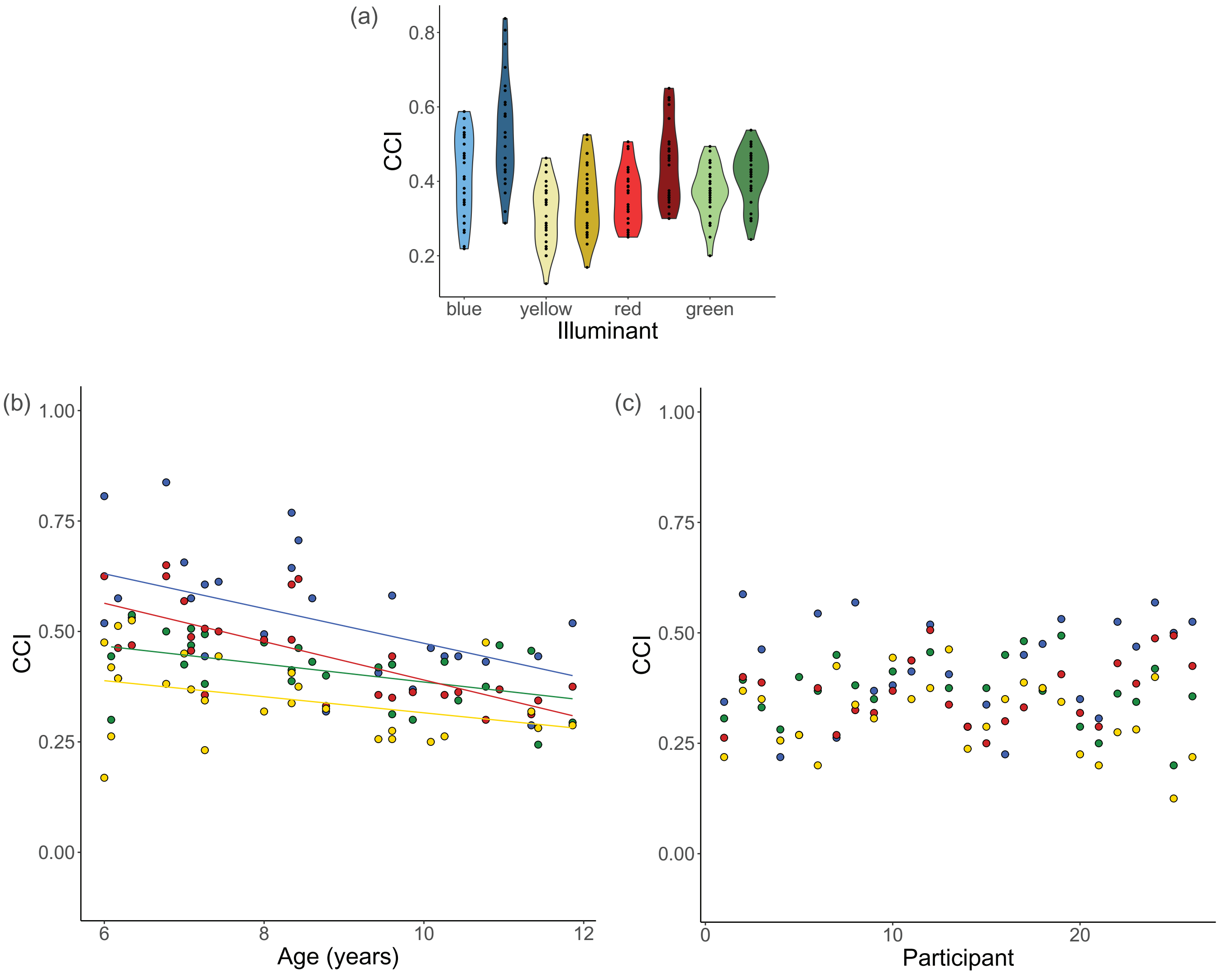

FIGURE 5 (a): Mean colour constancy indices (CCIs) for children (dark bars) and adults (light bars) collapsed across sweet colours, for each illumination condition separately. (b): Age against mean CCI across all sweet colours for each child and illumination condition separately. Colours of points represent illumination colour. Regression lines are shown for each illumination condition. (c): Mean CCI across all sweet colours for each adult and illumination condition. Participants are sorted in ascending age order

on average, higher CCIs (closer to 1) than adults (light violins) under all four illuminants (collapsed across sweet colours). Figure 5 shows the same pattern, comparing individual children (b) with adults (c). In a lme model with age group (children or adults), illumination, and sweet colour as predictors of CCI (Equation 2), the main effect of age group was significant ($p < 0.001$, $t = 9.605$), with higher CCIs in children (mean = 0.436, SD = 0.373) than adults (mean = 0.364, SD = 0.303) (Figure 6).

To determine whether CCIs changed over childhood, addressing our second question, we ran robust regressions on the children's data using the rlm function of the MASS package in R (Venables & Ripley, 2002). We used age as a predictor of CCI for each illumination separately, collapsed across sweet colour and repetitions. All four regressions (Figure 5b) were significant and negative (Blue: $B = -0.0390$, $p = 0.0147$; Yellow: $B = -0.0252$, $p = 0.0054$; Green:

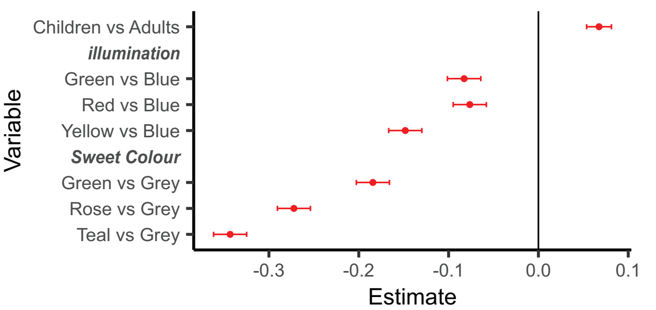

FIGURE 6 Main effects model results. Red circles show estimates; error bars show 95% confidence intervals. Black vertical line at Estimate = 0 reflects non-significance. Estimates greater than 0 indicate the first variable (e.g., "green" in "green vs blue" predicts higher CCIs; estimates less than 0 suggest the variable on the right hand side predicts higher CCIs

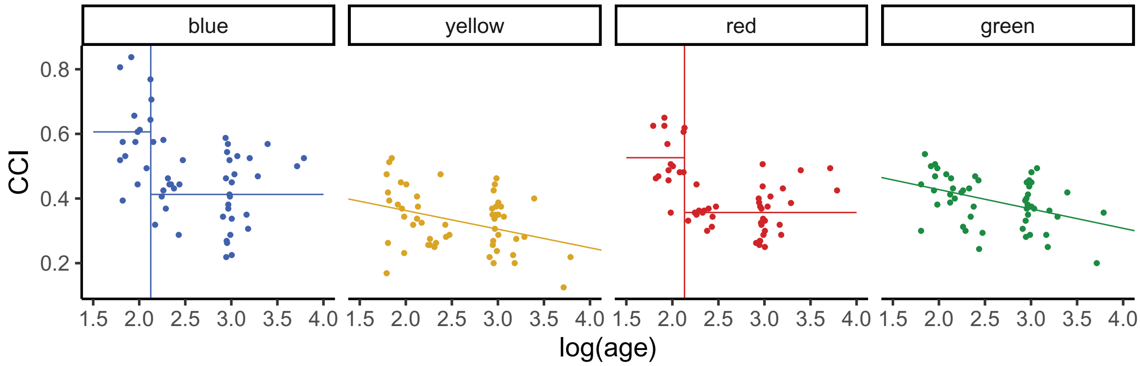

FIGURE 7 CCIs against log(age) with each of the best fitting models for each illumination. For Blue and Red illuminations, the best fitting model is the jump model, with the vertical line indicating the jump and the horizontal lines indicating the mean CCI below and above this age. For Green and Yellow, the best fitting model is linear, with the regression line shown

$B = -0.0241$, $p = 0.0061$; Red: $B = -0.0436$, $p < 0.001$), suggesting that CCIs decrease with age between 6 and 11 years under all four illuminations. B denotes the change in CCI for each year of age; for the Blue illumination, a 7-year-old's CCI is 0.0390 lower than a 6-year-old's, on average.

To determine more specifically the shape of the developmental trajectory, we asked which of four distinct models best fit the CCIs as a function of log(age) for all data (children and adults): linear; jump model, which predicts a step-like change in CCIs; hockey stick model, which predicts a linear change up to a certain age, after which CCIs plateau; and no change with age. We used log(age) to reduce the age gap between adult and child observers. The models were fitted to the data for each illumination condition separately, to determine the interaction between age and illumination. The BIC was determined for each model to determine which best fitted the data, and are all given in the Supplementary Material. The models with the lowest BIC, and therefore the best fits, were the jump model for Blue and Red illuminations, and the linear model for Green and Yellow. For Blue, the jump was at 8.6 years, with the mean CCI 0.609 below this age and 0.422 above. For Red the jump was at 8.4 years, with CCIs higher below this age (mean = 0.526) than above (mean = 0.357). For green, β was -0.0599, and for yellow β was -0.576. These models are plotted alongside the data in Figure 7.

To test whether the pattern of CCIs across illuminations and reflectances differed between adults and children—our third research question—we ran a second lme model to look for interactions (Figure 8), described in detail in the Supplementary Material.

The two lme models show that CCIs differ significantly across illuminations (highest under Blue and lowest under Yellow) and sweet colours (highest for Grey and lowest for Teal), and that there are significant interactions between the two. None of the two-way interactions involving age group are significant, indicating that the patterns across illuminations and reflectances are similar across adults and children. The two significant three-way interactions children|Green$_{sweet}$|Red$_{illumination}$ (p = 0.007); children|Teal$_{sweet}$|Yellow$_{illumination}$ (p = 0.003) indicate that the two-way interactions for Green$_{sweet}$ |Red$_{illumination}$ and Teal$_{sweet}$|Yellow$_{illumination}$ are more pronounced for children than adults.

2.3 | Interim discussion

In this object selection task, we found that 6- to 11-year-old children's colour constancy indices were *higher* than adults' and decreased with age between 6 and 11 years. Although surprising, the result agrees with some previous reports of superior constancy in children (Katz, 2013; Meneghini & Leibowitz, 1967).

As well as the overall change in CCIs with age, we were interested in the pattern across illuminations and reflectances for adults and children. Both adults and children had the highest CCIs under the Blue illumination, possibly due to a daylight prior, in agreement with Delahunt and Brainard (2004). Furthermore, children and adults had the highest CCIs for grey sweets, in line with previous research (Olkkonen et al., 2009). Interestingly, there was a strong interaction between sweet colour and illumination, such that CCIs were generally higher when the sweet's reflectance and illumination had similar chromaticities (e.g., Green sweets under Green illumination). As with the main effects of sweet colour and illumination, this interaction did not vary by age group. Whilst there were few significant interactions involving age group in the lme model, different models best explained the developmental trajectory under Blue and Red illuminants (jump model) compared to Yellow and Green (linear model). The developmental trajectory may therefore vary across illuminations.

This experiment provided a useful way to measure colour constancy in children and adults in a simple, controlled scene, using a novel task. However, the simplicity of the stimuli limits their relevance to colour constancy in real life, as many cues are missing, and observers might not interpret the backgrounds as different illuminations. Therefore, in Experiment 2 we used three-dimensional rendered scenes, to determine whether their additional cues to the illumination geometry would be exploited by children and/or adults.

Developmental Science WILEY

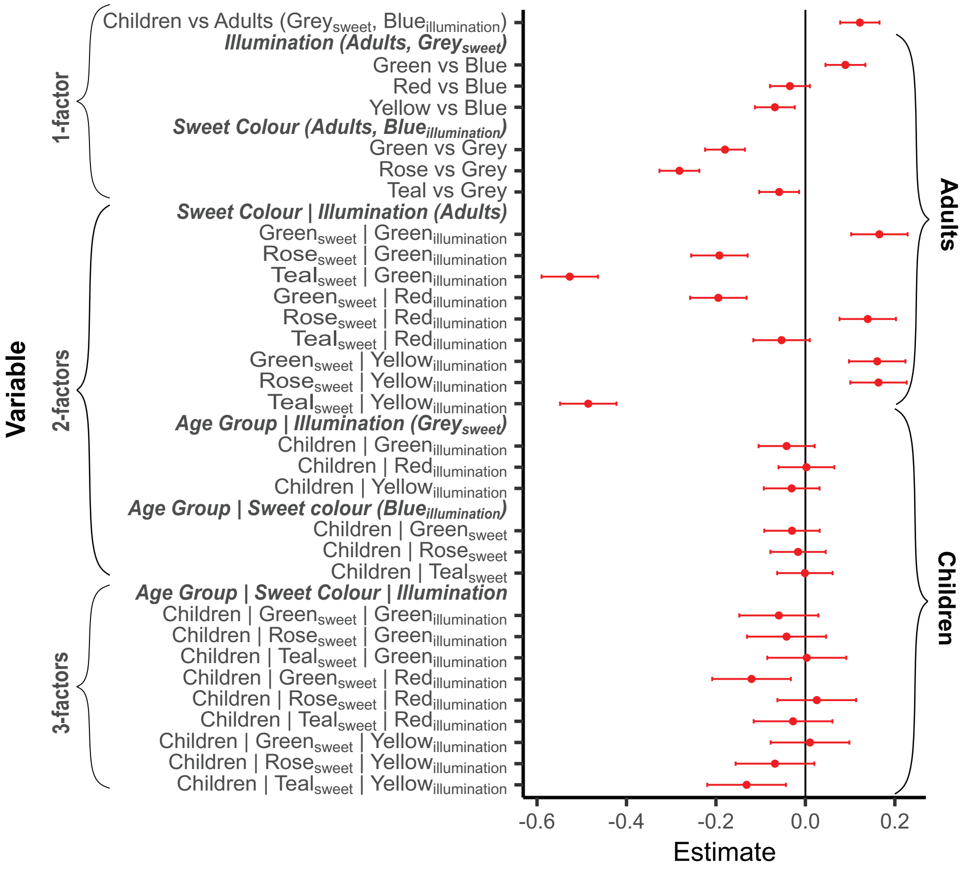

FIGURE 8 Interaction effects model results. Red circles show estimates; error bars show 95% confidence intervals. Bold, italic text is added to divide the different factors. Parentheses show the level of the fixed factors, such that the first effect is a comparison of children compared to adults with grey sweets under the blue illumination. See text for full explanation of results

3 | EXPERIMENT 2

To measure colour constancy in a more realistic environment, in Experiment 2 we used three-dimensional computer rendered stimuli. The task, illuminations, and sweet colours were the same as in Experiment 1.

3.1 | Method

3.1.1 | Observers

A further 26 adults (mean age = 25.2 years, SD = 10.1; seven male, 19 female) and 40 children aged 6–11 years (15 male, 25 female) participated. All observers were naive to the purposes of the study, apart from one adult who was involved in collecting data.

3.1.2 | Materials and apparatus

The same materials were used as in Experiment 1. The computer and testing room were those employed in the second half of Experiment 1.

3.1.3 | Stimuli

An example of the stimuli is shown in Figure 9. The geometry of the three-dimensional scenes was created using Blender (https://www.blender.org/), and they were rendered using Mitsuba (http://www.mitsuba-renderer.org/) compiled for spectral rendering. Mutual illumination was excluded from the rendering process by limiting the number of bounces to one, to ensure the chromaticities of the sweets were not affected by surrounding sweets or walls.

The scenes consisted of two boxes with open fronts. A square area light encompassed the ceiling of each box. The lights were hidden by wall fragments on the top fronts. As in Experiment 1, the left box was always illuminated by D57, and the right box was illuminated by either a Blue, Yellow, Red, or Green illumination—each $30\Delta E_{u*v*}$ from D57. The boxes subtended approximately 11 × 10 degrees of visual angle.

All surfaces within the boxes had a spectrally neutral reflectance of 0.5. The sweet shape was identical for all competitors and generated by adding cones to either side of an ellipsoid with the vertices randomly deformed to look like the sweet wrappers. A single target sweet was placed in the left-hand box, roughly in the middle, with two rows of four competitor sweets in the right-hand box. Each sweet subtended

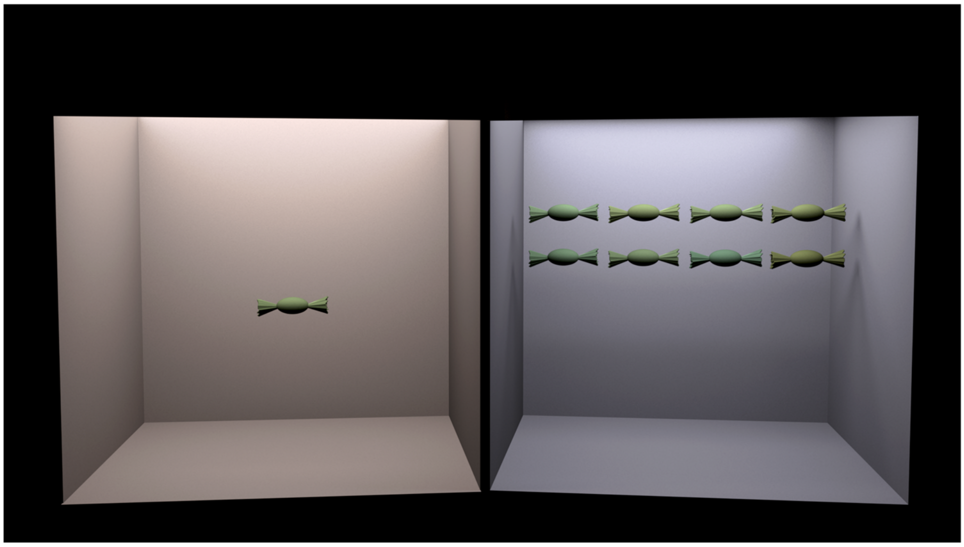

FIGURE 9 Example of a 3D scene containing green sweets under the blue illumination on the right

approximately 1.8 × 0.4 degrees of visual angle at a viewing distance of 60 cm.

The sweets' reflectances were generated using the rough plastic material in Mitsuba, with an alpha of 1, which is perfectly rough, and a specular reflectance of 0, which is perfectly matte. The chromaticities of the target sweets under the neutral illumination were the same as in Experiment 1: Grey, Green, Rose, and Teal. For each chromatic illumination and each target sweet, the colour differences between the eight competitor sweets were evenly spaced in L*u*v*, and corresponded to a reflectance match (R), a tristimulus match (T), and six alternative surfaces, including one over-constant and two under-constant matches, as in Experiment 1. To generate the 3D rendered sweets, spectral reflectance functions for each alternative sweet were computed, which would deliver the desired evenly spaced chromaticities; for details see Supplementary Material.

As in Experiment 1, in two control conditions both boxes were illuminated by the Neutral illuminant (D57). See Supplementary Material for details.

For each of the 24 conditions (four sweets × six illumination conditions), 10 scenes were rendered with the eight competitors in a different, random, order. The hyperspectral rendered scenes were converted to RGB images using the calibration file described in Experiment 1. As some pixels in the rendered images were out of gamut, these were truncated by converting any pixels with R, G, or B greater than 1 to 1, and converting any values less than 0 to 0.

During rendering and truncation, the sweets' chromaticities shifted slightly from the predicted chromaticities. Therefore, during analysis, the chromaticity of the rendered sweet selected was used, as shown in the Supplementary Material.

3.1.4 | Task

The task was the same as in Experiment 1. The observer's selected sweet was indicated by a black arrow instead of an enlargement.

TABLE 3 Mean and SD number of blocks completed by children of each age (out of 6)

Age (years)	Mean number of blocks completed	SD blocks completed
6	5.29	1.50
7	5.22	1.56
8	6	0
9	6	0
10	6	0
11	6	0

3.1.5 | Procedure

All adults completed all conditions within a single session lasting roughly an hour. Most children completed all conditions. Four children, all aged 6 or 7 years, did not complete all conditions due to fatigue, for a total of 12 conditions (out of a possible 240). The number of blocks completed by age is shown in Table 3.

The procedure was identical to that in Experiment 1, except additional instructions were given before the main instructions. These consisted of a demonstration scene containing the boxes from the main experiment with perfectly reflective (white), specular, objects inside: a cone, a sphere, and a cube (Figure 10). The experimenter explained to observers that the light would change colour on every round, while manipulating the illumination in the right-hand box to illustrate the effects. This ensured that observers understood that it was the illumination, rather than the wall, changing colour. A three-dimensional rendered image of Derek the dragon gave the remainder of the instructions, with the same text as in Experiment 1.

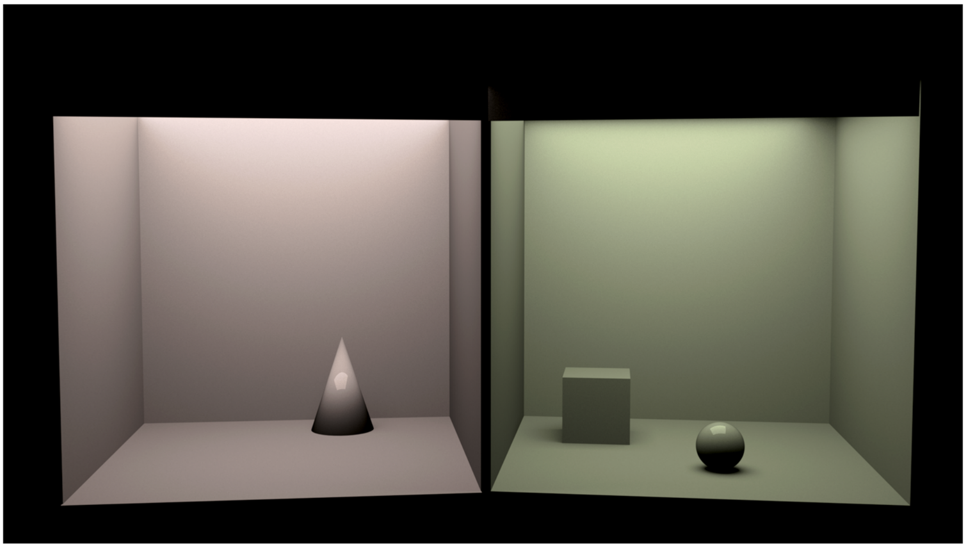

FIGURE 10 Demonstration scene used to show the light changing colour. In this example, the light in the right-hand room is green, and the light in the left-hand room is neutral. In this and all other representations of stimuli, images have been tonemapped for illustrative purposes

3.1.6 | Data analysis

The same exclusion criteria as in Experiment 1 were applied. Ten illumination conditions from children were better fit by a Uniform than a Gaussian distribution, and the SD of responses was greater than 2.32 in 22 conditions from children. Due to overlap between the two criteria, 24 conditions, representing 10.5% of those completed by children, were excluded. No adults' data were excluded by either of these criteria. However, two adults were removed from analysis as their mean CCIs were more than three SD from the mean of all adults; one of these was a non-naive observer who collected part of the data.

When calculating CCIs, the responses were converted to the chromaticity (in u'v') of the sweet in that particular rendering. "T" was defined as the chromaticity of the rendered target sweet. "R" was the chromaticity the rendered target sweet would have under the chromatic illuminant (see Supplementary Material for further details).

As the chromaticities of the rendered sweets did not always fall on a straight line between T and R (as defined above), CCIs were calculated by defining *a* as the (signed) vector projection of the line joining R to the response, onto the line joining R to T, using Equation 1. In separate analyses, we calculated CCIs by defining *a* as the (signed) Euclidean distance from R to the response, but the results were no different to those reported here.

3.2 | Results

The mean chromaticities of the rendered sweets selected by adults and children are shown separately for each sweet colour in Figure 11a, and together in Figure 11b. These show a similar pattern across adults and children which varies dramatically across sweet colours.

To check that observers could discriminate the sweets and perform accurate matches without colour constancy demands, we ran an ANOVA on the data from the neutral control conditions. Across all observers and both control conditions, the mean $\Delta E_{u'v'}$ was 0.00494. Deviations did not differ significantly along the blue-yellow vs red-green axes. There was a significant difference between the sweet colours, with the smallest error for Grey sweets, and the largest for Teal sweets. Children had overall higher $\Delta E_{u'v'}$ than adults ($p < 0.001$), but the pattern across sweet colours was consistent across age groups.

To answer our first research question, and determine whether CCIs differed between children and adults, we ran a lme model with age group, illumination, and sweet colour as predictors of CCI. As can be seen in Figure 12a, children (dark violins) have higher CCIs, on average, than adults (light violins) in all four illumination conditions. Individual children's CCIs for each illumination condition are shown in Figure 12b, and adults' CCIs in 12c. The results of the lme model in Figure 13 show a significant main effect of age group ($p < 0.001$), with higher CCIs in children (mean = 0.315, SD = 0.362) than in adults (mean = 0.245, SD = 0.294).

To determine whether colour constancy changes over childhood (our second research question), robust regressions were fitted to the children's data for each illumination condition separately. Under both Blue ($\beta = -0.036$, $p = 0.019$) and Red ($\beta = -0.031$, $p = 0.020$), but not Yellow or Green, age negatively predicted CCI. To further explore the developmental trajectory, we fit four models (linear, jump, hockey stick, and no change) to the CCIs as a function of log(age). The linear model was the best fit for the Blue ($\beta = -0.0644$), Green ($\beta = -0.0515$), and Red ($\beta = -0.116$) illuminations. Under the Yellow illumination, the no change model was the best fit. The best fitting models are plotted in Figure 14. The BIC for each model are shown in the Supplementary Material.

The main effects model found significant differences in CCIs across illuminations (highest for Blue) and sweet colours (highest for Grey). To determine whether the pattern across illuminations and sweets differed across age groups, addressing our third research question,

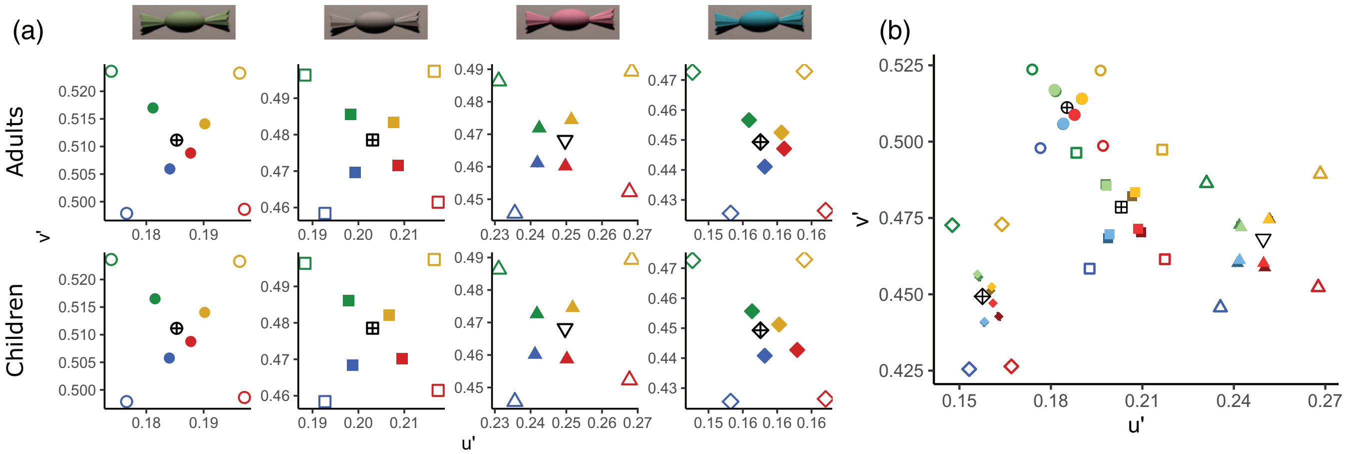

FIGURE 11 Mean chromaticities of sweets selected. (a) shows the chromaticities for each sweet colour separately: Os are green; squares are grey; triangles are rose; diamonds are teal. The top row shows the adults' results and the bottom row shows the children's results. (b) shows the mean chromaticities selected by both adults (pale symbols) and children (dark symbols) for all sweet colours. In all plots, black symbols are tristimulus matches, open coloured symbols are reflectance matches, and filled coloured symbols are mean matches

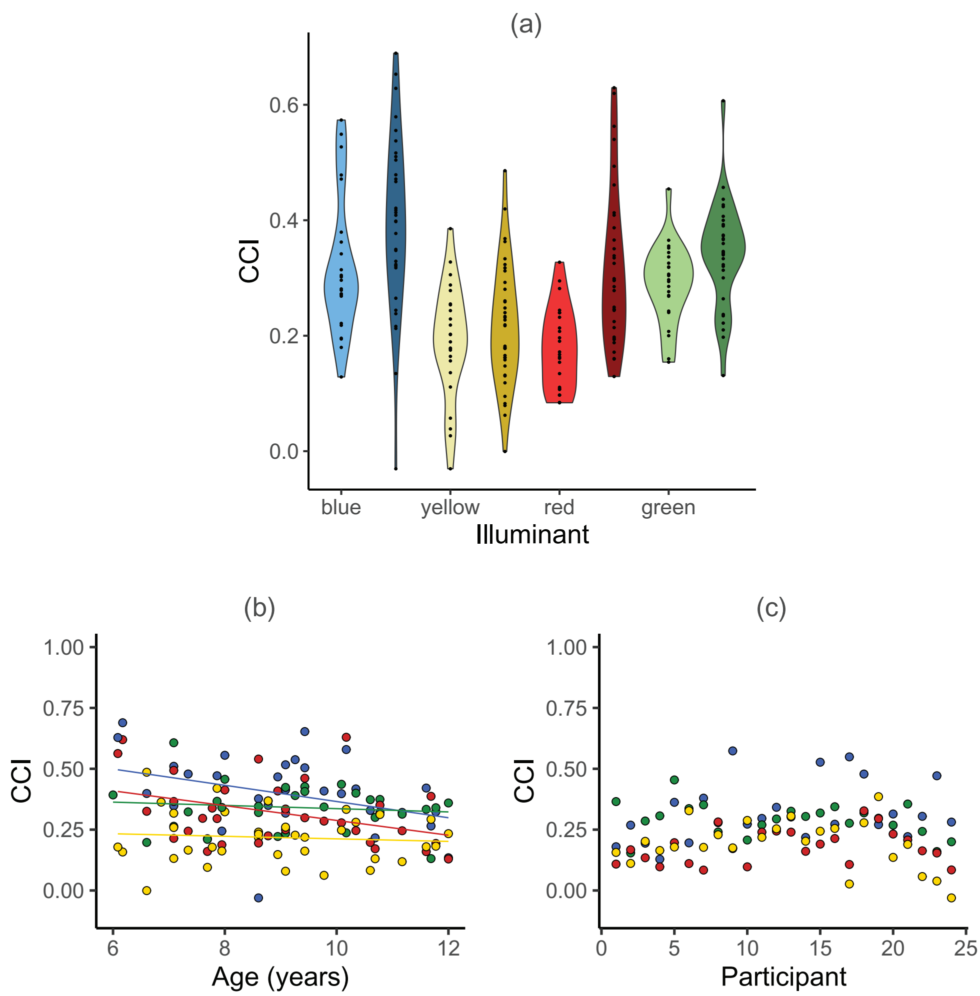

FIGURE 12 (a): Mean CCIs for adults (pale) and children (dark) for each illumination (b): Individual children's CCIs against age, with regression lines for each illumination. Colours reflect illuminations. (c): Individual adult's CCIs for each illumination condition. Horizontal axis is age rank of observer

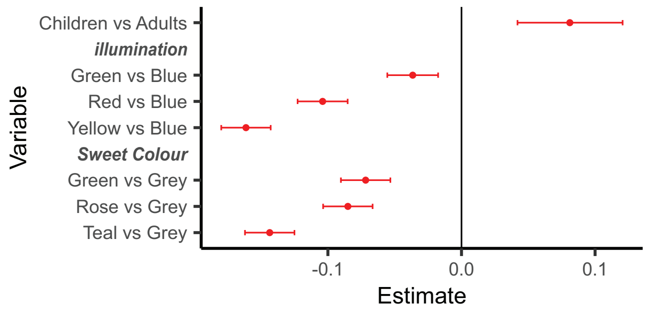

FIGURE 13 Main effects model results. Red circles are estimates, with error bars showing 95% CIs. For estimates greater than 0, the comparison condition predicts higher CCIs; for estimates lower than 0 the reference condition predicts higher CCIs

we ran a lme model to test for any interaction effects. Illumination, sweet colour, and age group were added to the model as interacting predictors of CCIs (Figure 15). There was a significant interaction between illumination and sweet colour, visible in Figure 11. For Grey sweets, children had higher CCIs than adults under the Blue illumination, but the Age Group | Illumination interaction suggests that the difference between children and adults was significantly smaller under the Yellow illumination. Furthermore, the difference between adults and children under the Blue illumination did not depend on sweet colour, but there was a significant positive estimate for children|Teal$_{sweets}$|Red$_{illumination}$. This means that, while adults have a negative interaction for Teal$_{sweet}$|Red$_{illumination}$, this effect is either smaller, non-significant, or reversed in children.

3.3 | Interim discussion

In this experiment we measured colour constancy in children and adults using realistic three-dimensional computer rendered stimuli, to answer three research questions. First, we found that 6- to 11-year-old children had higher colour constancy indices than adults, as in Experiment 1. Second, we found that under the Red and Blue, but not Yellow or Green, illuminations CCIs decreased with age from 6 to 11 years. This contrasts with Experiment 1 where constancy decreased with age under all four illuminations.

Third, considering the pattern across illuminations and sweets, the highest CCIs were under the Blue illumination and for Grey sweets, in agreement with Experiment 1. The superior performance under the Blue illumination may indicate use of a daylight prior, although it does not extend to the Yellow illumination used here, which elicited the lowest CCIs. As in Experiment 1, there was a significant interaction between sweet colour and illumination. In Figure 11, CCIs appear to be highest when the chromaticities of the illumination and sweet are similar. There were few interactions involving age group, suggesting a similar pattern across sweets and illuminations. However, there was a significantly smaller difference between the Yellow and Blue illuminations for children compared to adults. Additionally, model fitting found the developmental trajectory was best fit by a linear model for the Blue, Green, and Red illuminations whereas the data were best fit by a no change model under the Yellow illumination, suggesting CCIs do not change with age. Taken together, these findings suggest that the developmental trajectory is different under the Yellow illumination compared to the others.

Overall, the results generally agreed with those from Experiment 1, although the CCIs were somewhat lower.

4 | DISCUSSION

In this study, we ran two experiments, using a novel measure of colour constancy, to determine the nature of the development of colour constancy. This task was designed to be appropriate for measuring colour constancy in 6- to 11-year-old children. Overall, the pattern of results was similar across the two experiments.

We first asked whether constancy differed between adults and children, and examined the developmental trajectory across childhood. In both experiments we found a significant difference between age groups, with children demonstrating better colour constancy than adults, as shown by picking sweets closer to the reflectance match, resulting in higher CCIs. Robust regressions found that CCIs decreased with age from 6 to 11 years under all illumination conditions with two-dimensional stimuli (Experiment 1), and under Blue and Red illuminations with three-dimensional rendered stimuli (Experiment 2). In Experiment 1, model fitting showed a step-like decrease in colour constancy at around 8.5 years under Blue and Red illuminations and a linear decline with age under Green and Yellow. In Experiment 2, the data were best fit by a linear model under all illuminants except Yellow, in which CCIs did not change with age. Taken together, these findings suggest that colour constancy decreases with age. Whilst this finding may appear counter-intuitive, and differs from the findings in toddlers (Rogers et al., 2020), it agrees with findings described by Katz (2013). Katz reports that Brunswik found colour constancy to peak between 8 and 15 years, after which it decreases up to adulthood. Interestingly, in another set of experiments, Katz notes that Burzlaff found that the developmental trajectory depends on the experimental setup, such that children are as good as adults on an object selection task, with almost perfect constancy, but performance on an adjustment task improves with age. The task used here is more like the former, in which children demonstrate high degrees of colour constancy. It would therefore be interesting to determine whether the results found here apply to other experimental setups. However, matching tasks have additional cognitive and motor demands, which could explain the developmental trajectory found by Katz. Furthermore, the findings reported by Katz are to be taken with caution, as he does not report any statistical analyses.

It is important to consider other possible explanations for the decreasing CCIs with age, apart from declining colour constancy. Firstly, children's data are generally more noisy than adults' data, which could be due to them responding more randomly. If an observer were responding entirely at random, we would predict CCIs of 0.375. In Experiment 1, adults' CCIs were closer to this (0.361) than

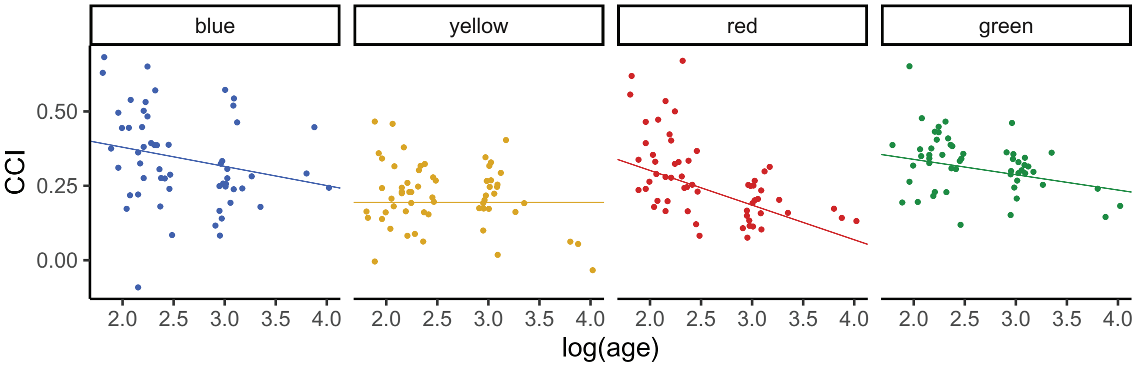

FIGURE 14 CCIs against log(age) with the best fitting models for each illumination. For Blue, Red, and Green the best fitting model is linear, with the regression line shown. For Yellow, the best fit is no change, with the mean across all ages plotted as a horizontal line

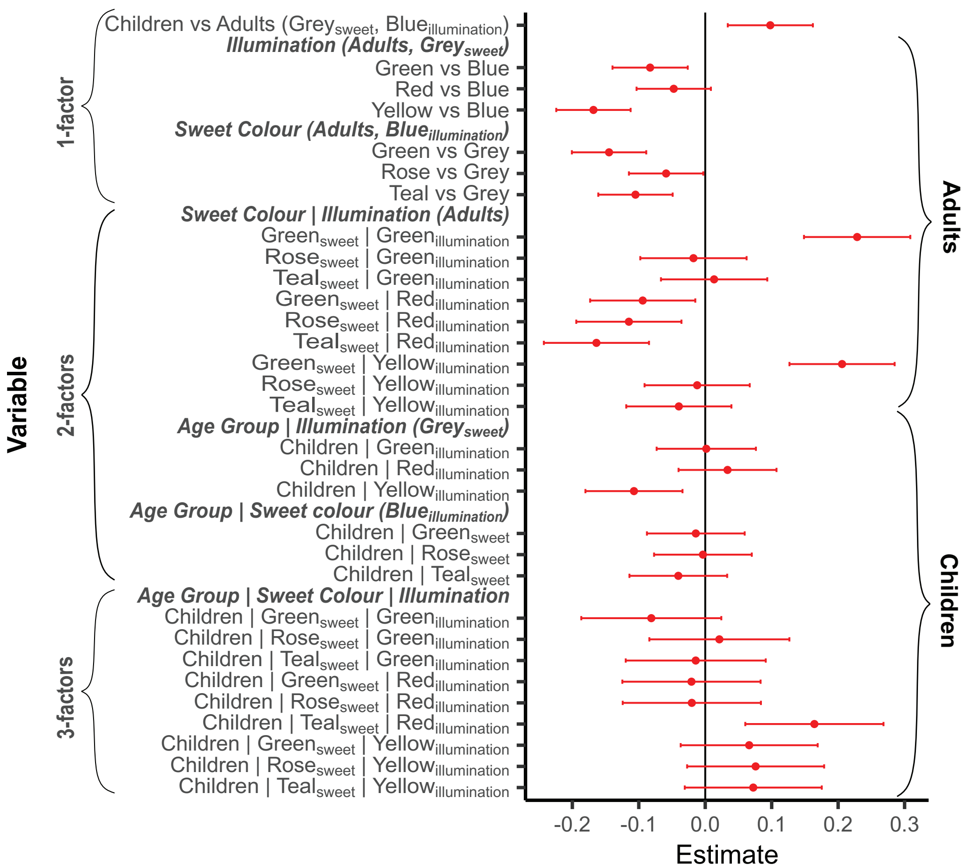

FIGURE 15 Interaction model effects. Red circles show estimates, with 95% CI error bars. See text in Supplementary Material for full explanation

children's (0.425), whereas in Experiment 2, the lower CCIs meant children were closer to random (0.315) than adults (0.245). It is, however, unlikely that children would respond randomly in one experiment only. Furthermore, we excluded conditions with random responses before analysis.

A more likely alternative explanation is a difference in strategy use due to interpretation of the deliberately ambiguous instructions. Observers were told to find the sweet that Derek the Dragon would "like best." They were not told whether to match based on the hue and saturation, or to pick the sweet which has the same reflectance as the target (paper match), and there was no explicit mention of colour. This was to determine what observers would do in the real world, without an explicit strategy. However, if observers were attempting to make a hue/saturation match, the "correct" response would be the tristimulus match, whereas for a paper match the correct response would be the reflectance match. The greater cognitive effort required in making a hue/saturation match might make it harder for children to adopt that strategy, since they must override in-built colour constancy mechanisms. Future studies manipulating instructions may help to determine whether different strategies drive differences in performance, at least in adults. If a difference in strategy use is driving these developmental differences, it suggests that as age increases, people can more flexibly switch between different methods of processing. This aligns with task-switching literature, which has found this executive function to develop during childhood (Davidson et al., 2006; Gupta et al., 2009).

Relatedly, a difference in perception of the stimuli could have impacted on task strategy. Whilst the stimuli were designed to simulate different illuminations, they could have been perceived as different wall colours by some observers. In situations in which the background changes, the "correct" match would be a hue/saturation match. It is possible that there are developmental changes in likelihood of perceiving the difference as due to illumination or wall colour, which might have influenced the results.

A difference in adaptation to the test illumination, or in simultaneous contrast, may also explain the difference in CCIs between children and adults. Adaptation is known to be one of the major mechanisms of colour constancy (Fairchild & Lennie, 1992; H. Smithson & Zaidi, 2004; Werner, 2014). Children might have adapted more to the test illumination, either due to having poorer divided attention and executive control (Klenberg et al., 2001; Rosario Rueda et al., 2004; Shepp & Barrett, 1991), causing more looking to the right-hand side of the screen; having smaller receptive fields; or using more local, as opposed to global, processing (Balas et al., 2020; Poirel et al., 2008). All these factors could similarly impact simultaneous contrast. However, there is evidence that by 5 years children's receptive fields from V1 to VO1 are adult-like (Gomez et al., 2018). To test whether looking behaviour can explain the results, future studies might use eyetracking to determine whether there is a difference in looking behaviour between age groups, or haploscopic viewing to ensure all observers are fully adapted in one eye.

In summary, the developmental findings suggest that, by the age of 6 years, children are better on this task, which might imply they are better able to take the illumination into account when selecting objects than adults. This extends findings from infant studies which suggest that infants have a rudimentary form of colour constancy at a couple of months old. Better performance on this task might arise from different weights on low-level versus higher-level mechanisms to colour constancy. But the difference might be in task strategy, rather than colour constancy per se. Further studies using physical, rather than computer-generated, stimuli are needed to determine the cause of this difference.

Our use of four different sweet colours and four illuminations allowed us to investigate the pattern of colour constancy across different conditions, and determine whether this pattern changed with age. In both experiments, colour constancy was significantly higher under the blue daylight illuminant than any other illuminants. Such a "blue bias" has been found in previous research (Delahunt & Brainard, 2004; Weiss et al., 2017), and is in line with a broad daylight prior, as modelled by D. H. Brainard et al. (2006). The fact that this effect did not extend to the Yellow illumination, under which observers had the lowest CCIs, could be due to a skew in the distribution of daylights such that highly saturated blues are more common than highly saturated yellows (Hernández-Andrés et al., 2001; Nascimento et al., 2016). Thus, observers may be more likely to attribute the blue illuminated box to a difference in illumination while attributing the other illuminations to a difference in wall colour. In both experiments the pattern across illuminations was similar for adults and children. However, in Experiment 2 (three-dimensional stimuli) there was a significant interaction between illumination and age group such that children had a smaller advantage over adults under the Yellow illumination than all other illuminations. This is in agreement with the model fitting which suggested no change in colour constancy with age under the Yellow illumination. Considering the developmental trajectory, we found a divide between Red and Blue versus Yellow and Green illuminations, such that there was a step-like change in colour constancy with age under Blue and Red illuminations but a linear change under Green and Yellow in Experiment 1. In Experiment 2 there was no change with age across childhood under Yellow or Green but a linear decline with age under Blue and Red. Overall, these results suggest that any daylight prior is already present in children, but the divide between daylight and non-daylight illuminations requires further investigation.

As with illuminations, the pattern across sweet colours remained fairly consistent across age groups and experiments, with both adults and children demonstrating the highest constancy for the Grey sweets, followed by Green, Rose, and poorest constancy for Teal. There were no significant interactions between age group and sweet colour in either experiment. Interestingly, we found a consistent significant interaction between surface reflectance and illumination, which was similar across age groups. When the sweet and illumination had similar chromaticities, such as the Green sweet under Green illumination, CCIs were higher than when the chromaticities were opposed. This effect is clearly visible for adults and children in Figures 4 and 11 (b). A possible explanation is that observers have a bias towards more saturated competitors. Under the Green illuminant, the most saturated Green competitor would be the over-constant one whereas the most saturated Rose competitor would be under-constant. An alternative explanation with similar predictions is that observers are biased towards competitors which have a greater cone contrast against the background. Clearly observers are not simply picking the sweet with the highest saturation or greatest cone contrast, as these would result in CCIs of 1.25 or −0.5 depending on the condition, but they may have a bias in that direction. Radonjić et al. (2015a) used a similar methodology to that used here, with blue and yellow illuminations, and found interactions between the set of reflectances and illuminations. Under their blue illumination, colour constancy was higher for the set of reflectances used in the present experiments than for more natural reflectances, whereas for the natural reflectances, colour constancy was higher under the yellow illuminant. The natural reflectances they used appear more yellow under a neutral illumination. Therefore, although not discussed or explicitly tested, their findings are broadly in agreement with those found in the present study. However, future studies are needed which manipulate the surface reflectances in a more controlled manner, such as using chromaticities aligned with, or orthogonal to, the illuminations, to determine whether saturation or cone contrast can explain these findings. Furthermore, the interaction seen here suggests future experiments investigating a daylight prior would benefit from manipulating surface reflectances in addition to illuminations.

The CCIs found in these experiments were generally low, with mean CCIs of 0.425 and 0.361 in Experiment 1, and 0.315 and 0.245 in Experiment 2, for children and adults respectively. Several factors might explain this level of performance. Firstly, as noted above, the instructions to observers were deliberately ambiguous and did not mention colour. Instructions have been shown to influence measured levels of performance. Hue/saturation matches generally show poorer constancy than paper matches (L. E. Arend et al., 1991). It is possible that both children and adults opted for hue/saturation matches, with adults coming closer to this tristimulus match, as discussed above. This cannot be the whole explanation, because CCIs would then be even closer to zero for both (L. Arend & Reeves, 1986). It is more likely that the constancy task was difficult for both because of the nature of the stimuli. In Experiment 1, the stimuli were two-dimensional and the backgrounds might not have been perceived as differing in illumination. Radonjić et al. (2015b) found very low CCIs for two-dimensional stimuli. It is surprising that the CCIs in Experiment 2, with three-dimensional rendered stimuli and an explanation of the changing illumination, were lower than in Experiment 1. Hedrich et al. (2009) found higher constancy for three-dimensional compared to two-dimensional scenes, and de Almeida et al. (2010) found no effect of dimensionality, both for real, rather than rendered scenes. In some conditions of Experiment 2, it was impossible to achieve perfect constancy as the chromaticities of the competitors did not contain a reflectance match, although CCIs above 0.9 in all but 1 condition were still possible (see Supplementary Materials). Additionally, the 3D realism of these rendered scenes might have been reduced by the lack of mutual reflections. Because the sweets were "floating," they also might have seemed to belong to a different illumination framework from the boxes. Furthermore, the stimuli were not immersive in either experiment; they were viewed on a monitor without the stereoscopic viewing deployed in Radonjić et al. (2015a). The task used here was also harder than in Radonjić et al. (2015b) where observers had only two rather than eight competitors to choose from on a given trial. Many observers commented on the difficulty of the task.

In summary, we found a consistent pattern in colour constancy across illuminations and sweets across ages, with adults and children showing a "blue bias," and higher constancy for sweets with a more neutral reflectance. Furthermore, for both age groups the effect of illumination was mediated by the surface reflectance. Importantly, we found a surprising decline in colour constancy with age from children to adults. This suggests that the cognitive contributions to colour constancy—or to task strategy—change during development.

ACKNOWLEDGMENTS

This work was funded by Leverhulme Trust Research Project Grant RPG-2017-097.

CONFLICTS OF INTEREST

All authors declare that they have no conflicts of interest.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are openly available in OSF at osf.io/fnuzv. citation: Wedge-Roberts, R. (2020, December 17). Developmental changes in colour constancy in a naturalistic object selection task. Retrieved from osf.io/fnuzv.

ETHICS STATEMENT

This work was granted ethical approval by Durham University's Psychology Ethics Committee (reference 17/17).

REFERENCES

Arend, L., & Reeves, A. (1986). Simultaneous color constancy. *Journal of the Optical Society of America A, Optics and Image Science*, *3*(10), 1743–1751. ISSN 1084–7529. https://doi.org/10.1364/JOSAA.3.001743

Arend, L. E., Reeves, A., Schirillo, J., & Goldstein, R. (1991). Simultaneous color constancy: Papers with diverse Munsell values. *Journal of the Optical Society of America A*, *8*(4), 661. ISSN 1084–7529. https://doi.org/10.1364/josaa.8.000661

Balas, B., Auen, A., Thrash, J., & Lammers, S. (2020). Children's use of local and global visual features for material perception. *Journal of Vision*, *20*(2), 10. ISSN 15347362. https://doi.org/10.1167/jov.20.2.10

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. 10.18637/jss.v067.i01

Beau Lotto, R. (2004). Visual development: Experience puts the colour in life. *Current Biology*, *14*(15), 619–621. ISSN 09609822. https://doi.org/10.1016/j.cub.2004.07.045

Beck, J. (1966). Age differences in lightness perception. *Psychonomic Science*, *4*(6), 201–202. ISSN 0033–3131. https://doi.org/10.3758/bf03342249

Bone, R. A., Landrum, J. T., Fernandez, L., & Tarsis, S. L. (1988). Analysis of the macular pigment by hplc: Retinal distribution and age study. *Investigative Ophthalmology & Visual Science*, *29*(6), 843–849.

Brainard, D. (1998). Color constancy in the nearly natural image. 2. Achromatic loci. *Journal of the Optical Society of America A*, *15*(2), 307. ISSN 1084–7529. https://doi.org/10.1364/JOSAA.15.000307

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

Brainard, D. H., Brunt, W. A., & Speigle, J. M. (1997). Color constancy in the nearly natural image. 1. asymmetric matches. *Journal of the Optical Society of America A*, *14*(9), 2091–2110.

Brainard, D. H., & Freeman, W. T. (1997). Bayesian color constancy. *Journal of the Optical Society of America A*, *14*(7), 1393. ISSN 1084–7529. https://doi.org/10.1364/josaa.14.001393

Brainard, D. H., Longére, P., Delahunt, P. B., Freeman, W. T., Kraft, J. M., & Xiao, B. (2006). Bayesian model of human color constancy. *Journal of Vision*, *6*(11), 1267–1281. https://doi.org/10.1167/6.11.10

Brainard, D. H., Pelli, D. G., & Robson, T. (2002). Display characterization. In: J. P. Hornak editor, *Encyclopedia of imaging science and technology*. Wiley, pp. 172–188.

Brislin, R. W., & Leibowitz, H. W. (1970). The effect of separation between test and comparison objects on size constancy at various age-levels. *The American Journal of Psychology*, *83*(3), 372–376. ISSN 00029556.

Brown, A. M. (1990). Development of visual sensitivity to light and color vision in human infants: A critical review. *Vision Research*, *30*(8), 1159–1188. ISSN 00426989. https://doi.org/10.1016/00426989(90)90173-I

Chambers, C., Sokhey, T., Gaebler-Spira, D., & Kording, K. P. (2018). The development of Bayesian integration in sensorimotor estimation. *Journal of Vision*, *18*(12), 1–16. ISSN 15347362. https://doi.org/10.1167/18.12.8

Concetta Morrone, M., David Burr, C., & Fiorentini, A. (1990). Development of contrast sensitivity and acuity of the infant colour system. *Proceedings of the Royal Society of London Series B: Biological Sciences*, *242*(1304), 134–139.

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. ISSN 0012–1649. https://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), 255–267. ISSN 00220965. https://doi.org/10.1016/0022-0965(87)90033-6

Davidson, M. C., Amso, D., Anderson, L. C., & Diamond, A. (2006). Development of cognitive control and executive functions from 4 to 13 years: Evidence from manipulations of memory, inhibition, and task switching. *Neuropsychologia*, *44*(11), 2037–2078.

de Almeida, V. M. N., Fiadeiro, P. T., & Nascimento, S. M. C. (2010). Effect of scene dimensionality on colour constancy with real three-dimensional scenes and objects. *Perception*, *39*(6), 770–779.

Delahunt, P. B., & Brainard, D. H. (2004). Does human color constancy incorporate the statistical regularity of natural daylight? *Journal of Vision*, *4*(2), 1. ISSN 1534–7362. https://doi.org/10.1167/4.2.1

Fairchild, M. D., & Lennie, P. (1992). Chromatic adaptation to natural and incandescent illuminants. *Vision Research*, *32*(11), 2077–2085. ISSN 00426989. https://doi.org/10.1016/00426989(92)90069-U

Field, J., & Collins, J. K. (1977). A developmental study of the effects of instructions on visual shape judgements. *British Journal of Psychology*, *68*(3), 297–304. ISSN 20448295. https://doi.org/10.1111/j.2044-8295.1977.tb01593.x

Foster, D. H. (2011). Color constancy. *Vision Research*, *51*(7), 674–700. ISSN 1878–5646. https://doi.org/10.1016/j.visres.2010.09.006; http://www.ncbi.nlm.nih.gov/pubmed/20849875

Gomez, J., Natu, V., Jeska, B., Barnett, M., & Grill-Spector, K. (2018). Development differentially sculpts receptive fields across early and high-level human visual cortex. *Nature Communications*, *9*(1), 1–12.

Granrud, C. E. (2009). Development of size constancy in children: A test of the metacognitive theory. *Attention, Perception & Psychophysics*, *71*(3), 644–654. ISSN 1943–3921. https://doi.org/10.3758/APP

Granrud, C. E., & Schmechel, T. T. N. (2006). Development of size constancy in children: A test of the proximal mode sensitivity hypothesis. *Perception & Psychophysics*, *68*(8), 1372–1381. ISSN 0031–5117. https://doi.org/10.3758/BF03193736

Gupta, R., Kar, B. R., & Srinivasan, N. (2009). Development of task switching and post-error-slowing in children. *Behavioral and Brain Functions*, *5*(1), 1–13.

Hedrich, M., Bloj, M., & Ruppertsberg, A. I. (2009). Color constancy improves for real 3d objects. *Journal of Vision*, *9*(4), 16–16.

Hernández-Andrés, J., Lee, R. L., & Romero, J. (1999). Calculating correlated color temperatures across the entire gamut of daylight and skylight chromaticities. *Applied Optics*, *38*(27), 5703–5709. ISSN 0003–6935. https://doi.org/10.1364/AO.38.005703; http://www.ncbi.nlm.nih.gov/pubmed/18324081

Hernández-Andrés, J., Romero, J., Nieves, J. L., & Lee, R. L. (2001). Color and spectral analysis of daylight in Southern Europe. *Journal of the Optical Society of America A*, *18*(6), 1325–1335.

Hollants-Gilhuijs, M. A. M., Ruijter, J. M., & Spekreijse, H. (1998). Visual half-field development in children: Detection of colour-contrast-defined forms. *Vision Research*, *38*(5), 645–649. ISSN 00426989. https://doi.org/10.1016/S0042-6989(97)00203-4

Hui Lin Chien, S., Bronson-Castain, K., Palmer, J., & Teller, D. Y. (2006). Lightness constancy in 4-month-old infants. *Vision Research*, *46*(13), 2139–2148. ISSN 00426989. https://doi.org/10.1016/j.visres.2005.11.027

Ishihara, S. (2006). *Ishihara's tests for colour deficiency*. Kanehara Trading Inc.

Johnson, S. P., & Aslin, R. N. (2000). Infants' perception of transparency. *Developmental Psychology*, *36*(6), 808.

Judd, D. B., MacAdam, D. L., Wyszecki, G., Budde, H. W., Condit, H. R., Henderson, S. T., & Simonds, J. L. (1964). Spectral distribution of typical daylight as a function of correlated color temperature. *Journal of the Optical Society of America*, *54*(8), 1031. ISSN 0030–3941. https://doi.org/10.1364/JOSA.54.001031

Kaess, D. W., Haynes, S. D., Craig, M. J., Pearson, S. C., & Greenwell (1974). Effect of distance and size of standard object on the development of shape constancy. *Journal of Experimental Psychology*, *102*(1), 17–21. ISSN 00221015. https://doi.org/10.1037/h0035715

Katz, D. (2013). *The world of colour*, *vol.* 44. Routledge.

Kavšek, M. (2011). Infants' responsiveness to lightness changes on a dynamic three dimensional surface. *Vision Research*, *51*(23-24), 2398–2404. ISSN 00426989. https://doi.org/10.1016/j.visres.2011.09.018

Kavšek, M., & Granrud, C. E. (2012). Children's and adults' size estimates at near and far distances: A test of the perceptual learning theory of size constancy development. *i-Perception*, *3*(7), 459–466. ISSN 20416695. https://doi.org/10.1068/i0530

Kleiner, M., Brainard, D., & Pelli, D. (2007). What's new in psychtoolbox-3? *Perception*, 36 ECVP Abstract Supplement.

Klenberg, L., Korkman, M., & Lahti-Nuuttila, P. (2001). Differential development of attention and executive functions in 3-to 12-year-old finnish children. *Developmental Neuropsychology*, *20*(1), 407–428.

Meneghini, K. A., & Leibowitz, H. W. (1967). Effect of stimulus distance and age on shape constancy. *Journal of Experimental Psychology*, *74*(2 PART 1), 241–248. ISSN 00221015. https://doi.org/10.1037/h0024558

Nascimento, S. M. C., Amano, K., & Foster, D. H. (2016). Spatial distributions of local illumination color in natural scenes. *Vision Research*, *120*, 39–44. ISSN 18785646. https://doi.org/10.1016/j.visres.2015.07.005

Olkkonen, M., Hansen, T., & Gegenfurtner, K. R. (2009). Categorical color constancy for simulated surfaces. *Journal of Vision*, *9*(12), 6–6.

Olkkonen, M., Saarela, T. P., & Allred, S. R. (2016). Perception-memory interactions reveal a computational strategy for perceptual constancy. *Journal of Vision*, *16*(3), 1–21. ISSN 15347362. https://doi.org/10.1167/16.3.38

Pearce, B., Crichton, S., Mackiewicz, M., Finlayson, G. D., & Hurlbert, A. (2014). Chromatic illumination discrimination ability reveals that human colour constancy is optimised for blue daylight illuminations. *Plos One*, *9*(2), e87989. https://doi.org/10.1371/journal.pone.0087989

Peeples, D., & Teller, D. Y. (1975). Color vision and brightness discrimination in two-month-old human infants. *Science*, *189*(4208), 1102–1103.

Pelli, D. G. (1997). The videotoolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, *10*, 437–442.

Poirel, N., Mellet, E., Houdé, O., & Pineau, A. (2008). First came the trees, then the forest: Developmental changes during childhood in the processing of visual localglobal patterns according to the meaningfulness of the stimuli. *Developmental Psychology*, *44*(1), 245.

R Core Team (2020). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Radonjić, A., & Brainard, D. H. (2016). The nature of instructional effects in color constancy. *Journal of Experimental Psychology: Human Perception and Performance*, *42*(6), 847–865. ISSN 19391277. https://doi.org/10.1037/xhp0000184

Radonjić, A., Cottaris, N. P., & Brainard, D. H. (2015a). Color constancy in a naturalistic goal-directed task. *Journal of Vision*, *15*(13), 3.1–3.21. ISSN 1534–7362. https://doi.org/10.1167/15.13.3.doi

Radonjić, A., Cottaris, N. P., & Brainard, D. H. (2015b). Color constancy supports crossillumination color selection. *Journal of Vision*, *15*(6), 13. ISSN 1534–7362. https://doi.org/10.1167/15.6.13

Rapoport, J. (1967). Attitude and size judgment in school age children. *Child Development*, *38*(4), 1187–1192. ISSN 00093920. https://doi.org/10.2307/1127116

Rogers, M. R., Witzel, C., Rhodes, P., & Franklin, A. (2020). Color constancy and color term knowledge are positively related in early childhood. *Journal of Experimental Child Psychology*, *196*, 104825.

Rosario Rueda, M., Fan, J., McCandliss, B. D., Halparin, J. D., Gruber, D. B., Lercari, L. P., & Posner, M. I. (2004). Development of attentional networks in childhood. *Neuropsychologia*, *42*(8), 1029–1040.

Shepp, B. E., & Barrett, S. E. (1991). The development of perceived structure and attention: Evidence from divided and selective attention tasks. *Journal of Experimental Child Psychology*, *51*(3), 434–458.

Skelton, A. E., Catchpole, G., Abbott, J. T., Bosten, J. M., & Franklin, A. (2017). Biological origins of color categorization. *Proceedings of the National Academy of Sciences*, *114*(21), 5545–5550.

Smithson, H., & Zaidi, Q. (2004). Colour constancy in context: Roles for local adaptation and levels of reference. *Journal of Vision*, *4*(9), 693–710. ISSN 15347362. https://doi.org/10.1167/4.9.3

Smithson, H. E. (2005). Sensory, computational and cognitive components of human colour constancy. *Philosophical Transactions of the Royal Society of London Series B, Biological Sciences*, *360*(1458), 1329–1346. ISSN 0962–8436. https://doi.org/10.1098/rstb.2005.1633

Spitschan, M., Aguirre, G. K., Brainard, D. H., & Sweeney, A. M. (2017). Erratum: Variation of outdoor illumination as a function of solar elevation and light pollution. *Scientific Reports*, *7*, 46930. ISSN 20452322. https://doi.org/10.1038/srep26756

Stone, J. V. (2011). Footprints sticking out of the sand. Part 2: Children's Bayesian priors for shape and lighting direction. *Perception*, *40*(2), 175–190. ISSN 03010066. https://doi.org/10.1068/p6776

Sugita, Y. (2004). Experience in early infancy is indispensable for color perception. *Current Biology*, *14*, 1267–1271. ISSN 0950–1991. https://doi.org/10.1016/j; http://linkinghub.elsevier.com/retrieve/pii/S0960982205000989

The Math Works, Inc. (2018). *MATLAB version 9.5.0,878302 (R2018b)*. The Math Works, Inc.

Thomas, R., Nardini, M., & Mareschal, D. (2010). Interactions between "light-fromabove" and convexity priors in visual development. *Journal of Vision*, *10*(8), 1–7. ISSN 15347362. https://doi.org/10.1167/10.8.6

Tronick, E., & Hershenson, M. (1979). Size-distance perception in preschool children. *Journal of Experimental Child Psychology*, *27*(1), 166–184. ISSN 00220965. https://doi.org/10.1016/0022–0965(79)90068-7

Venables, W. N., & Ripley, B. D. (2002). *Modern Applied Statistics with S*, (4th ed.). Springer. http://www.stats.ox.ac.uk/pub/MASS4/ ISBN 0-387-95457-0

Weiss, D., Witzel, C., & Gegenfurtner, K. (2017). Determinants of colour constancy and the blue bias. *i-Perception*, *8*(6), ISSN 20416695. https://doi.org/10.1177/2041669517739635

Werner, A. (2014). Spatial and temporal aspects of chromatic adaptation and their functional significance for colour constancy. *Vision Research*, *104*, 80–89. ISSN 18785646. https://doi.org/10.1016/j.visres.2014.10.005

Yang, J., Kanazawa, S. o., & Yamaguchi, M. K. (2010). Perception of Munker-White illusion in 4-8-month-old infants. *Infant Behavior and Development*, *33*(4), 589–595. ISSN 01636383. https://doi.org/10.1016/j.infbeh.2010.07.010

Yang, J., Kanazawa, S. o., 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), 126–136. ISSN 00220965. https://doi.org/10.1016/j.jecp.2012.11.013

Yang, J., Kanazawa, S. o., Yamaguchi, M. K., & Motoyoshi, I. (2015). Pre-constancy vision in infants. *Current Biology*, *25*(24), 3209–3212. ISSN 0960–9822. https://doi.org/10.1016/j.cub.2015.10.053

Yonas, A., & Hagen, M. (1973). Effects of static and motion parallax depth information on perception of size in children and adults. *Journal of Experimental Child Psychology*, *15*(2), 254–265. ISSN 00220965. https://doi.org/10.1016/0022-0965(73)90147-1

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Wedge-Roberts, R., Aston, S., Beierholm, U., Kentridge, R., Hurlbert, A., Nardini, M., & Olkkonen, M. (2022). Developmental changes in colour constancy in a naturalistic object selection task. *Developmental Science*, e13306. https://doi.org/10.1111/desc.13306