# Age-Tailored Colour Palettes for Inclusive Visual Communication

Akvile Sinkeviciute
Computing
Northeastern University London
Devon House, 58 St Katharine's Way
London E1W 1LP, UK
*a.sinkeviciute@nulondon.ac.uk*

Dimitris Mylonas
Computing
Northeastern University London
Devon House, 58 St Katharine's Way
London E1W 1LP, UK
*dimitris.mylonas@nulondon.ac.uk*

**Colour discrimination and vocabulary both decline with age, undermining visual communication that relies on colour. We ask which colour palettes most effectively support colour communication across age groups. Analysing 103,294 unconstrained colour-naming responses from an online experiment with 2,418 English-speaking participants aged 16 to 87, we identified focal colours for a younger (16–39) and a mature (40+) group. Each group's focal colours were ranked by three criteria: perceptual discriminability ($\Delta E_{00}$), categorical distinction ($\Delta h$) and codability ($-H$), to construct three palettes per group, evaluated against each age group's own naming model. We also scored 10 widely used categorical palettes under both models to check for an existing age bias. The focal colours shifted with age ($\Delta E_{00} = 5.7$ on average), and colour-naming vocabulary richness declined. Age-tailoring robustly improved categorical distinction for the younger group out of sample; for the mature group, whose smaller vocabulary left less room to tailor, the benefit held only on the data it was built from. A majority (8/10) of widely used external categorical palettes were more codable for younger than for mature speakers, motivating age-tailored palettes for inclusive visual communication.**

*Colour naming, Colour palettes, Data visualisation, Ageing, Visual communication*

## 1. INTRODUCTION

Populations are ageing, increasing the need for designs that remain usable across the lifespan (World Health Organization 2015), and inclusive-design research has long argued that interfaces should accommodate the changing abilities of older users (Gregor et al. 2002; Newell et al. 2011). This concern extends to data visualisation, where colour is central: chromatic discrimination declines across the lifespan, most steeply along the yellow–blue axis (Pokorny et al. 1987; Knoblauch et al. 2001; Paramei and Oakley 2014), as the lens yellows (Weale 1988; Werner 1982), the pupil narrows (Winn et al. 1994) and S-cone pathways weaken (Shinomori and Werner 2012). The language of colour changes too, and people use fewer distinct colour terms as they age (Griber et al. 2021; Sinkeviciute et al. 2024).

For colour to communicate information, viewers must both discriminate colours and agree on what to call them, an agreement that must survive changes in display and observer (Mylonas and MacDonald 2012). Colour-naming distributions have been used to construct palettes for visualisation directly: a probabilistic colour-naming model built from a large crowdsourced naming survey can support colour selection and palette design (Heer and Stone 2012a; Munroe 2010), and categorical palettes can be generated to balance perceptual discriminability with preference (Gramazio et al. 2017). Such naming-based methods pool responses across the population, yet the colour categories they rest on are not universal: they have been shown to vary across languages (Kim et al. 2019), and, we argue, across ages. Because younger participants both dominate most crowdsourced naming datasets and command richer colour vocabularies (Griber et al. 2021; Sinkeviciute et al. 2024), a pooled naming model is implicitly tuned to younger namers and may not serve older viewers well.

In this study, we rank colours into palettes by three criteria, one perceptual and two linguistic. Perceptual discriminability is whether colours can be told apart visually (Szafir 2018), the baseline that precedes naming. The two linguistic criteria capture different aspects of colour naming. Categorical distinction is a **between-colour** measure: the degree to which two colours fall into separate

*© Sinkeviciute et al. Published by BCS Learning and Development Ltd.*

naming categories rather than overlapping as near-synonyms (Mylonas et al. 2021). Codability is a **within-colour** measure: the degree to which a single colour's own naming concentrates on one term or splits across several, with consistently named colours recalled and communicated more reliably (Brown and Lenneberg 1954; Lantz and Stefflre 1964) and ambiguous ones available to de-emphasise information (Mylonas and MacDonald 2017). Because age reshapes both colour vision and colour language, and because each of these criteria can be estimated separately for a given group of namers, they raise the question that motivates this work: **what are the most effective colour palettes for facilitating visual communication across age groups?**

## 2. METHODS

### 2.1. Participants

Colour-naming responses were collected between 2018 and 2024 from English-speaking participants, recruited both as unpaid volunteers (via social media and word of mouth) and as paid participants (via Prolific) to ensure adequate representation of older age groups (40+); the two samples were combined, as task performance is comparable between paid and unpaid respondents in carefully designed tasks (Mao et al. 2013). From an initial 2,619 participants (119,002 responses), we excluded participants, in sequence, for (1) congenital colour vision deficiency on the Dynamic Color Vision Test, with more conservative thresholds for female participants, given the much lower prevalence of red–green deficiency in females ($\sim$0.5 per cent) than males ($\sim$8 per cent) (Birch 2012); (2) systematically anomalous naming; (3) an age under 16 or missing; and (4) upbringing outside a native English-speaking country (United Kingdom, Canada, United States, Australia, New Zealand), a purely linguistic criterion not based on ethnicity or national origin. These participant-level exclusions left 108,560 responses. The 2,418 remaining participants were divided into the two age groups detailed in Table 1. The final sample had a mean age of 28.34 years ($\mathrm{SD} = 14.50$, range 16–87): 1,562 female (64.6 per cent), 714 male (29.5 per cent) and 139 (5.7 per cent) identifying with other gender categories. All participants gave informed consent online (Varnhagen et al. 2005). The study received continuous ethical approval from 20 January 2017, initially from University College London (3387/001; renewed annually) and from the current host institution from 19 May 2023 (No. 0001; renewed annually).

***Table 1:*** *Participants and responses by age bin. The sample is divided into younger and mature groups at the study's 40-year cutoff (Section 2.4).*

| Age bin | Participants | Responses | Age, mean (SD) |
|---|---|---|---|
| Younger | 1,980 | 58,023 | 22.3 (5.9) |
| Mature | 438 | 45,271 | 55.5 (10.3) |

### 2.2. Stimuli and procedure

Colour naming was collected through a web-based experiment (available at colornaming.net) whose full five-step procedure has been described previously (Mylonas and MacDonald 2012; Mylonas et al. 2026) (Figure 1). Participants first adjusted their display until all 11 steps of a greyscale ramp were visible, then completed a web-based dynamic colour-vision screening test. In the naming task, on each trial a participant viewed, in random order, a single uniformly coloured, 2-degree disc against a neutral grey background and produced any descriptor (a single word, a compound or a term with modifiers) with no predefined options. Participants then reported their device, viewing conditions and demographic details and finally received a summary of their responses. The 606 simulated Munsell stimuli were approximately uniformly distributed across the colour solid.

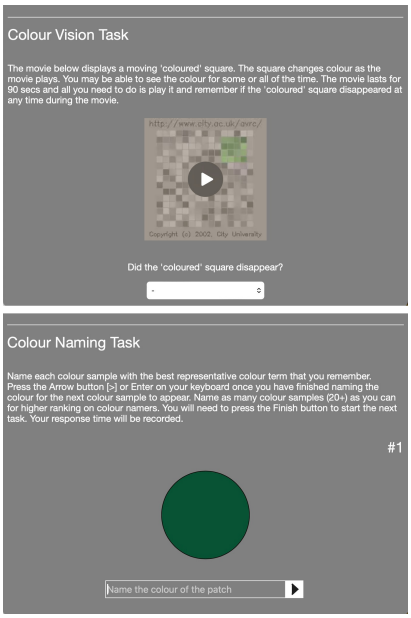


***Figure 1:*** *The web-based colour-naming experiment interface (colornaming.net), showing Steps 2–3 of the five-step procedure: the colour-vision screening test and the unconstrained naming task in which participants name each colour disc shown on a neutral grey background.*

### 2.3. Colour-name cleaning

Before analysis, the 108,560 colour-name responses were cleaned. Names were lower-cased, stripped

of punctuation, with separators standardised to underscores, single characters and digits removed, spelling corrected against a 765-entry dictionary and the American 'gray' mapped to 'grey'. We then discarded responses that were empty, only a single character, composed of repeated letters, without vowels, or produced by just one participant (idiosyncratic to model reliability); this removed 5,266 responses in total (4.9 per cent, of which 4,775 were single-participant terms), leaving 103,294 analysed responses across 1,950 unique colour terms.

### 2.4. Identifying focal colours by age group

To build the age-tailored palettes, we identified, separately for each group, the focal colours, namely the best examples of an optimal set of colour names (Mylonas et al. 2026) essential to label all colours. Constructing focal colours requires aggregating responses within groups; we therefore dichotomised at 40 years, above which yellow–blue thresholds rise more steeply due to reduced short-wavelength retinal illuminance with age (Barbur and Rodriguez-Carmona 2016), following a lifespan minimum in tritan discrimination in the mid-to-late twenties (Knoblauch et al. 2001). This cutoff also reflects the sample size available for finer subdivision: 438 mature participants versus 1,980 younger participants (Table 1).

### 2.5. Palette construction and evaluation

**Criteria.** We used three criteria, each capturing a different aspect of what makes a categorical palette effective. Perceptual discriminability ($\Delta E_{00}$) is the mean pairwise CIEDE2000 colour difference (Luo et al. 2001); being observer-independent, it serves as a control that cannot by construction express an age effect, and we therefore report it once and do not treat it as an age-sensitive measure. The other two are linguistic and can therefore vary with age. Both are computed from each colour's naming distribution $P(w \mid c)$, the probability of name $w$ given colour $c$ estimated from the naming responses (Chuang et al. 2008): categorical distinction ($\Delta h$) is the mean pairwise Hellinger distance between these distributions, rising as they diverge, and codability ($-H$) is the mean negative Shannon entropy of each, rising towards zero as a colour's naming concentrates on one term. These criteria follow established naming-derived measures (Heer and Stone 2012a; Gramazio et al. 2017), but we estimated them from our own unconstrained-naming data, kept them as separate rankings and made them age-specific.

**Construction.** We ranked each group's age-specific focal colours by each criterion to form the palettes shown in Figure 2. The achromatic terms (black, white and grey) were retained, as their strong naming consensus ranks them among the top colours for codability and categorical distinction.

**Evaluation.** Palettes were assessed at 21 colours (the size of the mature focal-colour set, the largest equal size both groups support) on all three criteria. Perceptual discriminability ($\Delta E_{00}$) is observer-independent, so each palette takes a single value; categorical distinction ($\Delta h$) and codability ($-H$) depend on how colours are named and are therefore scored against each age group's naming model. Here $\Delta h$ is scored as the mean Hellinger distance from each palette colour to all focal colours of the model group, rather than the within-palette pairwise distance used in the between-group comparison. Higher $\Delta E_{00}$ and $\Delta h$, and $-H$ nearer zero, are better. To test whether the evaluation generalised beyond the participants the models were built on, we ran split-half cross-validation. We re-derived focal colours and naming models on a random half of the participants and tested on the held-out half, over 200 splits. We repeated this with two further schemes that vary the training/test ratio, and also tested the 40-year dichotomisation point itself by repeating the same in-sample analysis at cutoffs of 50 and 65 (Section 3.2). We additionally scored 10 widely used external categorical palettes (Table 4) against both naming models using these same three criteria, following a previously published approach (Heer and Stone 2012b,a).

**AI use.** Generative AI (Claude, Anthropic) edited text and code the authors had already drafted and helped build the supplementary HTML demo. All analysis decisions and interpretation are the authors' own.

## 3. RESULTS

### 3.1. Age effects on colour naming

Comparing younger (16–39) and mature (40+) namers, three changes in colour naming accompanied age (Figure 3). First, the colour lexicon contracted: the younger group supported 31 focal colours against the mature group's 21, and vocabulary richness, quantified as the conditional entropy of colour naming $H(W \mid C)$, declined with age ($r = -0.69$, $p < .001$ across integer ages). Second, the focal colours shifted: across the 19 categories named by both groups, the focal colours differed by a mean of $\Delta E_{00} = 5.73$ ($\mathrm{SD} = 4.86$; one-sample $t(18) = 5.14$, $p < .001$), a reliable age-related shift. Third, category boundaries moved: the younger group named 'cyan' and 'turquoise' as two separate, neighbouring categories, but the mature group had only 'turquoise', anchored at the very chip the younger group called

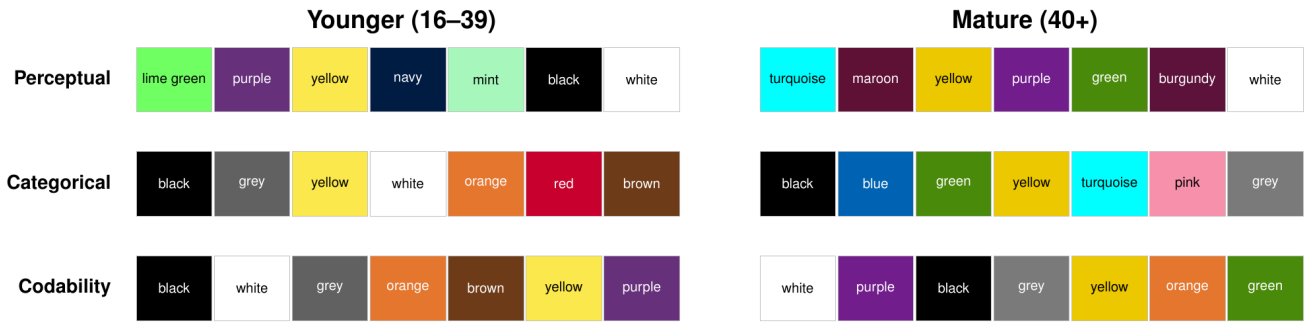


***Figure 2:*** *Age-tailored palettes for the younger (16–39) and mature (40+) groups, each ranked by perceptual discriminability ($\Delta E_{00}$), categorical distinction ($\Delta h$) and codability ($-H$); the leading seven focal colours per criterion are shown. Each row is one criterion; compare the younger (left) and mature (right) blocks within a row to see how membership and hue shift with age.*

'cyan'. This indicates that 'cyan' was absorbed into the neighbouring 'turquoise' category with age.

Comparing the two groups' focal colours on the two linguistic criteria (setting aside the achromatic black, white and grey, which are conventionally reserved for text and background elements in visualisation design rather than used as data colours, and whose near-zero chroma also carries no hue information), the younger focal colours were equally codable ($-H$: $-1.81$ vs $-1.84$; $t(31.8) = 0.11$, $p = .92$) but significantly more categorically distinct ($\Delta h$: $0.92$ vs $0.91$; Welch's $t(28.7) = 2.29$, $p = .03$).[1] Categorical distinction was thus the only one of the three criteria to differ reliably between age groups.

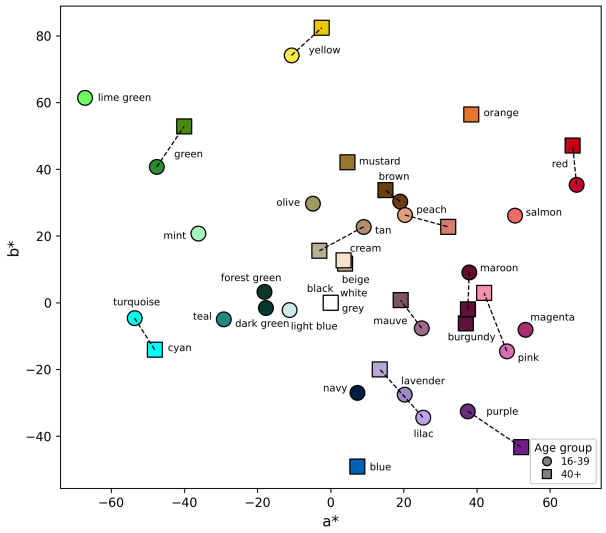


***Figure 3:*** *Focal colours for the younger (circles) and mature (squares) groups in the $a^*b^*$ plane. Dashed lines link the two focal colours of each colour name shared by both groups.*

[1]The observer-independent control showed no reliable between-group difference in perceptual spread ($\Delta E_{00}$: $42.71$ vs $39.58$; Welch's $t(34.5) = 1.67$, $p = .10$).

### 3.2. Palette evaluation

***Table 2:*** *Mean scores per 21-colour palette, under each age group's naming model.*

| | | Younger model | | Mature model | |
|---|---|---|---|---|---|
| Palette | $\Delta E_{00}$ | $\Delta h$ | $-H$ | $\Delta h$ | $-H$ |
| Agnostic | 44.68 | 0.96 | $-1.36$ | 0.95 | $-1.60$ |
| Younger | 50.10 | 0.96 | $-1.24$ | 0.94 | $-1.66$ |
| Mature | 41.67 | 0.95 | $-1.64$ | 0.96 | $-1.65$ |

The age effect appears in categorical distinction (Table 2): each palette was more categorically distinct under its own group's model than the other's. This held for the younger palette ($\Delta h = 0.97$ vs $0.94$; Wilcoxon $p < .001$), and went in the same direction for the mature palette ($0.96$ vs $0.95$) although it did not reach significance ($p = .27$). The younger advantage was robust, holding at every palette size tested (7–21 colours, all $p < .05$); the mature trend was significant at two of four sizes (7 and 15; Table 3). The younger crossover survives Bonferroni correction across these tests, whereas the mature results are reported as trends rather than corrected effects. Cross-validation confirmed this asymmetry out of sample: the younger crossover held in 191 of 200 splits (sign test $p < .001$), whereas the mature crossover was not reliable (9 of 200 splits), consistent with the smaller mature lexicon. The same asymmetry persisted under two further schemes with larger training folds (4-fold and 10-fold): the held-out mature crossover never exceeded 13 per cent of folds, while thin test folds (fewer than five responses per colour) rose from 0 per cent (split-half) to 100 per cent (10-fold), confirming the mature null is a small-sample ceiling bound by the group's 438 participants, not an artefact of the evaluation scheme. The 40-year dichotomisation point is similarly robust: splitting at 50 or 65 instead reproduces the same in-sample crossover pattern ($p < .05$ throughout). The age-agnostic palette fell

between the two under both models, optimal for neither.

***Table 3:*** *Mean categorical distinction ($\Delta h$) per palette, under its own naming model and the other group's, by size.* $^{*}\,p < .05$, $^{**}\,p < .01$, $^{***}\,p < .001$.

| Palette | Size | $\Delta h$ own | $\Delta h$ other | Wilcoxon $p$ |
|---|---|---|---|---|
| Younger | 7 | 0.984 | 0.952 | $0.016^{*}$ |
| Younger | 10 | 0.979 | 0.946 | $0.002^{**}$ |
| Younger | 15 | 0.972 | 0.946 | $0.002^{**}$ |
| Younger | 21 | 0.965 | 0.942 | $0.001^{***}$ |
| Mature | 7 | 0.979 | 0.955 | $0.031^{*}$ |
| Mature | 10 | 0.975 | 0.961 | $0.131$ |
| Mature | 15 | 0.967 | 0.954 | $0.035^{*}$ |
| Mature | 21 | 0.955 | 0.949 | $0.272$ |

We then scored 10 widely used categorical palettes (Heer and Stone 2012b), including ColorBrewer (Harrower and Brewer 2003) and Tableau, under both age models (Table 4). For a fixed external palette, codability is the only one of the three measures that can separate the age models: perceptual discriminability is observer-independent and so identical under both, and categorical distinction showed no age bias across the 10 palettes (median $\Delta h$ gap $-0.003$; Wilcoxon signed-rank $p = .32$, ns). On codability, existing palettes were more codable under the younger model (median $-H$ gap $0.16$), Wilcoxon signed-rank $p = .02$ (rank-biserial $r = .82$, Cohen's $d = .95$); existing palettes reliably suit younger namers.

***Table 4:*** *Codability ($-H$) of 10 widely used categorical palettes under each age group's model. Gap $= -H$(younger)$- -H$(mature), positive favouring younger; bold marks the more codable value per row.*

| | | Codability ($-H$) | | |
|---|---|---|---|---|
| Palette | Size | Younger | Mature | Gap |
| Map of the Market | 7 | $\mathbf{-1.14}$ | $-1.52$ | $+0.38$ |
| ColorBrewer-Q9 | 9 | $\mathbf{-1.12}$ | $-1.37$ | $+0.25$ |
| Numbers-6 | 6 | $\mathbf{-1.45}$ | $-1.69$ | $+0.24$ |
| ColorBrewer-Q10 | 10 | $\mathbf{-1.73}$ | $-1.91$ | $+0.18$ |
| Tableau-10 | 10 | $\mathbf{-1.46}$ | $-1.62$ | $+0.16$ |
| The Economist | 8 | $\mathbf{-1.94}$ | $-2.10$ | $+0.16$ |
| Excel-10 | 10 | $\mathbf{-1.95}$ | $-2.08$ | $+0.13$ |
| ManyEyes | 19 | $\mathbf{-2.29}$ | $-2.33$ | $+0.04$ |
| Tableau-20 | 20 | $-1.88$ | $-1.88$ | $+0.00$ |
| ColorBrewer-Q12 | 12 | $-2.30$ | $\mathbf{-2.15}$ | $-0.15$ |

### 3.3. Palette application

We apply the categorical-distinction palette to bar, line, scatter and donut charts on synthetic data, and, using Altair (VanderPlas et al. 2018), to three real, geographically grounded visualisations built on real unemployment and transit data from vega-datasets (Vega 2023) (a US state-level choropleth, a US county-level choropleth and the London Underground network) for both groups at $n = 3$ and $n = 10$ categories. Achromatic colours are excluded throughout, as in §3.1. At three categories the younger and mature palettes are clearly distinct, since the leading focal colours of the two models differ, so identical data is coloured differently depending on whose naming model built the palette. As the number of categories grows the palettes draw on increasingly overlapping focal colours and the contrast narrows, since both groups saturate at larger sizes, echoing that the advantage is strongest at small palette sizes (Table 3).

A supplementary interactive HTML file, the **Palette Explorer**, lets a reader explore every palette across all three criteria (perceptual discriminability, categorical distinction and codability) and across every visualisation type above, for either age group and any number of colours (Figure 4); it is available at `https://osf.io/3jec4/`.

## 4. DISCUSSION

Our findings demonstrate that categorical distinction is the most effective criterion for constructing age-tailored colour palettes. Whereas perceptual discriminability was intentionally observer-independent and codability showed little difference between age groups, only categorical distinction consistently improved palette performance, suggesting that age-related variation in colour communication is expressed primarily through shifts in colour category boundaries rather than changes in naming consistency. In doing so, our findings extend previous evidence that colour categories vary across languages (Kim et al. 2019; Mylonas et al. 2026) to demonstrate that they also vary systematically across age groups within the same language.

The contrasting behaviour of the three criteria helps explain this pattern. Perceptual discriminability, an observer-independent baseline, could not by design capture age-related differences in colour naming. Codability remained remarkably stable between the younger and mature groups' focal colours despite the overall reduction in colour vocabulary that accompanies ageing (Griber et al. 2021; Sinkeviciute et al. 2024): although older adults use fewer colour terms, they remain just as consistent in naming their retained categories' best examples, consistent with evidence that best-example colour loci stay stable across the lifespan (Enoch et al. 1999). Categorical distinction, by contrast, depends on how neighbouring colours partition into lexical categories, and these boundaries shifted with age, so tailoring palettes on this criterion better aligned colours with each group's naming system. The benefit was strongest for the younger palette; the mature palette showed only directional improvements, likely

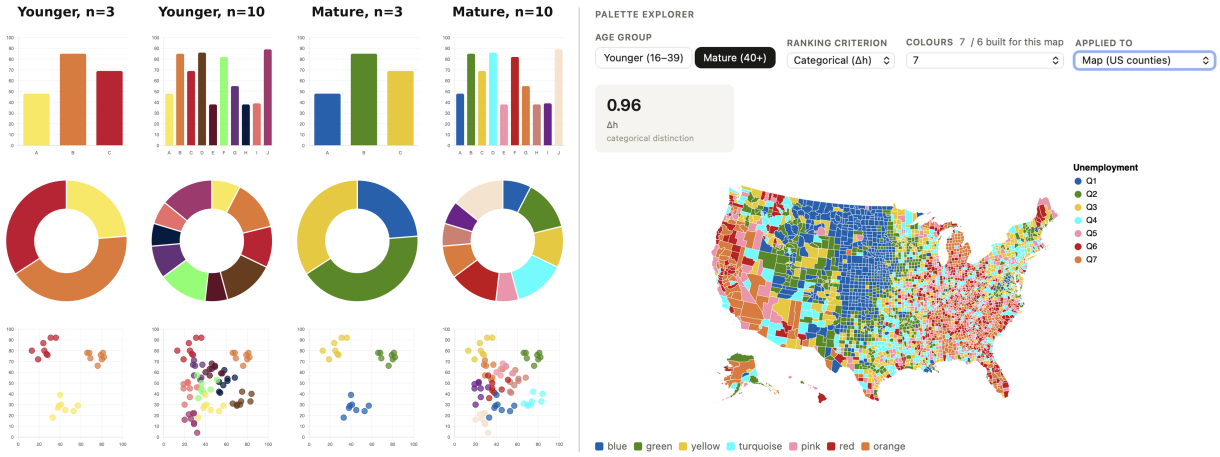


***Figure 4:*** *The categorical-distinction palette applied to common chart types for both groups at $n = 3$ and $n = 10$ (left), and, from the* ***Palette Explorer*** *supplementary demo, to the US county-level map (right; mature group, $n = 7$).*

reflecting its smaller focal-colour set (21 vs. 31). We therefore treat the mature palette's benefit as preliminary, consistent with the cross-validation analyses (§3.2).

Our evaluation of widely used categorical palettes confirms the concern that motivated this study: pooled naming data implicitly reflects the younger namers who dominate crowdsourced datasets (Griber et al. 2021; Sinkeviciute et al. 2024). Extending previous naming-based approaches to palette construction (Heer and Stone 2012a; Gramazio et al. 2017), our findings show that effective colour communication depends on the population interpreting the colours, not only on the colours selected. Consistently named colours already support reliable recall and communication (Brown and Lenneberg 1954; Lantz and Stefflre 1964); age-tailoring adds a further advantage when the target audience is known, preserving age-specific categorical distinctions in line with broader calls for inclusive design that accommodates lifespan changes in user abilities (Gregor et al. 2002; Newell et al. 2011).

Because age-tailoring changes only the colours assigned to categories, it offers a low-cost, drop-in intervention requiring no new hardware or interaction design, as demonstrated by the Palette Explorer (Figure 4). The same evaluation (Table 4) can also be incorporated into accessibility audits alongside existing colour-vision-deficiency checks (Wong 2011), addressing an aspect of age-related inclusivity current guidance (World Wide Web Consortium 2023) does not yet cover.

Several limitations point to future work. First, these palettes still need behavioural validation of communication accuracy, efficiency and recall. Second, larger older-adult samples would enable finer modelling within the 40+ range, including a possible split around age 60, where lens yellowing accelerates (Pokorny et al. 1987) — distinct from our 40-year cutoff; our current sample cannot support such finer bins, and falls further short of continuous age: with few responses per age, focal-colour sets would vary sharply between ages, reflecting sampling noise, not ageing. Finally, our cross-sectional, English-speaking sample leaves open whether these patterns generalise across languages and cultures.

## 5. CONCLUSION

We asked which colour palettes best facilitate visual communication across age groups. As colour vocabulary contracted with age, the number of focal colour categories also decreased, from younger to older participants. The focal categories shared between the two age groups also differed significantly. We generated age-specific palettes from each group's focal colours and ranked them using perceptual discriminability, codability and categorical distinction. Perceptual discriminability was fixed by construction and therefore served as the baseline ranking method. Of the two linguistic measures, only categorical distinction improved palette performance. Ten established colour palettes also showed a significant codability bias towards younger namers. Categorical distinction is therefore the most beneficial tailoring criterion, and existing pooled palettes warrant caution given their bias towards younger namers.

## ACKNOWLEDGEMENTS

DM was supported by the Leverhulme Trust (Cones2Words project, RPG-2024-096).

## REFERENCES

Barbur, J. L. and Rodriguez-Carmona, M. (2016). Color vision changes in normal aging. In Elliot, A. J., Fairchild, M. D., and Franklin, A., editors, *Handbook of Color Psychology*, pages 180–196. Cambridge University Press, Cambridge, UK.

Birch, J. (2012). Worldwide prevalence of red-green color deficiency. *Journal of the Optical Society of America A*, 29(3):313–320.

Brown, R. W. and Lenneberg, E. H. (1954). A study in language and cognition. *Journal of Abnormal and Social Psychology*, 49(3):454–462.

Chuang, J., Stone, M., and Hanrahan, P. (2008). A probabilistic model of the categorical association between colors. In *Color and Imaging Conference*, volume 16, pages 6–11. Society for Imaging Science and Technology.

Enoch, J. M., Werner, J. S., Haegerstrom-Portnoy, G., Lakshminarayanan, V., and Rynders, M. (1999). Forever young: Visual functions not affected or minimally affected by aging: A review. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*, 54(8):B336–B351.

Gramazio, C. C., Laidlaw, D. H., and Schloss, K. B. (2017). Colorgorical: Creating discriminable and preferable color palettes for information visualization. *IEEE Transactions on Visualization and Computer Graphics*, 23(1):521–530.

Gregor, P., Newell, A. F., and Zajicek, M. (2002). Designing for dynamic diversity: interfaces for older people. In *Proceedings of the Fifth International ACM Conference on Assistive Technologies (ASSETS '02)*, pages 151–156. ACM.

Griber, Y. A., Mylonas, D., and Paramei, G. V. (2021). Intergenerational differences in Russian colour naming in the globalized era: a linguistic analysis. *Humanities and Social Sciences Communications*, 8(1):1–19.

Harrower, M. and Brewer, C. A. (2003). ColorBrewer.org: An online tool for selecting colour schemes for maps. *The Cartographic Journal*, 40(1):27–37.

Heer, J. and Stone, M. (2012a). Color naming models for color selection, image editing and palette design. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '12)*, pages 1007–1016. ACM.

Heer, J. and Stone, M. (2012b). Color palette analyzer. Stanford Visualization Group. `http://vis.stanford.edu/color-names/analyzer/`.

Kim, Y., Thayer, K., Gorsky, G. S., and Heer, J. (2019). Color names across languages: Salient colors and term translation in multilingual color naming models. In *EuroVis 2019 – Short Papers*. The Eurographics Association.

Knoblauch, K., Vital-Durand, F., and Barbur, J. L. (2001). Variation of chromatic sensitivity across the life span. *Vision Research*, 41(1):23–36.

Lantz, D. and Stefflre, V. (1964). Language and cognition revisited. *Journal of Abnormal and Social Psychology*, 69(5):472–481.

Luo, M. R., Cui, G., and Rigg, B. (2001). The development of the CIE 2000 colour-difference formula: CIEDE2000. *Color Research & Application*, 26(5):340–350.

Mao, A., Kamar, E., Chen, Y., Horvitz, E., Schwamb, M. E., Lintott, C. J., and Smith, A. M. (2013). Volunteering versus work for pay: Incentives and tradeoffs in crowdsourcing. In *Proceedings of the First AAAI Conference on Human Computation and Crowdsourcing (HCOMP)*, pages 94–102.

Munroe, R. (2010). Color survey results. `http://blog.xkcd.com/2010/05/03/color-survey-results/`. Accessed: 2026.

Mylonas, D., Ahmed, R., Sinkeviciute, A., and Koliousis, A. (2026). Colour in translation: Data, models, and benchmarking for cross-linguistic colour naming. In *Proceedings of the 2026 CHI Conference on Human Factors in Computing Systems (CHI '26)*, Barcelona, Spain. ACM.

Mylonas, D., Koliousis, A., and Uusküla, M. (2021). Synonymy in the language of colour. In *Proceedings of the 14th Congress of the International Colour Association (AIC)*, Milan, Italy.

Mylonas, D. and MacDonald, L. (2012). Colour naming for colour communication. In *Colour Design*, pages 254–270. Woodhead Publishing.

Mylonas, D. and MacDonald, L. W. (2017). Colour naming for colour design. In Best, J., editor, *Colour Design: Theories and Applications*. Woodhead Publishing, Cambridge, UK, 2 edition.

Newell, A. F., Gregor, P., Morgan, M., Pullin, G., and Macaulay, C. (2011). User-sensitive inclusive design. *Universal Access in the Information Society*, 10(3):235–243.

Paramei, G. V. and Oakley, B. (2014). Variation of color discrimination across the life span. *Journal of the Optical Society of America A*, 31(4):A375–A384.

Pokorny, J., Smith, V. C., and Lutze, M. (1987). Aging of the human lens. *Applied Optics*, 26(8):1437–1440.

Shinomori, K. and Werner, J. S. (2012). Aging of human short-wave cone pathways. *Proceedings of the National Academy of Sciences*, 109(33):13422–13427.

Sinkeviciute, A., Ferguson, H., and Mylonas, D. (2024). Ageing and colour language in English speakers. In Tekavčič Pompe, M., Vrabič, N., and Rozman, A., editors, *27th International Colour Vision Society Meeting: Book of Abstracts with Programme*, page 201, Ljubljana. University Eye Clinic; International Colour Vision Society.

Szafir, D. A. (2018). Modeling color difference for visualization design. *IEEE Transactions on Visualization and Computer Graphics*, 24(1):392–401.

VanderPlas, J., Granger, B. E., Heer, J., Moritz, D., Wongsuphasawat, K., Satyanarayan, A., Lees, E., Timofeev, I., Welsh, B., and Sievert, S. (2018). Altair: Interactive statistical visualizations for python. *Journal of Open Source Software*, 3(32):1057.

Varnhagen, C. K., Gushta, M., Daniels, J., Peters, T. C., Parmar, N., Law, D., Hirsch, R., Sadler Takach, B., and Johnson, T. (2005). How informed is online informed consent? *Ethics & Behavior*, 15(1):37–48.

Vega (2023). vega-datasets. `https://github.com/vega/vega-datasets`. County-level unemployment data originally from the U.S. Bureau of Labor Statistics; public domain. Accessed 2026.

Weale, R. A. (1988). Age and the transmittance of the human crystalline lens. *The Journal of Physiology*, 395:577–587.

Werner, J. S. (1982). Development of scotopic sensitivity and the absorption spectrum of the human ocular media. *Journal of the Optical Society of America*, 72(2):247–258.

Winn, B., Whitaker, D., Elliott, D. B., and Phillips, N. J. (1994). Factors affecting light-adapted pupil size in normal human subjects. *Investigative Ophthalmology & Visual Science*, 35(3):1132–1137.

Wong, B. (2011). Points of view: Color blindness. *Nature Methods*, 8(6):441.

World Health Organization (2015). *World Report on Ageing and Health*. World Health Organization, Geneva.

World Wide Web Consortium (2023). Web content accessibility guidelines (wcag) 2.2. W3C Recommendation, 5 October 2023. `https://www.w3.org/TR/WCAG22/`.