A Procedural Method for Generating Pseudoisochromatic Plates in the Browser, and the Calibration Limits of Screen-Based Color-Vision Screening

Vince Gonzalez

Independent researcher · OpticQuiz (opticquiz.com)

ORCID: 0009-0005-3640-014X

Correspondence: vincegonzalez@me.com

Code(archived): https://doi.org/10.5281/zenodo.21311369
· source: github.com/zengineco/opticquiz.com

Abstract

Background. Pseudoisochromatic plate tests, exemplified by the Ishihara plates, remain the most widely used screen for red-green color vision deficiency. Online reproductions are now abundant, but most reuse a small set of fixed, scanned plate images, and few state clearly what a consumer display can and cannot measure.

Method. This paper specifies *Procedural Pseudoisochromatic Plate Generation* (PPPG), a method for generating pseudoisochromatic plates entirely in the web browser at run time. A target figure is rendered as a binary mask; a spatial-hash dot-packing routine fills the plate with non-overlapping dots; each dot is colored so that figure and background are constructed to differ predominantly along a red-green or blue-yellow (tritan) confusion direction; and dot lightness is randomized within a shared band on every run so that no *designed* luminance cue distinguishes figure from ground. No plate is stored as an image, and no two runs are identical.

Limitations. The method's ceiling is not its design but the display. On an uncalibrated, three-primary emissive screen with unknown transfer function and 8-bit-per-channel quantization, the intended chromatic relationships are not the relationships seen; the deficiency is further compounded by *observer metamerism*, because a match constructed for the standard observer is not guaranteed to hold for the anomalous observer being screened. The method therefore constitutes a *screening* stimulus generator, not a measurement instrument, and cannot classify deficiency type or grade severity.

Contribution. The contribution is not a new psychophysical model of color vision. It is (1) an openly specified, reproducible stimulus-generation pipeline whose assumptions and implementation are fully inspectable, and which removes the memorization and luminance-cue weaknesses of fixed-image online tests; and (2) an explicit account of the calibration limits that

bound any screen-based color screen. No clinical validation study was performed, and none is claimed.

Keywords: pseudoisochromatic plates; color vision deficiency; Ishihara test; color-confusion lines; sRGB; screen calibration; observer metamerism; color-vision screening; procedural generation

---

1. Introduction

Color vision deficiency (CVD) affects roughly 8% of men and 0.4% of women of European ancestry (Birch, 2012). The most familiar screening tool is the pseudoisochromatic plate (PIP): a field of colored dots in which a figure — classically a digit — is set apart from its background using colors that a person with typical color vision distinguishes but a person with a red-green deficiency does not. The Ishihara plates are the canonical example and have been in clinical use for a century (Birch, 2001).

The migration of this test to the web is now essentially complete: a search for "online color blind test" returns dozens of sites. The overwhelming majority share two properties. First, they present a fixed set of scanned or reproduced plate images — the same public-domain plates, served as static files. Second, they say little or nothing about the one factor that governs whether a screen-based color test means anything at all: the display. A scanned plate designed for printing under a standardized illuminant is not the same stimulus when it is re-encoded into sRGB and shown on an arbitrary, uncalibrated monitor or phone.

This paper takes the opposite stance on both points. It specifies a method that generates plates procedurally, in the browser, at the moment the test is taken — so that the stimulus is fresh, unmemorizable, and free of the designed luminance cues that leak into reused images — and it states, as a first-class part of the method, exactly where and why a screen-based screen reaches its limit.

The thesis is deliberately modest: **a screen can generate a legitimate red-green screening stimulus, but it cannot be a color-vision measurement instrument, and the honest design of an online test consists precisely in respecting that boundary.** This is a methods and limitations report. It contains no clinical validation study, makes no diagnostic claim, and should not be read as establishing the sensitivity or specificity of the described tool.

2. Background

2.1 The pseudoisochromatic principle

A pseudoisochromatic plate exploits the structure of color-confusion. In dichromacy and anomalous trichromacy, certain pairs of colors that are distinct to typical observers become confusable because they differ chiefly along the axis served by the missing or shifted photopigment. If a figure is composed of dots drawn from one side of such a confusion pair and the background from the other, and if all *other* cues — dot size, spacing, and crucially luminance — carry no information about the figure, then the figure is visible only to an observer whose color vision separates the two colors. Classical plate design achieves this not by exact equiluminance, which is fragile because individual luminance sensitivity varies, but by *randomizing* dot

luminance so that luminance is uninformative; the term "pseudoisochromatic" refers to this deliberate matching of everything except hue (Birch, 2001).

Two consequences follow, both established in the clinical literature. First, PIP tests are *screening* tools: the National Research Council's Committee on Vision concluded that "the number of errors on pseudoisochromatic tests tells us little about the type or extent of a color vision defect" (NRC Working Group 41, 1981). A plate test can indicate that a red-green deficiency is *likely present*; it cannot say which type or how severe. Second, the validity of a PIP test is illuminant-dependent by construction — the same report notes that the spectral quality of the light source affects readings, so printed plates are specified for viewing under a standardized illuminant. Any change to the light — or, in the screen case, to the display's spectral output — changes the stimulus.

2.2 Confusion lines and the color basis

The geometric basis for choosing figure/ground color pairs is the system of dichromatic confusion lines: sets of chromaticities that a given class of dichromat cannot distinguish, converging on a copunctal point characteristic of protan, deutan, or tritan deficiency (Wyszecki & Stiles, 1982, §5). Red-green deficiencies (protan and deutan) are common; tritan (blue-yellow) deficiency is rare and usually acquired. A plate intended to screen red-green deficiency therefore separates figure from ground along a red-green confusion direction; a tritan plate separates along the blue-yellow direction.

2.3 Simulating confusion on a display

There is a mature line of work on reproducing dichromatic color relationships on digital displays. Brettel, Viénot and Mollon (1997) gave the standard algorithm: stimuli are transformed into a cone-response (LMS) space and projected onto a reduced surface defined by a neutral axis and anchor wavelengths, yielding the appearance of a stimulus to a dichromat. Viénot, Brettel and Mollon (1999) adapted this into practical colormaps for checking display legibility for protanopes and deuteranopes. Machado, Oliveira and Fernandes (2009) generalized the approach into a single physiologically-based model spanning normal vision, anomalous trichromacy, and dichromacy through a severity parameter from 0.0 to 1.0. These models define, in principle, how to place figure and ground colors on opposite sides of a confusion axis on a given display. PPPG draws on the same principle; Section 4 is explicit about where it uses a principled derivation and where it uses an empirical approximation.

2.4 Why the display is the limit

Two facts bound any screen-based color test, and — importantly — the binding constraint is not gamut *volume*. Pseudoisochromatic figure/ground pairs are low-chroma colors that sit comfortably inside both the sRGB and typical print gamuts, so gamut size is a secondary concern; the sRGB and offset-print gamuts in any case interpenetrate rather than one containing the other. The real limits are chromatic *accuracy* and *resolution* on an uncharacterized device.

First, a consumer display is a three-primary emissive device. Its encoding — most commonly sRGB (IEC 61966-2-1:1999), though many current phones ship wider (P3) gamuts under operating-system color management — fixes the colorimetric meaning of a code value, but only a *calibrated* device actually reproduces it: the true primaries, white point, transfer function (gamma), peak luminance, and black level of an arbitrary screen are unknown to the generator, so the code-value-to-XYZ mapping is unknown. Compounding this, a confusion-axis separation

is a *small* chromatic step, and 8-bit-per-channel quantization limits the number of distinct, correctly-ordered chromatic steps available along a confusion direction, which can band or collapse after any gamma or white-point remapping.

Second, real displays vary and adapt: the chromaticity of a TFT-LCD's primaries and white point shifts measurably with viewing angle (Pardo et al., 2004), and phones add ambient-light sensors, adaptive white-point features, "vivid" saturation presets, and blue-light filters, each of which silently alters the stimulus. A study of an online Ishihara test found that color-deficient participants could pass plates they should have failed, a discrepancy the authors attributed to "the spectral quality of the light source, illumination of the plates or the impact of the screen resolution" (van Staden et al., 2018). The point is not that screens are useless — a controlled study of characterized panels concluded that (2004-era) TFT-LCDs are adequate for anomaly *screening* (Pardo et al., 2004) — but that their validity is conditional on factors an in-browser test cannot control or even observe.

3. Related online tests

Existing online color tests occupy a spectrum. At one end are fixed-image reproductions of the Ishihara plates: convenient, but subject to two avoidable weaknesses. Because the images are static and public, a returning or motivated user can memorize the figures; and because they are compressed raster images, small luminance differences introduced by scanning, re-encoding, or scaling can leak a residual figure cue that is not purely chromatic — a known hazard in the reproduction of pseudoisochromatic material (Dain, 2004). At the other end are commercially operated tests, some using proprietary cone-isolation stimuli that report per-cone scores; these are more sophisticated but are typically embedded in a funnel toward a product (for example, color-correcting glasses), which introduces an incentive orthogonal to measurement.

Procedural generation of plate-like stimuli is not itself novel. PPPG's contribution is the *combination*: procedural generation (removing memorization), luminance randomization (removing the luminance cue), explicit and openly-specified calibration-limit disclosure, and the absence of a product-sales funnel. It does not claim to be more *accurate* than its alternatives — accuracy on an uncalibrated display is bounded for all of them by the same physics — but it does claim to be more honest about that bound, and more robust against the two specific failure modes of reused images.

4. Method: the PPPG pipeline

PPPG produces a set of plates at run time on an HTML canvas. A single plate is generated by four algorithms — figure masking, dot packing, palette assignment, and lightness randomization — followed by a scoring rule over the completed run. All parameters below are the values used in the reference implementation, expressed relative to the internal render size `C` for scale-invariance (`C = 480` px; the canvas is downscaled by CSS to roughly 300 px so that on a high-density display each dot is rendered above the pixel grid). Rendered example plates — a standard plate, the lightness-only control, and a grayscale conversion in which the figure vanishes — can be viewed in the live implementation at opticquiz.com/color/ and reproduced from the open-source repository.

Table 1 — Geometry and packing parameters.

Parameter	Symbol	Value
Internal render size	`C`	480 px
Plate radius	`R`	`C/2 − 3` = 237 px
Dot radius range	`[R_min, R_max]`	`[0.0085C, 0.020C]` = [4.08, 9.60] px
Minimum center separation	—	`r_i + r_j + 1` px (a 1-px edge gap)
Spatial-hash cell size	`c`	`0.032C` = 15.36 px
Neighbor search	—	1-ring (3×3 cells)
Max placement attempts	—	70,000
Max dots per plate	—	5,200
Dot placement	—	uniform over disk: `θ ~ U(0, 2π)`, `ρ = R·√U(0,1)`
Figure font	—	Arial Black 900; size `0.60C` (1-digit) / `0.46C` (2-digit); stroke `0.028C`
Figure membership test	—	mask red channel < 128 ⇒ *figure*

Table 2 — Palettes (HSL ranges). Hue separates figure from ground along each axis; saturation and lightness are sampled per dot from the listed ranges. Lightness is *shared* between figure and ground within the red-green and tritan axes, so it carries no designed cue; the control axis deliberately differs in lightness so it reads by luminance alone.

Axis	Region	H (°)	S (%)	L (%)
red-green	figure (green)	96–145	40–60	50–68
red-green	ground (orange/tan)	15–44	46–68	50–68
tritan	figure (blue)	205–240	45–66	56–74
tritan	ground (pink/rose)	328–352	40–62	56–74
control	figure	20–38	10–24	36–48
control	ground	28–46	4–12	78–88

Algorithm 1 — Figure masking. The target figure — a digit for the standard test, or a simple shape (star, heart, square, circle, triangle) for a pre-reader variant — is drawn to an offscreen buffer as a solid binary mask `M`. A candidate dot at center `p` is later classified as *figure* if `M(p)` is set, else *ground*. Using a rendered mask rather than a font-glyph lookup keeps the method uniform across digits, shapes, and arbitrary paths, and makes the figure set trivially extensible.

Algorithm 2 — Spatial-hash dot packing. Dots are placed by randomized dart-throwing under a minimum-separation constraint, accelerated by a spatial hash. Dot radii are sampled from `[R_min, R_max]` so that dot size does not correlate with figure membership. Separation is imposed center-to-center as the sum of the two radii plus a 1-px gap; each candidate is tested only against dots in the surrounding 3×3 cells; and packing terminates at either 5,200 placed dots or 70,000 failed attempts.

```
grid ← {}                                          # cell-key → list of (x, y, r)
placed ← 0 ; attempts ← 0
while attempts < 70000 and placed < 5200:
    attempts ← attempts + 1
    θ ← U(0, 2π) ;  ρ ← R · √U(0,1)                # uniform over the disk
    x ← C/2 + ρ·cos θ ;  y ← C/2 + ρ·sin θ
    r ← U(R_min, R_max)
    if dot (x,y,r) is not fully inside the plate circle:  continue
    if not FITS(x, y, r):  continue
    color ← INFIGURE(mask, x, y) ? PICK(fg) : PICK(bg)     # Algorithm 3
    draw filled circle (x, y, r, color)
    grid[cell(x,y)].append((x, y, r)) ;  placed ← placed + 1

FITS(x, y, r):
    for each of the 9 cells in the 3×3 block around cell(x,y):
        for each stored (xi, yi, ri) in that cell:
            if (x−xi)² + (y−yi)² < (ri + r + 1)²:  return false
    return true
```

Honest packing note. Because `R_max = 0.020C` and `c = 0.032C`, two maximum-radius dots collide at a center distance up to `R_max + R_max + 1 ≈ 0.040C + 1`, which slightly exceeds one cell width; a candidate near a cell boundary can therefore, in rare cases, collide with a max-radius dot roughly 1.25 cells away that the 1-ring search does not visit. Guaranteed non-overlap is thus not strict. In practice the dots are small relative to the cell and any residual overlap is sub-pixel and visually negligible; the routine targets even coverage, not certified circle packing. This is disclosed rather than claimed.

Algorithm 3 — Palette assignment. Each dot is colored by its mask class. Figure and ground colors are drawn from palettes *constructed to differ predominantly along a chosen confusion direction* — red-green for the standard plates, blue-yellow for the tritan plates. Two instantiations are possible within the same procedure: (a) a *principled* variant that derives each figure/ground pair by projection in a cone-space model (Brettel–Viénot–Mollon, 1997; Machado et al., 2009) on the assumed display primaries, placing the pair on a computed confusion line; and (b) an *empirical* variant, used by the current implementation, in which the palettes are hand-constructed to lie nominally along the confusion direction. The empirical variant is a convenience of the current build, not a principled equivalent (see §5). Each dot's color is sampled independently from its region's ranges — `PICK(spec): h ~ U(spec.hMin, spec.hMax); s ~ U(spec.sMin, spec.sMax); l ~ U(spec.lMin, spec.lMax); return HSL(h, s%, l%)`. A single *control* plate is generated whose figure differs from ground in lightness alone, along no confusion axis; it is legible to all observers regardless of color vision and serves as an attention and comprehension check. (Note that while lightness is shared figure-to-ground within an axis, the saturation ranges differ slightly, e.g. red-green figure S 40–60% vs. ground 46–68%; this, and the fact that hue itself carries luminance, is why §5 declines to claim strict photometric equiluminance.)

Algorithm 4 — Shared-band lightness randomization. Within each plate, the lightness of every dot is sampled from a shared band `[L_lo, L_hi]` that spans figure and ground alike. Because figure and ground draw lightness from the same distribution, no *designed* luminance cue distinguishes them, and the figure carries no mean-luminance signal. (The claim is exactly this and no more: on an uncalibrated display, equal encoded lightness does not guarantee equal

photometric luminance, so a residual, chroma-dependent luminance difference cannot be excluded without measurement; the nominal analysis in §5 addresses this.) The band is resampled on every run, so no two generated plates are identical.

Scoring. Plates are presented one eye at a time, with the non-tested eye covered, because color vision can differ between eyes, particularly in acquired deficiency; per-eye presentation is standard clinical practice (Birch, 2001) and costs nothing to reproduce. A run comprises eight plates per eye in the standard configuration: five along the red-green axis, two along the tritan axis, and one lightness-only control. Scoring is a plain count of plates read correctly on each axis, reported as a coarse band per axis, with an explicit statement that a single misread plate, or an uncalibrated display, can move the band. The classification is a fixed decision over the red-green miss count (of 5) and tritan miss count (of 2), gated by the control plate: a missed control returns *inconclusive* (a display or attention problem, not a color finding); otherwise red-green misses of ≥4-with-both-tritan-missed, ≥4, =3, and ≥2 map to *broad loss*, *strong*, *moderate*, and *mild* respectively, a tritan miss with intact red-green maps to a *tritan pattern*, and zero relevant misses to *normal*. There is no weighting, no probability model, and no inference of deficiency type. Because the tritan axis carries only two plates, its band is explicitly labeled indicative-only; and because a five-plate red-green screen is short relative to a 24- or 38-plate clinical Ishihara, the report notes that plate count limits test-retest reliability. The report states plainly that the screen cannot separate protanopia from deuteranopia and is not a diagnosis.

5. Limitations

The limitations are not incidental; they are the second half of the contribution, and they are stated without hedging.

The display bounds everything, and calibration — not gamut — is why. As established in §2.4, the confusion-axis colors lie inside the sRGB gamut, so the binding constraints are the *unknown* calibration of an arbitrary device (primaries, white point, transfer function, luminance, black level) and 8-bit quantization of a small chromatic step. On an uncalibrated screen the true chromaticity of every dot is unknown to the generator. Uncalibrated gamma changes the effective luminance band; ambient light and glare raise the black level and desaturate; and display color modes — night shift, blue-light filters, "vivid" presets, adaptive white point — alter or destroy the intended relationships. Tritan (blue-yellow) plates are especially vulnerable, because their S-cone-directed separation is carried predominantly by the blue channel that blue-light filters and night modes attenuate directly; red-green pairs are also perturbed, but comparatively less. The method cannot detect any of these conditions.

Observer metamerism — the deepest reason a screen is not a plate. A three-primary emissive display can, at best, produce a *metameric* match to an intended color — a match of tristimulus values, not of spectral power (Wyszecki & Stiles, 1982, §3). Two consequences follow. First, the printed-plate stimulus and its screen rendering are, in general, metamers rather than identical stimuli. Second, and more fundamentally, a metameric match constructed for the CIE standard observer is *not guaranteed to be a match for an observer whose cone fundamentals are shifted* — that is, for precisely the anomalous observer the plate is meant to screen. The screen's stimulus can therefore fail to reproduce the intended confusion for the very population under test. This is a structural limit of emissive displays, not a calibration artifact.

The empirical-palette gap. As stated in §4 (Algorithm 3), the current build constructs its confusion palettes empirically rather than deriving each pair per-display from a cone-space model. Unknown display primaries degrade the principled and empirical approaches *equally* and so do not favor the empirical choice; the principled variant still strictly dominates, because under assumed nominal primaries it places the pair on a computed confusion line while the empirical one does not. The empirical separation is therefore *nominal, not measured*. This nominal alignment is nonetheless checkable without any human study. Converting the palette range-midpoints (Table 2) from HSL to CIE 1931 *xy* under assumed nominal sRGB primaries, and measuring the angular deviation of each figure-to-ground separation vector from the relevant confusion-line direction (the line from the type's copunctal point through the figure color), gives a first-order estimate (copunctal points are the standard CIE 1931 confusion points; Wyszecki & Stiles, 1982; the estimate uses the palette range-midpoints): the red-green palette sits close to both the protan and deutan confusion lines — deviations of approximately 6° and 5° respectively — which *supports the confusion-axis claim for the primary red-green screen*. The current tritan palette, however, deviates from the tritan confusion line by approximately 53°, indicating that the blue-yellow plates as presently constructed are a weak, off-axis tritan screen. This corroborates the report's existing caution that any tritan finding must be confirmed clinically, and it identifies a concrete improvement: re-deriving the tritan palette from a cone-space model (the principled variant of Algorithm 3) so that it lies on the tritan confusion line. A fuller analysis that samples the palette ranges rather than their midpoints, together with a rendered chromaticity-diagram figure, is identified as the next analytical step; the estimate above is reported here rather than withheld because it materially qualifies the tritan claim.

Screening, not measurement, not diagnosis. As established for PIP tests generally (NRC Working Group 41, 1981), a plate count indicates the likely *presence* of a red-green deficiency and nothing more. It does not classify type (protan vs. deutan) — that requires an anomaloscope, which for the red-green axis is the recognized gold standard and "the only clinical method capable of classifying the color defects by their presumed genetic entities" (NRC Working Group 41, 1981); the analogous instrument for the blue-yellow axis is the Moreland match. Nor does a plate count grade severity, which is the province of instruments such as the CAD test (Barbur et al., 2009). A separate caution on sensitivity is warranted: standard plates under-detect deficiency relative to *genetic* testing (sensitivity ≈ 0.50; Arnegard et al., 2022), but this figure reflects mild genotypes whose phenotype a functional plate is, by design, expected to pass, and must not be read as "plates miss half of color-deficient people" relative to the anomaloscope, against which plate sensitivity is far higher.

No validation was performed. This is the most important limitation. No study of the described tool's sensitivity, specificity, or agreement with a reference test (anomaloscope, CAD, or genotyping) has been conducted. PPPG is presented as a stimulus-generation technique and a limitations analysis, not as a validated screening instrument. Any such claim would require a controlled study against a reference standard on characterized displays, which is identified here as future work and explicitly not asserted.

6. Reproducibility and availability

PPPG is implemented in client-side JavaScript with no server component; all computation, including scoring, occurs in the browser and no response data is transmitted or stored. The full source is open and versioned in the project repository (github.com/zengineco/opticquiz.com),

and an archived snapshot is deposited at https://doi.org/10.5281/zenodo.21311369. Because the method is procedural, a reader can reproduce the stimulus family by running the code; because the code is open, the packing parameters, palette construction, and luminance bands are all inspectable and modifiable — and are given in full in Tables 1–2 and the Section 4 pseudocode, so the specification here is independent of any single implementation. A fuller range-sampled confusion-line analysis and a chromaticity-diagram figure (Section 5) are identified as the next analytical step.

7. Discussion

PPPG occupies a specific and defensible niche. It is a better *screen* than a fixed-image online test — unmemorizable, free of the designed luminance cue, fresh on every run — and it is more transparent than a commercial test about the ceiling that the display imposes on all of them. It is not, and does not pretend to be, a substitute for clinical assessment.

There is a broader observation worth stating. Much of the online color-testing landscape is characterized by overclaiming: fixed plates presented as if the display were a calibrated instrument, scores reported with unjustified precision, and screening framed as diagnosis. The corrective is not a better secret algorithm; it is disclosure. A test that generates its stimuli openly, randomizes away the non-chromatic cues, and states its display-bound limits in plain language gives the user something the polished alternatives often do not: an accurate sense of how much to trust the result. The value proposition of an honest screen is not maximal sensitivity — that is unattainable on uncalibrated hardware — but calibrated trust.

Future work follows directly from the limitations. The nominal confusion-line deviation analysis (§5) is the immediate next step and requires no participants. A principled reimplementation would derive figure/ground pairs from the Brettel–Viénot–Mollon or Machado models on measured display primaries; a display-characterization step (for example, a card-based or reference-swatch calibration) could narrow the uncertainty band; and a validation study against a reference standard on characterized displays would be required before any sensitivity claim could be made. None of these is claimed as complete here.

8. Conclusion

A pseudoisochromatic screening stimulus can be generated procedurally in the browser, removing the memorization and luminance-cue weaknesses of reused plate images, and doing so is straightforward and fully reproducible. What such a stimulus cannot do is escape the display: on uncalibrated, three-primary, variably-lit consumer screens — and, more deeply, under observer metamerism — a color test is a screen and not a measurement, able to suggest the presence of a red-green deficiency but unable to classify or grade it. The contribution of this paper is to provide the method and, with equal weight, the honest boundary around it. An online color test earns trust not by claiming the precision of a clinic, but by stating clearly the one thing most of its peers omit: what a screen cannot measure.

Acknowledgments

The author thanks the maintainers of the open color-science implementations of the Brettel–Viénot–Mollon and Machado models, whose public code clarified the confusion-axis literature.

Data and code availability

No human-subjects data were collected. The generator source is openly available in the project repository (github.com/zengineco/opticquiz.com) and archived at https://doi.org/10.5281/zenodo.21311369. This manuscript is released under CC BY 4.0.

Conflict of interest

The described method is implemented on OpticQuiz (opticquiz.com), a free website the author operates, which is supported by advertising and by a small number of affiliate links on its written guides (never on the tests themselves). This is a commercial incentive and is disclosed as such; it is not, however, a product-sales funnel — no glasses, subscription, diagnosis, or premium product is sold through the color tests, and the method's design objective is an honest screen rather than a conversion. No funding was received for this work.

References

Arnegard, S., Baraas, R. C., Neitz, M., Hagen, L. A., & Neitz, J. (2022). The limitation of standard pseudo-isochromatic plates in identifying colour vision deficiencies when compared with genetic testing. *Acta Ophthalmologica*. PMC9614865.

Barbur, J. L., Rodriguez-Carmona, M., Evans, S., & Milburn, N. (2009). *Minimum Color Vision Requirements for Professional Flight Crew* (Report DOT/FAA/AM-09/11). Federal Aviation Administration.

Birch, J. (2001). *Diagnosis of Defective Colour Vision* (2nd ed.). Butterworth-Heinemann. ISBN 0-7506-4174-6.

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

Brettel, H., Viénot, F., & Mollon, J. D. (1997). Computerized simulation of color appearance for dichromats. *Journal of the Optical Society of America A, 14*(10), 2647–2655. https://doi.org/10.1364/JOSAA.14.002647

Dain, S. J. (2004). Clinical colour vision tests. *Clinical and Experimental Optometry, 87*(4–5), 276–293. https://doi.org/10.1111/j.1444-0938.2004.tb05057.x

Farnsworth, D. (1943). The Farnsworth-Munsell 100-Hue and Dichotomous Tests for Color Vision. *Journal of the Optical Society of America, 33*(10), 568–578. https://doi.org/10.1364/JOSA.33.000568

International Electrotechnical Commission. (1999). *IEC 61966-2-1:1999 — Multimedia systems and equipment — Colour measurement and management — Part 2-1: Colour management — Default RGB colour space — sRGB*.

Machado, G. M., Oliveira, M. M., & Fernandes, L. A. F. (2009). A physiologically-based model for simulation of color vision deficiency. *IEEE Transactions on Visualization and Computer Graphics, 15*(6), 1291–1298. https://doi.org/10.1109/TVCG.2009.113

National Research Council, Committee on Vision. (1981). *Procedures for Testing Color Vision: Report of Working Group 41*. National Academies Press. https://www.ncbi.nlm.nih.gov/books/NBK217823/

Pardo, P. J., Pérez, A. L., & Suero, M. I. (2004). Validity of TFT-LCD displays for colour vision deficiency research and diagnosis. *Displays, 25*(4), 159–163. https://doi.org/10.1016/j.displa.2004.09.006

van Staden, D., Mahomed, F. N., Govender, S., Lengisi, L., Singh, B., & Aboobaker, O. (2018). Comparing the validity of an online Ishihara colour vision test to the traditional Ishihara handbook in a South African university population. *African Vision and Eye Health, 77*(1), a370. https://avehjournal.org/index.php/aveh/article/view/370

Viénot, F., Brettel, H., & Mollon, J. D. (1999). Digital video colourmaps for checking the legibility of displays by dichromats. *Color Research & Application, 24*(4), 243–252. https://doi.org/10.1002/(SICI)1520-6378(199908)24:4<243::AID-COL5>3.0.CO;2-3

Wyszecki, G., & Stiles, W. S. (1982). *Color Science: Concepts and Methods, Quantitative Data and Formulae* (2nd ed.). Wiley.