

Categorical perception of colour signals in a songbird

Eleanor M. Caves^{1,2}, Patrick A. Green^{1,2}, Matthew N. Zippel¹, Susan Peters¹, Sönke Johnsen¹ & Stephen Nowicki^{1*}

In many contexts, animals assess each other using signals that vary continuously across individuals and, on average, reflect variation in the quality of the signaller^{1,2}. It is often assumed that signal receivers perceive and respond continuously to continuous variation in the signal². Alternatively, perception and response may be discontinuous³, owing to limitations in discrimination, categorization or both. Discrimination is the ability to tell two stimuli apart (for example, whether one can tell apart colours close to each other in hue). Categorization concerns whether stimuli are grouped based on similarities (for example, identifying colours with qualitative similarities in hue as similar even if they can be distinguished)⁴. Categorical perception is a mechanism by which perceptual systems categorize continuously varying stimuli, making specific predictions about discrimination relative to category boundaries. Here we show that female zebra finches (*Taeniopygia guttata*) categorically perceive a continuously variable assessment signal: the orange to red spectrum of male beak colour. Both predictions of categorical perception⁵ were supported: females (1) categorized colour stimuli that varied along a continuum and (2) showed increased discrimination between colours from opposite sides of a category boundary compared to equally different colours from within a category. To our knowledge, this is the first demonstration of categorical perception of signal-based colouration in a bird, with implications for understanding avian colour perception and signal evolution in general.

First described for the perception of phonemes in human speech⁶, categorical perception was later shown to function in the perception of auditory signals in other animals^{7–9}. With regard to colour, animals may discriminate among colours but nevertheless treat them as similar^{10–12}, and colour categorization¹³ may influence decision-making thresholds¹⁴. Thus, animals can categorize colours, forming discrete groups of similar yet discriminable variants across the visible spectrum. However, the hallmark of categorical perception^{15,16}—increased discrimination of variants between categories relative to variants from within—has not been demonstrated for natural variation in colour-based signals.

Carotenoid-based colouration is commonly used in visual signalling across many taxa, although its function in assessment signalling is best described in mate choice in birds and fish^{17,18}. Individuals vary in their ability to acquire¹⁹ and metabolize²⁰ carotenoids; therefore, variation in carotenoid-based colouration has been linked to variation in the quality of the signaller. Carotenoid-based beak colouration in male zebra finches ranges from light orange to dark red²¹, beak redness correlates positively with variation in cell-mediated immunity^{21,22}, and females show a mating preference for males with red versus orange beaks²³. Previous studies¹⁸ have tested how receivers respond to both ends of a carotenoid-based colour continuum, but whether they perceive variation continuously (responding differently to any detectable differences in colour) or exhibit categorical perception is unknown.

To create stimuli that vary continuously along a spectrum from red to orange, we selected eight Munsell colours (Pantone) previously used to describe the colour of zebra finch beaks^{24–26}. We modelled chromatic

distance (ΔS) using the receptor noise-limited model of colour discrimination²⁷ (Extended Data Tables 1, 2). To ensure that the selected colours are approximately equidistant from one another when accounting for zebra finch spectral sensitivity²⁸ and ambient light (Extended Data Figs. 1, 2 and Supplementary Information), we visualized ΔS in a chromaticity space²⁹ in which the Euclidean distance between stimuli plotted in an x - y plane equals chromatic distance for a trichromatic viewer (Fig. 1).

We used a food-reward protocol to test for categorization (which is sometimes referred to as ‘labelling’⁶) and discrimination of the eight stimuli spanning the orange–red colour spectrum. We created discs of Munsell paper comprising two semi-circular halves of either the same or different colours (hereafter ‘solid’ and ‘bicolour’, respectively). Once females had been trained to flip over these discs to access a food reward, we trained them to flip bicolour discs first, before any solid discs; in essence, birds learned to recognize ‘bicolour’ versus ‘solid’ discs, rather than particular colour combinations. In experimental trials, we presented females with a foraging grid that had twelve wells, six of which were covered with discs: two solid discs each for each of the two colours,

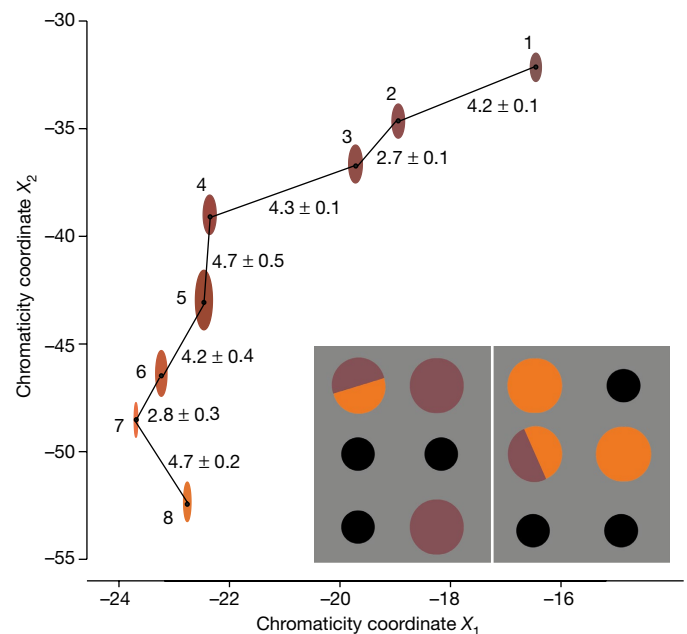


Fig. 1 | Munsell colours used to create stimuli. Colours were approximately equally spaced in chromaticity space and were closer to their nearest neighbour than to any other colour. Dots show mean chromaticity coordinates for each colour; ellipses show one standard deviation in the X_1 and X_2 dimensions ($n = 4$ measurements per colour; Supplementary Information); numbers between dots show chromatic distance (ΔS) between colours (mean \pm s.d.). Ellipse colour corresponds to relevant Munsell colour (exact colours in the figure may vary). Inset, foraging grid of an example trial.

¹Department of Biology, Duke University, Durham, NC, USA. ²These authors contributed equally: Eleanor M. Caves, Patrick A. Green. *e-mail: snowicki@duke.edu

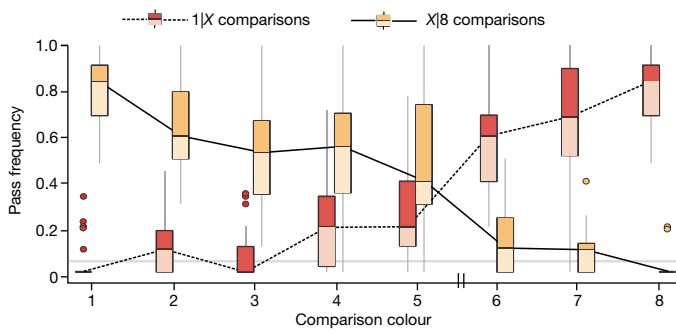


Fig. 2 | Categorization trials suggested a boundary between colours 5 and 6. The boundary is indicated by vertical lines on the x axis. Females were 31% more likely to pass 1|6 than 1|5 trials and 34% more likely to pass 5|8 than 6|8 trials. $n = 26$ birds in three independent cohorts. Box plots depict median (horizontal line inside box), 25th and 75th percentiles (box), 25/75th percentiles $\pm 1.5 \times$ interquartile range (whiskers), and outliers (circles). Horizontal grey line indicates the expected pass frequency if birds flip discs by chance.

and two bicolour discs comprising the same colours as the solid discs (Fig. 1 inset). Birds passed a trial if they flipped both bicolour discs before any solid discs, indicating that they perceived the two colours on the bicolour disc as different.

We first performed categorization experiments to establish the location of potential perceptual boundaries. We tested females using bicolour discs that included colour 1 in combination with all other colours (that is, 1|2, 1|3, 1|4 and so on) and, separately, colour 8 in combination with all other colours (that is, 8|7, 8|6, 8|5 and so on). We determined the proportion of trials passed for each comparison. For both the 1| X and 8| X comparisons (where X is any other colour), pass frequency increased when the chromatic distance increased between colours 1 or 8 and the comparison colour X . The greatest difference occurred, however, when comparing the pass frequencies for 1|5 and 1|6, and likewise between 8|5 and 8|6, suggesting that there is a potential boundary between colours 5 and 6 (Fig. 2). A linear mixed model (Table 1) demonstrated that comparing colour 1 or 8 with the colour immediately preceding versus immediately following the putative boundary resulted in the same change in pass frequency as would moving 10.5 ΔS units, approximately equal to three colour steps within a category (mean ΔS between colour steps = 3.6).

We next determined whether discrimination of colour differences crossing the putative 5–6 category boundary was increased compared to discrimination of equal colour differences that did not cross the

Table 1 | Mixed model results for categorization and greyscale discrimination data

Model	Parameter	Estimate	χ^2 (d.f. = 1)	P value
Colour categorization	Intercept	0.011		
	$\Delta S = 1$	0.022	77.8	<0.0001
	Across 5–6 boundary	0.24	30.8	<0.0001
Greyscale discrimination	Intercept	0.001		
	Michelson contrast	2.30	106	<0.0001

A linear mixed model built from categorization data demonstrates significant effects on pass rate of chromatic distance between colours and whether the comparison crossed the 5–6 boundary. Bird ID was included as a random effect. The binary parameter of crossing the boundary was collinear with differences in brightness; therefore, both could not be included in the same model. Including a random slope of whether the comparison crossed the 5–6 boundary significantly improved the model ($\chi^2(2) = 28.6$, $P < 0.0001$), indicating significant inter-bird variation in the magnitude of the effect of crossing the 5–6 boundary. This variation may be due to variation in boundary location between birds and/or variation in the strength of the effect of a comparison crossing a bird's categorical boundary. A model that also included a random slope of perceptual distance did not improve the model (likelihood ratio test, $P > 0.1$) or qualitatively change the results. A mixed effects model built from greyscale discrimination demonstrates a significant effect of Michelson contrast on pass rate in the absence of hue information. Bird ID was included as a random effect. P values were calculated based on ANOVA comparisons of models that do or do not include each parameter.

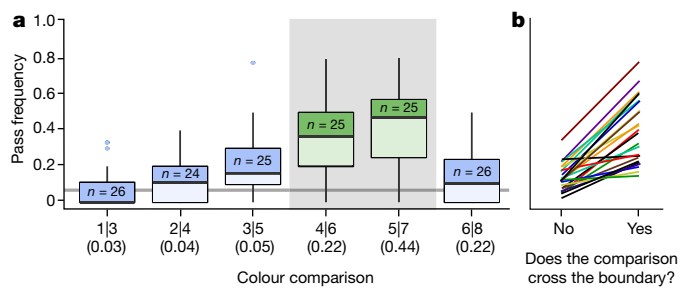


Fig. 3 | Discrimination of stimuli that are two colour spaces apart increased across the 5–6 boundary. **a**, **b**, Mean pass frequency across all individuals (**a**) and for each individual (**b**) was greater for comparisons that crossed the boundary versus those that did not. **a**, Green boxes in the grey-shaded area are comparisons that cross the boundary. Sample sizes (number of birds, across three independent cohorts) are shown within each box. Michelson contrast is shown in parentheses. Boxes, whiskers, circles and horizontal grey line are as described in Fig. 2.

boundary. We presented females with colour pairs that did or did not cross the hypothesized boundary, and that were two colour spaces apart (that is, 1|3, 2|4, 3|5 and so on; Fig. 3a). The pass frequency for comparisons that crossed the 5–6 boundary was 26 ± 6 percentage points higher (mean \pm s.d.) than comparisons that did not (paired t -test, $t_{25} = 9.26$, $P < 0.0001$; Fig. 3b). We found similar results when stimuli were one or three colour spaces apart (Extended Data Figs. 3, 4) and when we combined all categorization and discrimination data into a single linear mixed model (Extended Data Table 3).

Our colour stimuli varied in brightness because real zebra finch beaks of different colour are also not of equal brightness. Several lines of evidence support the conjecture that brightness alone does not completely explain our results. First, we built a linear mixed model comparable to the one presented in Table 1, but which—in addition to chromatic distance (ΔS)—included Michelson contrast (a measure of brightness ratios) between colour pairs to explain pass frequency instead of the binary variable indicating the 5–6 boundary. This model performed substantially worse (Δ Akaike information criterion = 13) than the model that included the binary parameter of crossing the boundary. Additionally, our data show that several discrimination comparisons had either similar Michelson contrasts but different pass frequencies (4|6 compared to 6|8 (Fig. 3) and 5|6 compared to 6|7 (Extended Data Fig. 3)), or equivalent pass frequencies with different Michelson contrasts (4|6 compared to 5|7 (Fig. 3) and 3|6 compared to 4|7 and 5|8 (Extended Data Fig. 4)). Lastly, we performed one-apart discrimination experiments ($n = 18$ birds) using shades of grey (Extended Data Fig. 5 and Supplementary Information) selected to match the red–orange colours in zebra finch double cone²⁸ quantum catch (Extended Data Table 1), an estimate of perceived brightness in passerines³⁰. In these greyscale experiments, pass frequency was significantly predicted by Michelson contrast: birds demonstrated increased discrimination for both 5|6 and 6|7 compared to greyscale pairs with a lower contrast (Table 1 and Extended Data Fig. 5). Together, these pieces of evidence suggest that brightness may contribute to category formation but cannot alone explain the perceptual categories that we observed.

Our results demonstrate categorical perception of colour associated with an assessment signal. We found that a category boundary influences the perception of two colour stimuli as similar or different, and results in differential discrimination between stimuli depending on their location relative to the category boundary. Discrimination improved with increasing chromatic distance between colours (that is, variants within a category were not perceived as identical), but increased most sharply across the category boundary (that is, variants from across the boundary were perceived as most different).

We did not explicitly test whether categorical perception arises at the level of the photoreceptor or higher, such as in the brain. However, a wavelength discrimination function derived from electroretinographic data from the pigeon *Columba livia*³¹, a reasonable proxy for the zebra

finch, suggests that photoreceptor sensitivity alone probably does not explain categorical perception (Extended Data Fig. 6). We also did not directly test female preference for males in relation to the categories found here. Nonetheless, these findings have important implications for the understanding of colour perception and encourage further work exploring whether and how categorical perception interacts with selection on signal form and function, particularly in the context of assessment (Extended Data Fig. 7).

Online content

Any Methods, including any statements of data availability and Nature Research reporting summaries, along with any additional references and Source Data files, are available in the online version of the paper at <https://doi.org/10.1038/s41586-018-0377-7>.

Received: 20 November 2017; Accepted: 21 June 2018;

Published online 1 August 2018.

- Maynard Smith, J. & Harper, D. *Animal Signals* (Oxford Univ. Press, Oxford, 2003).
- Searcy, W. A. & Nowicki, S. *The Evolution of Animal Communication* (Princeton Univ. Press, Princeton, 2005).
- Roff, D. A. The evolution of mate choice: a dialogue between theory and experiment. *Ann. NY Acad. Sci.* **1360**, 1–15 (2015).
- Bornstein, M. H. in *Categorical Perception: the Groundwork of Cognition* (ed. Harnad, S. R.) 287–300 (Cambridge Univ. Press, Cambridge, 1987).
- Studdert-Kennedy, M., Liberman, A. M., Harris, K. S. & Cooper, F. S. Motor theory of speech perception: a reply to Lane's critical review. *Psychol. Rev.* **77**, 234–249 (1970).
- Liberman, A. M., Harris, K. S., Hoffman, H. S. & Griffith, B. C. The discrimination of speech sounds within and across phoneme boundaries. *J. Exp. Psychol.* **54**, 358–368 (1957).
- Baugh, A. T., Akre, K. L. & Ryan, M. J. Categorical perception of a natural, multivariate signal: mating call recognition in tungara frogs. *Proc. Natl Acad. Sci. USA* **105**, 8985–8988 (2008).
- Wyttenbach, R. A., May, M. L. & Hoy, R. R. Categorical perception of sound frequency by crickets. *Science* **273**, 1542–1544 (1996).
- Nelson, D. A. & Marler, P. Categorical perception of a natural stimulus continuum: birdsong. *Science* **244**, 976–978 (1989).
- Wright, A. A. & Cumming, W. W. Color-naming functions for the pigeon. *J. Exp. Anal. Behav.* **15**, 7–17 (1971).
- von Frisch, K. *Bees: Their Vision, Chemical Senses, and Language* (Cornell Univ. Press, Ithaca, 1964).
- Sandell, J. H., Gross, C. G. & Bornstein, M. H. Color categories in macaques. *J. Comp. Physiol. Psychol.* **93**, 626–635 (1979).
- Jones, C. D., Osorio, D. & Baddeley, R. J. Colour categorization by domestic chicks. *Proc. R. Soc. B* **268**, 2077–2084 (2001).
- Hanley, D. et al. Egg discrimination along a gradient of natural variation in eggshell coloration. *Proc. R. Soc. B* **284**, 20162592 (2017).
- Harnad, S. R. *Categorical Perception: the Groundwork of Cognition* (Cambridge Univ. Press, Cambridge, 1987).
- Calder, A. J., Young, A. W., Perrett, D. I., Etcoff, N. L. & Rowland, D. Categorical perception of morphed facial expressions. *Vis. Cogn.* **3**, 81–118 (1996).
- Endler, J. A. Natural and sexual selection on color patterns in poeciliid fishes. *Environ. Biol. Fishes* **9**, 173–190 (1983).
- Hill, G. E. Female house finches prefer colourful males: sexual selection for a condition-dependent trait. *Anim. Behav.* **40**, 563–572 (1990).
- Hill, G. E., Montgomerie, R., Inouye, C. Y. & Dale, J. Influence of dietary carotenoids on plasma and plumage colour in the house finch: intra- and intersexual variation. *Funct. Ecol.* **8**, 343–350 (1994).
- Borel, P. Genetic variations involved in interindividual variability in carotenoid status. *Mol. Nutr. Food Res.* **56**, 228–240 (2012).
- Blount, J. D., Metcalfe, N. B., Birkhead, T. R. & Surai, P. F. Carotenoid modulation of immune function and sexual attractiveness in zebra finches. *Science* **300**, 125–127 (2003).
- George, D. B., Schneider, B. C., McGraw, K. J. & Ardia, D. R. Carotenoids buffer the acute phase response on fever, sickness behavior and rapid bill color change in zebra finches. *J. Exp. Biol.* **220**, 2957–2964 (2017).
- Collins, S. A. & ten Cate, C. Does beak colour affect female preference in zebra finches? *Anim. Behav.* **52**, 105–112 (1996).
- Burley, N. & Coopersmith, C. B. Bill color preferences of zebra finches. *Ethology* **76**, 133–151 (1987).
- Collins, S. A., Hubbard, C. & Houtman, A. M. Female mate choice in the zebra finch — the effect of male beak colour and male song. *Behav. Ecol. Sociobiol.* **35**, 21–25 (1994).
- Birkhead, T. R., Fletcher, F. & Pellatt, E. J. Sexual selection in the zebra finch *Taeniopygia guttata*: condition, sex traits and immune capacity. *Behav. Ecol. Sociobiol.* **44**, 179–191 (1998).
- Vorobyev, M. & Osorio, D. Receptor noise as a determinant of colour thresholds. *Proc. R. Soc. B* **265**, 351–358 (1998).
- Lind, O. Colour vision and background adaptation in a passerine bird, the zebra finch (*Taeniopygia guttata*). *R. Soc. Open Sci.* **3**, 160383 (2016).
- Hempel de Ibarra, N., Giurfa, M. & Vorobyev, M. Discrimination of coloured patterns by honeybees through chromatic and achromatic cues. *J. Comp. Physiol. A* **188**, 503–512 (2002).
- Osorio, D. & Vorobyev, M. Photoreceptor spectral sensitivities in terrestrial animals: adaptations for luminance and colour vision. *Proc. R. Soc. B* **272**, 1745–1752 (2005).
- Riggs, L. A., Blough, P. M. & Schafer, K. L. Electrical responses of the pigeon eye to changes in wavelength of the stimulating light. *Vision Res.* **12**, 981–991 (1972).

Acknowledgements We thank D. Osorio for comments on the manuscript; L. Jin, K. Kochvar, J. Bollinger and L. Deehan for assistance running trials; and R. Mooney and M. Booze for providing access to zebra finches from their colony. Funding for this study was provided by the Duke University Office of the Provost.

Reviewer information Nature thanks A. Kelber and the other anonymous reviewer(s) for their contribution to the peer review of this work.

Author contributions S.N. conceived and led the project, and secured funding; E.M.C., P.A.G. and M.N.Z. designed and carried out behavioural experiments, with assistance from S.P.; S.J. led the visual modelling and colour vision aspects of the study, with assistance from E.M.C. and P.A.G.; M.N.Z. conducted statistical analyses; P.A.G. prepared figures with assistance from M.N.Z.; S.P. managed the zebra finch population; E.M.C. wrote the initial draft of the manuscript, which was subsequently edited and approved by all authors.

Competing interests The authors declare no competing interests.

Additional information

Extended data is available for this paper at <https://doi.org/10.1038/s41586-018-0377-7>.

Supplementary information is available for this paper at <https://doi.org/10.1038/s41586-018-0377-7>.

Reprints and permissions information is available at <http://www.nature.com/reprints>.

Correspondence and requests for materials should be addressed to S.N.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

METHODS

The goal of this study was to test whether female zebra finches perceive colours along a red-to-orange spectrum in a continuous or categorical fashion. We selected eight colours from red to orange that correspond to male zebra finch beak colours and that are roughly equidistant from one another in a chromaticity space based on zebra finch spectral sensitivity. These eight colours were made into disc stimuli that were either solid in colour (that is, made of two halves of the same colour) or bicolour (that is, one colour on one half of the disc and another colour on the other half). Disc stimuli were used in a food-reward protocol. Female zebra finches were first trained to flip over discs covering wells in a foraging grid to find a food reward (millet seeds, as used previously³²), using discs made from the colours at the endpoints of our red–orange continuum (colours 1 and 8). Initially, all wells (both solid and bicolour) were baited with millet seeds, to reward the birds any time they successfully flipped a disc. After being trained on this task, females were further trained to flip only bicolour discs; we used the same stimuli made of colours 1 and 8, but this time baited only wells that were covered with bicolour discs. Once females passed six out of seven consecutive trials under this protocol, indicating they had learned to search for food only under bicoloured discs, we then conducted trials using different combinations of our eight selected colours. If females perceived two colours as distinct, they would flip the bicoloured discs first to gain a food reward; if they did not perceive two colours as distinct, they would not preferentially flip the bicoloured discs first. For details, see ‘Behavioural trials’.

Birds used in this study. All birds in this study were sexually mature female zebra finches (age range: 3–50 months at start of experimental testing) from a colony maintained by R. Mooney at Duke University (IACUC A258-14-10). After transfer from the colony, birds were housed singly in cages (46 × 23 × 23 cm³, Prevue Pet) outfitted with two wooden perches and a cuttlebone. Lighting was controlled during trials (see ‘Behavioural trials’) and food was removed 5 h before trials to ensure that birds would be motivated to perform the task. Outside of trial times, birds were kept on a 15 h:9 h light:dark cycle (consistent with the light cycle of the birds’ home colony), with overhead lighting provided by fluorescent bulbs (Ecolux with Starcoat SP 35/41, colour temperature 3,500–4,100 K, General Electric) with ballast (Hi-Lume 3D/Eco-10, Lutron Electronics) operating at 50–60 Hz. Birds were given zebra finch food (Kaytee Forti-Diet Pro Health Finch diet) and water ad libitum. Rooms were maintained at 25–27 °C. Testing was done under Duke University IACUC protocol A004-17-01.

Selecting stimulus colours. Previous work has shown that the range of beak colouration in zebra finch males can be represented by red and orange shades in the Munsell colour system^{24–26}. Munsell colours are defined by three parameters: hue, value (brightness) and chroma (saturation). Previous work on zebra finches identified a large set of Munsell colours that, by the human eye, approximate the colours of zebra finch beaks. In particular, this set of Munsell colours consists of colours with hues 6.25R–3.75YR, values 3.5–6 and chromas 10–14^{24,26}. Notably, the values used in these previous studies spanned a range of hues from yellow–orange to red, and specifically did not use Munsell colours of the same brightness because real male beaks of different hues are also of different brightness.

These Munsell colours^{24–26} are based on beak colours across two different colony populations of zebra finches and capture most of the variation within those populations, although the authors of the studies note that the beaks of occasional individuals were found to be outside that range. Because the goal of our study was to examine how female zebra finches perceive a range of colours that are similar to the range spanned by the colours of male zebra finch beaks, we used this previous work as a starting point to choose 40 Munsell colour samples from within the set outlined above. Reflectance spectra from each colour sample were measured using an integrating sphere with a built-in tungsten-halogen light source (ISP-REF; Ocean Optics). All measurements were taken with reference to a Spectralon 99% white reflectance standard (Labsphere).

For each of the Munsell colours, we calculated relative photon catches (that is, how many photons are absorbed by a given cone type in response to a visual scene³³) for zebra finch short-, medium- and long-wave cones. Photon catches were calculated over the wavelengths 400–700 nm, using zebra finch spectral sensitivity curves²⁸, an ambient light spectrum and the reflectance spectrum of each colour. Thus, photon catch Q for receptor type r in response to colour c was calculated using:

$$Q_{r,c}(\lambda) \propto \int_{400}^{700} S_r(\lambda) * R_c(\lambda) * I(\lambda) d\lambda$$

where S_r is the sensitivity of receptor type r , R_c is the reflectance of colour c , λ denotes the wavelength, and I is the irradiance of the illuminant. We use proportionality throughout, because we did not require an absolute quantum catch, and the constant factors remain the same across different receptor types. As for our measure of ambient light, we used a standard tungsten bulb illuminance

spectrum (CIE Illuminant A, colour temperature 2,856 K; spectrum in Supplementary Information), which is very similar to the spectrum of ambient light provided by the tungsten bulbs in our experiment (see below, Extended Data Fig. 1 and Supplementary Information). We used sensitivity data from the zebra finch²⁸ to calculate photon catches for the short-, medium- and long-wave and double cones (Extended Data Table 1). We did not calculate photon catches for the ultraviolet cone because (1) male zebra finch beaks reflect minimal ultraviolet light³⁴, (2) female zebra finches do not require ultraviolet radiation to discriminate and rank red–orange colouration³⁵, (3) our own measurements confirmed that Munsell paper does not reflect strongly in the ultraviolet portion of the spectrum, and (4) the reflected ultraviolet radiance from our Munsell chip stimuli, under experimental lighting conditions, was essentially zero (see Extended Data Fig. 2).

In tests of categorical perception, stimuli should be roughly equidistant from one another in perceptual space. No metric can fully describe the perceptual space of human colour vision, let alone that of a zebra finch. Therefore, to calculate the chromatic distance between colour stimuli, we used the receptor noise-limited (RNL) model of colour discrimination²⁷, which uses photon catches calculated from the spectral sensitivity of a relevant visual system to calculate ΔS , a measure of chromatic distance between two colours (equivalent to just-noticeable differences, JNDs). We visualized ΔS using a perceptually uniform, two-dimensional space based on both hue and saturation/chroma, in which the Euclidean distance between two colours is equivalent to the RNL model-derived chromatic distance (ΔS) between colours²⁹. Plotting chromatic distances in this two-dimensional space is only relevant for trichromatic vision, which in our case was appropriate given that we did not incorporate quantum catches from ultraviolet cones (see above). The RNL model²⁷ states that colour discriminability is primarily limited by photoreceptor noise; thus a Euclidean distance of 1.0 in the chromaticity space is approximately equal to one standard deviation of receptor noise, or one JND. Given that this method uses photon catches from single cones (short-, medium- and long-wave), we performed a separate analysis to examine perceived brightness based on double cones (see below).

In summary, the eight colours that we used (see Supplementary Information for reflectance spectra) were based on previously published comparisons with male zebra finch beaks and are approximately equally spaced in a chromaticity space based on zebra finch spectral sensitivity. Throughout, we refer to these colours as 1 (the darkest, red end of the range) through 8 (the brightest, orange end of the range).

Sensitivity analyses. To ensure that the eight colours that we selected are not equidistant in chromaticity space simply owing to the exact parameters of the zebra finch spectral sensitivities and the tungsten lighting conditions that we used, we additionally plotted the selected colours in chromaticity space using different conditions (all spectra and sensitivity curves are provided in the Supplementary Information), including: the zebra finch ultraviolet-light-sensitive (UVS) cone type retina²⁸ under both (1) tungsten and (2) daylight (D65) illuminants; (3) the zebra finch UVS cone type retina using the spectrum of halogen light present in the experimental rooms; (4) the zebra finch UVS cone type retina after applying von Kries adaptation based on a neutral grey background³⁶ to account for colour constancy mechanisms; (5) another UVS cone-type retina visual system (the starling *Sturnus vulgaris*³⁷); (6) the average UVS cone-type retina (data from previously published supplementary online material³³); and (7) the average violet-light-sensitive (VS) cone-type retina (data from previously published supplementary online material³³), which is the other primary type of retina found in birds (Extended Data Table 2). Overall, chromatic distances between selected Munsell colours remained relatively consistent even when changing the spectral sensitivity and lighting conditions, as well as when accounting for colour constancy, with the distance between 5 and 6 never being the largest jump. Thus, chromatic distance alone cannot explain the category boundary that we observed.

Behavioural trials. Room set-up. Birds were housed in individual cages, with up to 12 birds in each room. Before trials, we placed opaque barriers between adjacent cages so that birds could not see their closest neighbours perform any tasks. Individuals were able to see other birds across the room, although they could not see the task that birds were performing—this set-up prevented birds from seeing how other birds were passing the task while ensuring that they could see other birds, which was beneficial given that zebra finches are highly social. During trials, the room’s overhead lights were turned off and birds were allowed approximately 10 min to acclimatize to the experimental lighting. During experimental trials, each cage was illuminated from above by a halogen lamp (colour temperature 2,900 K, model number H&PC-61361, Philips Lighting) approximately 81 cm from the foraging grid (spectra in Supplementary Information). The light passed through vellum paper hung 8 cm from the light source to ensure that each cage had even and diffuse illumination (Extended Data Fig. 1). All trials were recorded on video (Logitech Webcam Pro 9000, Logitech). On trial days, food was removed from the cages at 09:00 and trials began at 14:00, to ensure that birds would be motivated to attempt the task.

Disc design. Our experimental stimuli were discs 2.5 cm in diameter, made of two semi-circles of Munsell colour sheets, glued with their straight edges together to create a full circle. The discs were covered with a clear epoxy cover. A clear vinyl disc (1.3 cm diameter, 0.3 cm high) was attached to the bottom of the disc to ensure it fit precisely into the wells of the foraging grid. We created two types of discs: solid, in which both semi-circular halves were the same colour; and bicolour, in which the two semi-circular halves were made of different colours. Bicolour discs were named for the colours of their two halves—that is, a disc made up of half colour 1 (the far red) and half colour 8 (the far orange) would be referred to as 1|8.

The experimental foraging grid consisted of two grey plastic blocks (13.5 × 9 cm² and 2.5 cm high) positioned adjacent to one another. Each block contained six identical wells (1.3 cm diameter, 0.8 cm deep). Birds first learned to search for food beneath the discs in five stages. In the first stage, we placed millet seeds in four randomly selected wells, with no discs present. In the second stage, four discs (two bicolour, and one each of the two solid colours that comprised the bicolour discs halves) were on the grid, adjacent to the baited wells. In the third stage, discs were placed halfway covering the baited wells. In stage four, discs were placed tipped into the baited wells, so as to hide the seed but with discs fitted only loosely into the well. In the fifth and final stage, discs were placed to completely cover the baited well, so a bird could only access the seed by flipping the disc off of the well using its beak. For each of the training tasks, success was defined as obtaining the seeds from any of the baited wells within a 20-min period. A subject had to pass three consecutive trials of each stage to progress to the next stage.

Once the birds had learned the basic task of searching for food under the discs, they were trained to associate only bicolour discs with a food reward, a stage that we call ‘bicolour association’. Zebra finches were trained in bicolour association using a total of six discs: two 1|8 bicolour discs (which were baited), and two solid discs each of colour 1 and 8 (which were not baited). Importantly, our behavioural data show that birds learned to recognize bicolour versus solid discs, rather than particular colour combinations of bicolour discs, as shown by their ability to extrapolate the colour-based task to a greyscale task (described below). Trials lasted for 2 min, during which time birds were allowed to flip discs, and we recorded the order of the first two discs that they flipped. Following the 2-min observation period, the grid and any remaining millet were left in for up to 20 min. To pass a bicolour association trial, birds had to flip over both bicolour (1|8) discs before either solid disc was flipped within the 2-min observation period. Birds had to pass six out of seven consecutive trials before we determined that they had learned the task, after which they progressed to data collection. If an individual failed more than one training trial out of seven, they were given additional training trials until the pass criterion was reached or until we determined the bird was unlikely to learn the task. In total, 26 out of 30 birds (87%) that reached the bicolour association trials passed and went on to data collection trials.

Data collection trials involved the same grid and 12-well set up as in training. Six wells were covered with discs (two bicolour, four solid), and only the two bicolour discs were baited. We used the ‘sample’ function in R³⁸ to create a set of randomized locations for each disc in each trial, up to 6 trials for each colour combination. Discs were placed in haphazard orientations (that is, the direction of the line bisecting the middle was not consistent from one trial to the next). During data collection, birds were allowed up to 2 min to flip discs, after which time the grid and any remaining seed were removed. We recorded both the order in which the first two discs were flipped, as well as the latency to flipping the first disc. Observers were not blind to the tasks; however, up to three of five (usually, one or two) observers collected data on a given day and no single observer consistently collected data for a given trial (for example, 5|6, 1|6). Results were consistent across the five observers who collected data during the experiment.

We conducted 3–5 trials per bird per day. Overall, we collected data from 26 birds (trained in three independent cohorts of $n = 10$, $n = 6$ and $n = 10$) over the course of 7,015 experimental trials. In 10.1% of trials ($n = 712$), birds flipped either no discs or one disc and then stopped flipping discs. In these cases, we did not count these trials as a failure, but rather removed the trial from the dataset. Most commonly, birds did not flip any discs or only flipped one disc if other birds in the room suddenly went quiet or if they were startled by the behaviour of another bird. Removing these incomplete trials from the dataset ensured that we did not bias our data towards an increased failure rate. At the end of each trial day, we performed a motivation check on each bird by returning the bird’s food dish to the cage and observing for up to 10 min to ensure that they would eat seeds. Birds typically began eating within 30 s of having their food returned, indicating that individuals still had high feeding motivation even at the end of our trials.

Categorization experiments. Categorization experiments tested whether birds responded in a continuous or discontinuous manner to colour variation, identifying the location of boundaries if they were present. As in all trials, we presented birds with two bicolour discs and two of each solid colour disc (six discs total

spread randomly across 12 possible wells on the grid). Each day, a bird was given one refresher trial (1|8 bicolour and 1 and 8 solid) and five experimental trials of a single colour combination. The next day, the bird would receive a refresher trial and five trials on a different colour combination, and so on. Colour combinations used in categorization trials included the combination of colour one and all other colours (that is, 1|2, 1|3, 1|4...) as well as the combination of colour eight and all other colours (that is, 8|7, 8|6, 8|5...). Additionally, as a control for the possibility that birds used olfactory cues to locate seeds under certain discs, we performed 1|1 and 8|8 trials to check that, in the absence of any variation in visual information, birds would not perform better than chance at flipping discs placed on baited well. Birds were shown colour combinations in an order that alternated relatively distant colour combinations (those that were five to seven steps apart) with relatively closer colour combinations (those that were one to four steps apart). Each bird was given two experimental days (10 total trials) with each colour combination, and one experimental day (5 total trials) with each of the 1|1 and 8|8 comparisons. The second experimental day for each combination occurred only after the first experimental day for all other combinations had occurred (that is, experimental days for a specific colour combination did not immediately repeat).

Preliminary data analysis identified a putative boundary between colours 5 and 6. To analyse these data, we calculated the proportion of trials that each bird passed for each comparison (15 total comparisons per bird). Using only data from categorization trials, we built a linear mixed model (throughout, linear mixed models were calculated using R package lme4³⁹) that included pass rate as its response variable, Euclidian chromaticity distance (ΔS) between the two colours being compared and whether the comparison crossed the putative 5–6 boundary as fixed effects, and bird ID as a random effect. The model included random slopes in addition to random intercepts, to account for variation among birds in the strength of the effect of crossing the 5–6 boundary.

We visually inspected the residuals of this model using a quantile–quantile plot, a histogram and a scatterplot of residuals versus predicted values to confirm that the residuals of the model were approximately normally distributed around zero and that they were homoscedastic.

Discrimination experiments. A second requirement for demonstrating categorical perception is that subjects show increased discrimination between stimuli that cross a category boundary compared to equally spaced stimuli that do not cross a boundary. Thus, for discrimination trials, bicoloured discs were made of colour pairs that were one space apart (that is, 1|2, 2|3, 3|4...; Extended Data Fig. 3), two spaces apart (that is, 1|3, 2|4, 3|5...; Fig. 3) or three spaces apart (that is, 1|4, 2|5, 3|6...; Extended Data Fig. 4). The experimental set-up and criteria for passing were the same as for categorization trials. For each bird, mean pass rates were calculated for all discrimination trials that did not cross the putative boundary and, separately, for those that did. We then calculated the difference between these means for each bird and used a two-sided *t*-test to determine whether the mean difference in pass rate was significantly different from zero.

Combining categorization and discrimination data in a single statistical model.

To present our data in the most easily interpreted format, we present separate analyses for categorization and discrimination trials in the main text. To bolster our conclusions, maximize statistical power and contain our analysis in a single model, we built a linear mixed model that includes all discrimination and categorization data. As reported in the main text, this model contains random effects of bird ID and includes crossing the 5–6 boundary as a random slope. As in the main text, significance tests were performed using ANOVA comparisons. The results of the single model can be found in Extended Data Table 3.

Greyscale discrimination experiments. The set of eight Munsell colours that we used in this experiment were selected primarily based on their colour: in particular, the colour of the Munsell colours aligned with those of actual beaks and were equally spaced in a chromaticity space that describes the hue and saturation of colours and from which brightness has been factored out. However, brightness is also an important part of how colour is perceived. To examine the effects that brightness has on structuring zebra finch perception of colours, we performed a second experiment in which we used shades of grey that matched our eight Munsell colours in brightness, as perceived by zebra finches.

To select appropriate shades of grey, we used an integrating sphere to measure reflectance spectra from a set of 72 grey paint swatches (Behr brand, Behr Process). We then calculated the quantum catch of the zebra finch double cone²⁸ for each of our eight Munsell colours and for each of the grey paint swatches; the double cone is believed to be the primary way in which birds encode brightness information³⁰. By finding shades of grey that matched each Munsell colour in double cone quantum catch, we selected a set of eight grey shades that were equivalent in bird-perceived brightness to the colour stimuli that we had been using.

Using this set of eight shades of grey, we created experimental discs for the 1|8-grey and each one-apart combination of grey shades. We followed the same experimental procedure as above to examine the birds’ discrimination ability when hue information had been removed (Extended Data Fig. 5).

Wavelength discrimination function. To investigate whether the categorical perception that we observed could be solely due to the wavelength discrimination function (WDF) of avian photoreceptors, we examined a WDF that was derived using electroretinography in the pigeon *C. livia*³¹. We know of no experimentally or behaviourally derived WDF for zebra finches. The pigeon is a reasonable substitute, however, given that the spectral sensitivity peaks for its medium- and long-wavelength sensitive cones (505 and 565 nm, respectively⁴⁰) are very similar to those of the zebra finch (507 and 565 nm, respectively²⁸).

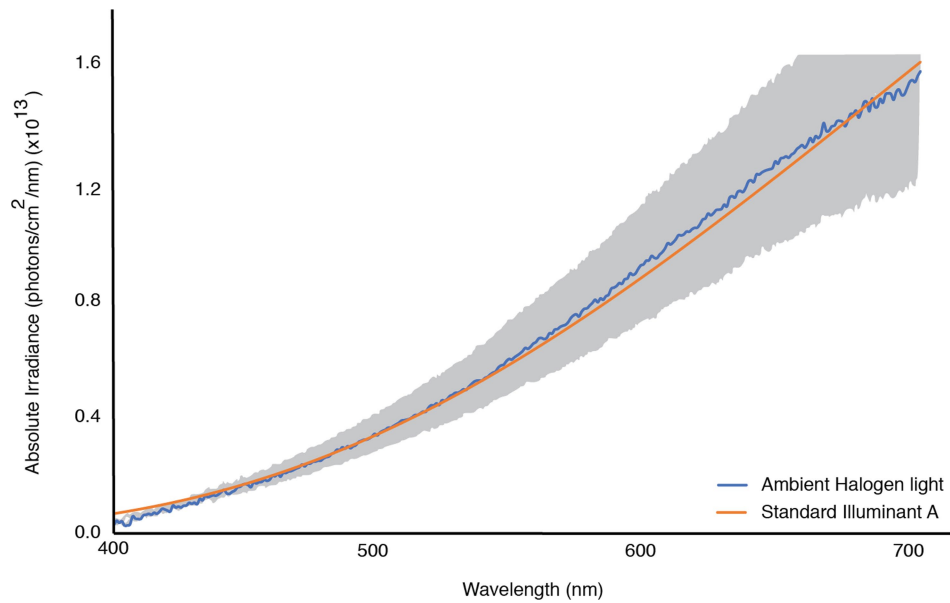
We plotted electroretinography-derived WDF data³¹ alongside the spectral sensitivity peaks of the zebra finches (Extended Data Fig. 6). The units of the WDF are arbitrary units, in keeping with the original publication, but show the general pattern of wavelength discrimination. One complication is that it is not possible to know precisely where in the visible spectrum the avian perception of 'orange' and 'red' would occur. However, in humans, orange and red both occur above 580 nm, in the region in which stimulation is primarily of the long-wavelength cone and secondarily the medium-wavelength cone. The WDF showed a discriminability peak at approximately 605 nm, followed by a relatively smooth decrease in discriminability in the region of the spectrum in which avian viewers probably see orange and red (Extended Data Fig. 6). The shape of these curves suggests that photoreceptor sensitivity on its own probably does not explain categorical perception. From this analysis alone, however, we cannot rule out that hue discrimination based on spectral sensitivity curves contributes to categorical perception. Ultimately, whether categorical perception arises at the level of the photoreceptor or retina, or is a higher-order process, does not affect our findings or interpretation, but suggests avenues for future research.

Data reporting. No statistical methods were used to predetermine sample size. The experiments were not randomized and the investigators were not blinded to allocation during experiments and outcome assessment.

Reporting summary. Further information on experimental design is available in the Nature Research Reporting Summary linked to this paper.

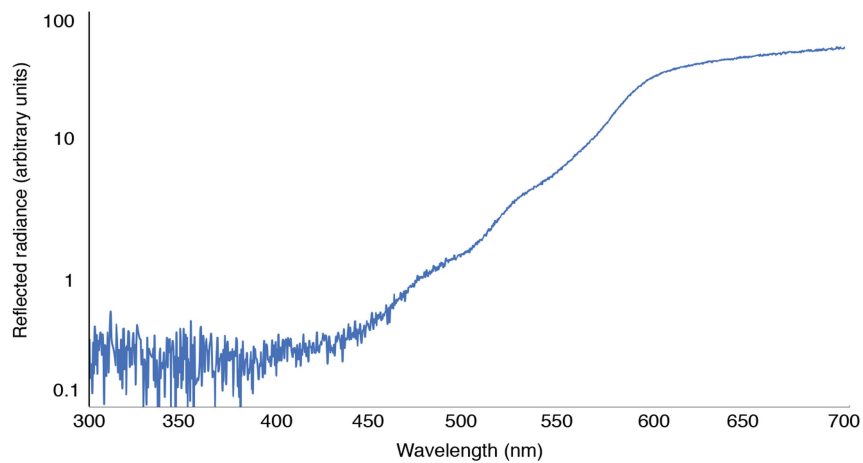
Data availability. The datasets generated and analysed during the current study are available in the Duke Data Repository: <https://doi.org/10.7924/r4rx96r99>.

32. Boogert, N. J., Anderson, R. C., Peters, S., Searcy, W. A. & Nowicki, S. Song repertoire size in male song sparrows correlates with detour reaching, but not with other cognitive measures. *Anim. Behav.* **81**, 1209–1216 (2011).
33. Endler, J. A. & Mielke, P. W. Comparing entire colour patterns as birds see them. *Biol. J. Linn. Soc.* **86**, 405–431 (2005).
34. Bennett, A. T. D., Cuthill, I. C., Partridge, J. C. & Maier, E. J. Ultraviolet vision and mate choice in zebra finches. *Nature* **380**, 433–435 (1996).
35. Hunt, S., Cuthill, I. C., Swaddle, J. P. & Bennett, A. T. D. Ultraviolet vision and band-colour preferences in female zebra finches, *Taeniopygia guttata*. *Anim. Behav.* **54**, 1383–1392 (1997).
36. Balkenius, A. & Kelber, A. Colour constancy in diurnal and nocturnal hawkmoths. *J. Exp. Biol.* **207**, 3307–3316 (2004).
37. Stoddard, M. C. & Prum, R. O. Evolution of avian plumage color in a tetrahedral color space: a phylogenetic analysis of new world buntings. *Am. Nat.* **171**, 755–776 (2008).
38. R Core Team. *R: A Language and Environment for Statistical Computing* <http://www.R-project.org/> (R Foundation for Statistical Computing, Vienna, 2016).
39. Bates, D., Mächler, M., Bolker, B. & Walker, S. Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* **67**, 1–48 (2015).
40. Vorobyev, M., Osorio, D., Bennett, A. T. D., Marshall, N. J. & Cuthill, I. C. Tetrachromacy, oil droplets and bird plumage colours. *J. Comp. Physiol. A* **183**, 621–633 (1998).



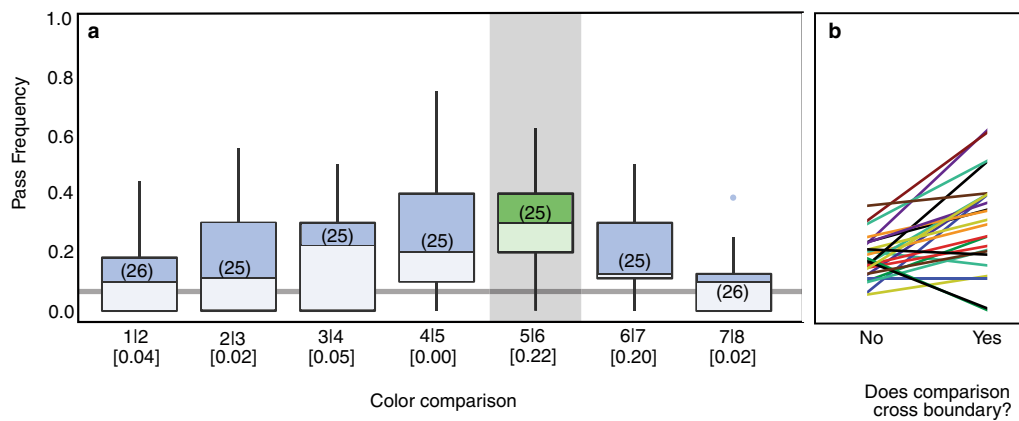
Extended Data Fig. 1 | Downwelling vector irradiance at the level of the foraging grid. Units are photons per cm² per nm. Tungsten bulbs were used to illuminate each cage from a set distance. There was some variation in irradiance between cages. The blue line represents the mean absolute irradiance of our halogen bulbs, and the grey-shaded region indicates one

standard deviation in either direction. The orange line illustrates standard Illuminant A, which we used throughout for visual modelling because it is a standard spectrum (and thus repeatable by other researchers) and closely matched the ambient lighting in our room.



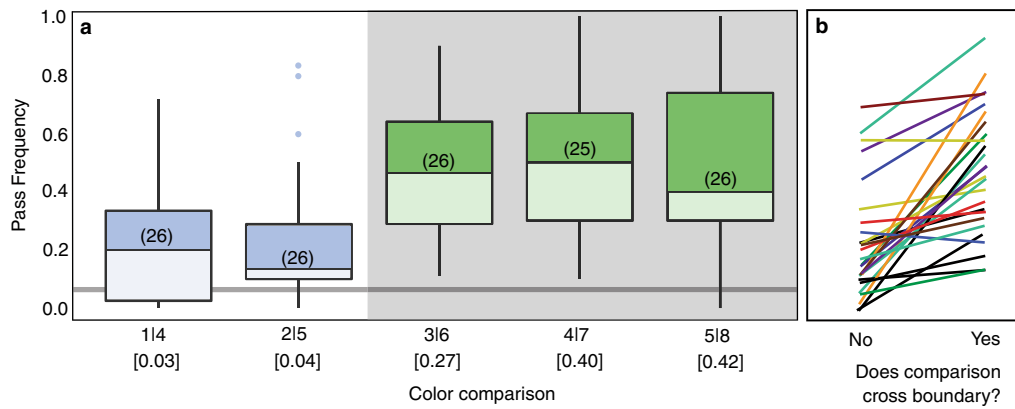
Extended Data Fig. 2 | Reflected radiance of experimental stimuli under experimental lighting conditions. Reflected radiance is in arbitrary energy units; stimulus is Munsell paper covered with an epoxy cover. Below 400 nm, the values are so low that they reach the noise floor of the

spectrometer, suggesting that the use of a trichromatic model (versus one that includes ultraviolet sensitivity) to examine chromatic distance between colours is appropriate.



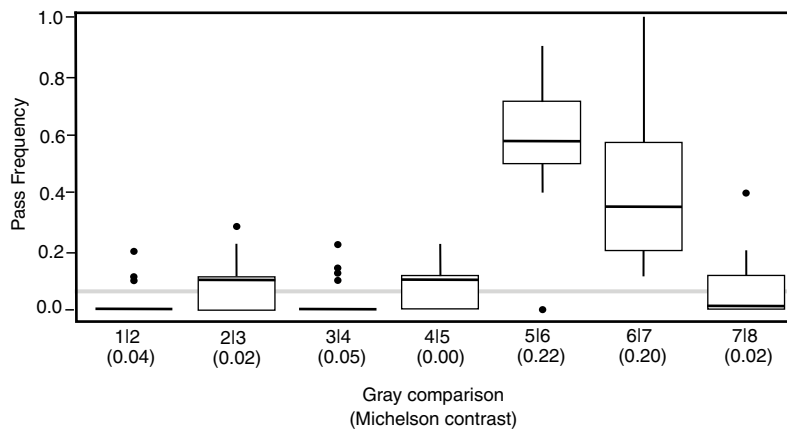
Extended Data Fig. 3 | Results from one-apart discrimination trials. **a, b**, Results show data across all birds (**a**) and for each bird individually (**b**). Overall, birds performed more poorly at all one-apart tasks compared to two- or three-apart tasks, perhaps suggesting that colours that are one step apart neared the physiological limit of colour discrimination. However, the pass rate for the 5|6 comparison (which crossed the category boundary; green box, grey-shaded area) was significantly higher than for comparisons that did not cross the category boundary

(white-shaded areas; blue boxes) (paired t -test, $t_{24} = 4.09$, $P < 0.01$). Numbers in parentheses inside boxes are number of birds that participated in each type of comparison. Box plots depict the median (horizontal line inside the box), 25th and 75th percentiles (box), 25th and 75th percentiles $\pm 1.5 \times$ interquartile range (whiskers) and outliers (circles). The horizontal grey line represents the expected pass frequency if birds flip discs by chance. Numbers in square brackets are Michelson contrast for a given colour pair.



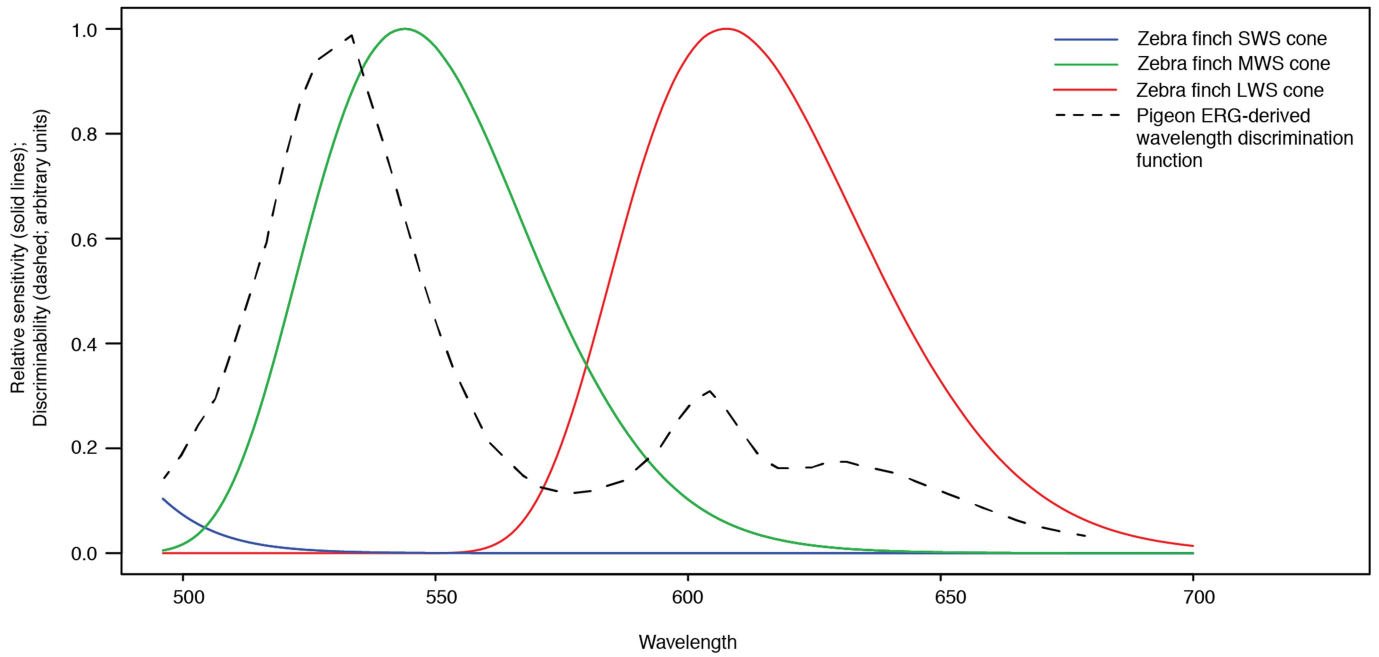
Extended Data Fig. 4 | Results from three-part discrimination trials. **a, b,** Results show data across all birds (**a**) and for each bird individually (**b**). Pass rate was significantly higher for the three comparisons that crossed the boundary (3|6, 4|7 and 5|8; green boxes, grey-shaded area) than for those that did not (1|4 and 2|5; blue boxes, white-shaded area) (paired t -test, $t_{25} = 6.07$, $P < 0.001$). Numbers in parentheses inside boxes

are number of birds that participated in each type of comparison. Box plots depict the median (horizontal line inside the box), 25th and 75th percentiles (box), 25th and 75th percentiles $\pm 1.5 \times$ interquartile range (whiskers) and outliers (circles). The horizontal grey line represents the expected pass rate if birds flip discs by chance. Numbers in square brackets are Michelson contrast values for a given colour pair.



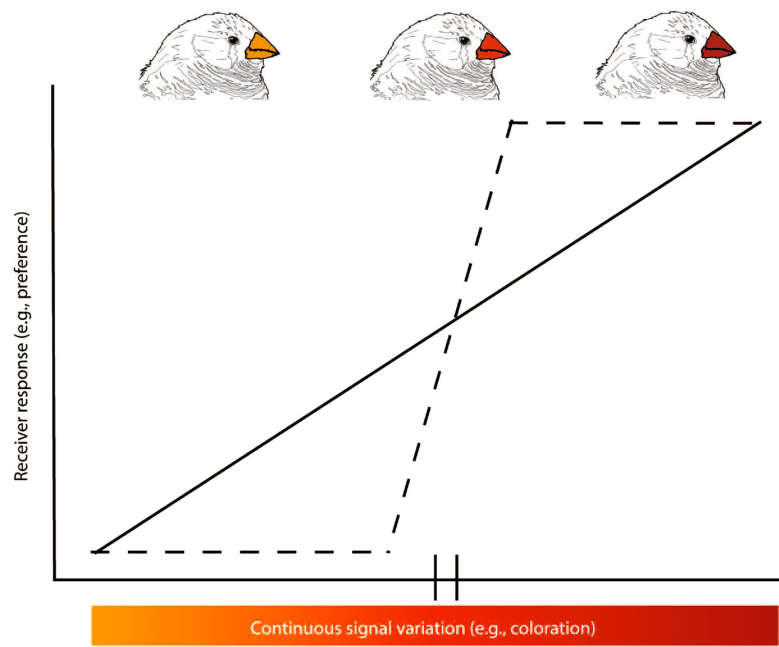
Extended Data Fig. 5 | Results from greyscale (that is, hue information removed) one-apart discrimination experiments. These greyscale experiments did not replicate the categories that we observed when hue information was included, indicating that categories are not structured based on brightness alone. Box plots depict the median (horizontal line inside the box), 25th and 75th percentiles (box), 25th and 75th percentiles $\pm 1.5 \times$ interquartile range (whiskers) and outliers (circles). Numbers

in parentheses below each comparison are Michelson contrast values. Linear mixed models showed that, in our greyscale experiments (that is, without chromaticity information), Michelson contrast between greyscale pairs significantly predicted pass rate. This finding is consistent with the possibility that category formation may be the result of both chromaticity and brightness information (see Table 1). Sample size was 18 birds for all comparisons.



Extended Data Fig. 6 | Wavelength discrimination of avian photoreceptors. The wavelength discrimination function of the pigeon *C. livia* (black dashed line) plotted against the spectral sensitivity peaks of the zebra finch (green, red, blue lines). Original data for the

electroretinography-derived wavelength discrimination function are from Riggs et al.³¹ and have been inverted so that higher numbers indicate greater discrimination.



Extended Data Fig. 7 | Schematic illustrating differences between continuous and categorical perception. Under continuous perception (solid line), receivers perceive and respond in a continuous fashion to signal variation, meaning that any change in a signalling trait is associated with a concomitant change in receiver response. Under categorical perception (dashed line), such as described here for female

zebra finches, receivers show enhanced discrimination of variants across a boundary (hash marks on x axis) compared to equally spaced variants within a category. Zebra finch line drawing by N. Silina licensed under a Attribution 4.0 International (CC BY 4.0) licence (<http://supercoloring.com/pages/zebra-finch>).

Extended Data Table 1 | Photon catch values for the short-, medium- and long-wave cones of the zebra finch

Colour	S	M	L	Double
1 (5R 4/4)	7.73E+06	3.07E+07	9.89E+07	9.89E+07
2 (5R 4/6)	6.99E+06	3.01E+07	1.15E+08	1.08E+08
3 (7.5R 4/6)	6.08E+06	2.83E+07	1.15E+08	1.05E+08
4 (7.5R 4/8)	5.72E+06	2.83E+07	1.38E+08	1.16E+08
5 (10R 4/8)	4.50E+06	2.81E+07	1.38E+08	1.16E+08
6 (10R 5/10)	5.67E+06	4.27E+07	2.20E+08	1.83E+08
7 (10R 6/12)	7.42E+06	6.31E+07	3.35E+08	2.73E+08
8 (2.5YR 6/12)	6.26E+06	6.90E+07	3.42E+08	2.85E+08
Grey 1 (Carbon/Behr)	-	-	-	1.03E+08
Grey 2 (Deep Onyx/Glidden)	-	-	-	1.10E+08
Grey 3 (Black/Behr)	-	-	-	1.06E+08
Grey 4 (Night Club/Behr)	-	-	-	1.15E+08
Grey 5 (Night Club/Behr)	-	-	-	1.15E+08
Grey 6 (Flagstone Gray/Glidden)	-	-	-	1.82E+08
Grey 7 (Flipper/Behr)	-	-	-	2.77E+08
Grey 8 (Downtown Gray/Behr)	-	-	-	2.83E+08

Values were calculated using the reflectance spectrum of each Munsell colour or shade of grey, spectral sensitivity of the zebra finch single and double cones, and standard Illuminant A. Photon catches for Munsell colours are the average of four measurements and have not been relativized. For Munsell colours, the hue group and saturation/chroma are given. For greys, the name and brand of the paint swatch used are given.

Extended Data Table 2 | Chromatic distances between selected Munsell colours under different spectral sensitivities and lighting conditions

Condition	1v2	2v3	3v4	4v5	5v6	6v7
(1) Zebra Finch, Illuminant A	4.2	2.7	4.3	4.7	4.2	2.8
(2) Zebra finch, Daylight	4.3	2.7	4.1	4.8	4.1	2.7
(3) Zebra finch, experimental room lighting (halogen)	4.3	2.7	4.3	4.7	4.2	2.8
(4) Zebra finch, Von Kries adaptation, Illuminant A	4.5	2.3	3.9	4.0	3.5	2.2
(5) Starling, Illuminant A	3.8	1.8	3.5	3.5	3.5	2.1
(6) Average UVS, Illuminant A	3.8	1.8	3.5	3.5	3.5	2.1
(7) Average VS, Illuminant A	4.3	2.1	3.5	3.1	3.7	2.3

Chromatic distances were calculated using (1) zebra finch *Taeniopygia guttata* UVS cone-type retina and standard Illuminant A; (2) zebra finch spectral sensitivity and standard daylight Illuminant D65; (3) zebra finch spectral sensitivity, ambient light spectrum from experimental trials; (4) zebra finch spectral sensitivity, standard Illuminant A, accounting for Von Kries adaptation using a neutral grey background; (5) starling *S. vulgaris* UVS cone-type retina and standard Illuminant A; (6) the average avian UVS cone-type retina with standard Illuminant A; and (7) the average avian VS cone-type retina with standard Illuminant A.

Extended Data Table 3 | Results of a single linear mixed model containing all labelling and discrimination data

Parameter	Estimate	χ^2 (1)	p value
Intercept	0.062		
One Colour Distance	0.020	147	<0.0001
Across 5v6 Boundary	0.21	39.1	<0.0001

As in the model presented in Table 1, this is a linear mixed effects model that includes bird ID as a random effect and the binary indicator (5|6) as random slopes.

Life Sciences Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form is intended for publication with all accepted life science papers and provides structure for consistency and transparency in reporting. Every life science submission will use this form; some list items might not apply to an individual manuscript, but all fields must be completed for clarity.

For further information on the points included in this form, see [Reporting Life Sciences Research](#). For further information on Nature Research policies, including our [data availability policy](#), see [Authors & Referees](#) and the [Editorial Policy Checklist](#).

▶ Experimental design

1. Sample size

Describe how sample size was determined.

Sample sizes were chosen so as to identify a strong statistical and biological effect while minimizing the number of individuals needed (following IACUC protocols). Furthermore, we replicated the experiment over three independent cohorts to ensure reproducibility of our results.

2. Data exclusions

Describe any data exclusions.

We excluded no data from the analyses as long as an individual completed our training protocol or flipped the sufficient number of chips for a trial to be completed. In some cases, an individual never passed our training criteria, in which case we did not progress that individual to experiments. Statements to this effect are made in the text to indicate where and why data have been excluded.

3. Replication

Describe whether the experimental findings were reliably reproduced.

We conducted three experimental replicates on cohorts of 10, 10, and 6 individuals. In each replicate, individuals did not interact with any other individuals that had previously completed the experiment. Furthermore, individuals were unable to see each other complete the experimental task.

4. Randomization

Describe how samples/organisms/participants were allocated into experimental groups.

Individuals were allocated to experimental groups randomly, based on our ability to receive them from Richard Mooney's zebra finch colony. During experimental trials, the order of trial presentation was consistent across individuals, the placement of color discs was randomized for each color comparison using a random sampling function, and the orientation of each disc was haphazard.

5. Blinding

Describe whether the investigators were blinded to group allocation during data collection and/or analysis.

Blinding was not possible in this experiment, given the necessity of observing the individuals flipping the color discs to collect data. For a given experimental day, up to three of five observers (usually, one or two) were collecting data, and no single observer consistently collected data on a given trial. Over the course of the experiment, five observers collected data, and results were consistent across them.

Note: all studies involving animals and/or human research participants must disclose whether blinding and randomization were used.

6. Statistical parameters

For all figures and tables that use statistical methods, confirm that the following items are present in relevant figure legends (or in the Methods section if additional space is needed).

n/a Confirmed

- The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement (animals, litters, cultures, etc.)
- A description of how samples were collected, noting whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- A statement indicating how many times each experiment was replicated
- The statistical test(s) used and whether they are one- or two-sided (note: only common tests should be described solely by name; more complex techniques should be described in the Methods section)
- A description of any assumptions or corrections, such as an adjustment for multiple comparisons
- The test results (e.g. P values) given as exact values whenever possible and with confidence intervals noted
- A clear description of statistics including central tendency (e.g. median, mean) and variation (e.g. standard deviation, interquartile range)
- Clearly defined error bars

See the web collection on [statistics for biologists](#) for further resources and guidance.

► Software

Policy information about [availability of computer code](#)

7. Software

Describe the software used to analyze the data in this study.

We collected and organized data using Microsoft Excel. We analyzed data using only the lme4 package in R software, and thus no custom codes were used. We have noted this in the text.

For manuscripts utilizing custom algorithms or software that are central to the paper but not yet described in the published literature, software must be made available to editors and reviewers upon request. We strongly encourage code deposition in a community repository (e.g. GitHub). *Nature Methods* [guidance for providing algorithms and software for publication](#) provides further information on this topic.

► Materials and reagents

Policy information about [availability of materials](#)

8. Materials availability

Indicate whether there are restrictions on availability of unique materials or if these materials are only available for distribution by a for-profit company.

No unique materials were used.

9. Antibodies

Describe the antibodies used and how they were validated for use in the system under study (i.e. assay and species).

No antibodies were used.

10. Eukaryotic cell lines

a. State the source of each eukaryotic cell line used.

No eukaryotic cell lines were used.

b. Describe the method of cell line authentication used.

No eukaryotic cell lines were used.

c. Report whether the cell lines were tested for mycoplasma contamination.

No eukaryotic cell lines were used.

d. If any of the cell lines used are listed in the database of commonly misidentified cell lines maintained by [ICLAC](#), provide a scientific rationale for their use.

No commonly misidentified cell lines were used.

► Animals and human research participants

Policy information about [studies involving animals](#); when reporting animal research, follow the [ARRIVE guidelines](#)

11. Description of research animals

Provide details on animals and/or animal-derived materials used in the study.

We used lab-raised, female zebra finches (*Taeniopygia guttata*) in this study. Birds were all sexually mature, between 3 and 50 months old at the start of the experiment. All birds originated from a colony maintained by Dr. Richard Mooney at Duke University. All birds were covered under Duke University IACUC protocol A004-17-01.

Policy information about [studies involving human research participants](#)

12. Description of human research participants

Describe the covariate-relevant population characteristics of the human research participants.

This study did not involve human participants.