DOT/FAA/TC-23/56

Federal Aviation Administration
William J. Hughes Technical Center
Atlantic City International Airport
New Jersey 08405

Adapting the FAA-HF-STD-010A Standard Color Palette to Daytime Illumination

David L. Post[1]
Nicole Racine[2]
Eve Perchanok[2]
Randy Sollenberger[3]

[1] ϕ
Dayton OH

[2] DSoft Technology
Colorado Springs CO

[3] Federal Aviation Administration
William J. Hughes Technical Center
Human-Systems Integration Branch, ANG-E5B
Atlantic City International Airport NJ

February 2024

Final Report

U.S. Department of Transportation
Federal Aviation Administration

NOTICE

This document is disseminated under the sponsorship of the U.S. Department of Transportation in the interest of information exchange. The U.S. Government assumes no liability for the contents or use thereof. The U.S. Government does not endorse products or manufacturers. Trade or manufacturers' names appear herein solely because they are considered essential to the objective of this report. The views expressed herein are those of the authors and do not reflect views of the United States (U.S.) Department of Transportation (DOT) or FAA. This document does not constitute FAA policy. Consult the FAA sponsoring organization listed on the Technical Documentation page as to its use.

This report is available at the Federal Aviation Administration William J. Hughes Technical Center's Full-Text Technical Reports page: https://www.faa.gov/data_research/library in Adobe Acrobat portable document format (PDF).

Technical Report Documentation Page

Form DOT F 1700.7 (8-72) Reproduction of completed page authorized

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
DOT/FAA/TC-23/56		

4. Title and Subtitle	5. Report Date
Title of Report: Adapting the FAA-HF-STD-010A Standard Color Palette to Daytime Illumination	February 2024
	6. Performing Organization Code ANG-E5B
7. Author(s) David L. Post, ϕ; Nicole Racine, DSoft Technology; Eve Perchanok, DSoft Technology; Randy Sollenberger, FAA	8. Performing Organization Report No.
9. Performing Organization Name and Address Federal Aviation Administration William J. Hughes Technical Center Human-Systems Integration Branch, ANG-E5B Atlantic City International Airport, NJ 08405	10. Work Unit No. (TRAIS) 11. Contract or Grant No.
12. Sponsoring Agency Name and Address ATO-PMO Analysis & Planning Team, AJM-1310	13. Type of Report and Period Covered Technical Report 14. Sponsoring Agency Code

15. Supplementary Notes

16. Abstract

This project demonstrated that the color palette specified in FAA-HF-STD-010A (2020) can be used unaltered in air traffic control (ATC) towers during bright daytime conditions. The FAA palette introduced standard colors for coding ATC displays. It was developed to accommodate controllers who have mild-to-moderate color-vision deficiencies without compromising usability for color normals. The test conditions were dark, which represents the majority of ATC environments, but that choice left the palette's suitability for use in towers during daytime unknown. We have shown that, under worst-case ambient illumination, controllers with normal and deficient color-vision can be expected to perform almost perfectly with the new Standard Terminal Automation Replacement System (STARS) Tower Display Monitors if they are allowed to adjust the display's luminance. We provide an equation that predicts the peak luminance requirement for any candidate display that uses the FAA standard palette.

17. Key Words	18. Distribution Statement
Air Traffic Controller, Color Vision Deficiency, Standard Color Palette, Visual Search Task	This document is available to the U.S. public through the National Technical Information Service (NTIS), Springfield, Virginia 22161. This document is also available from the Federal Aviation Administration William J. Hughes Technical Center at https://www.faa.gov/data_research/library

19. Security Classification. (of this report)	20. Security Classification (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified		

ACKNOWLEDGEMENTS

This research project was performed by DSoft Technology, Engineering & Analysis, Inc. under Federal Aviation Administration (FAA) contract 692M15-18-D-0000. It was completed with funding from the FAA NextGen Human Factors Division (ANG-C1) in support of the FAA Air Traffic Organization (ATO). There are many people we owe thanks to. Our program managers (Dan Herschler and Karl Kaufmann) and sponsor (Ben Willems) made valuable contributions to the experimental plan and final report. Carla Hackworth (AAM-500), Janine King, and Suzanne Thomas (CN3-S) loaned the ATCOV workstation we needed to screen participants. The ANG-E14 team provided tower simulation laboratory support: Albert Macias was in charge of laboratory management coordination, Laura Hamann assisted with calibrating the displays, Alex Perez made the modifications we needed to our data collection software, and John Dilks installed equipment and provided general laboratory support. Zak Renzi (ANG-E151) loaned the STARS Tower Display Monitors that were essential to the project. Jennifer Librizzi (ANG-E5B) provided administrative project support and Jon Rein (ANG-E5B) helped us, who were new to R (which is an open-source statistical programming language) learn enough of it to produce all of our figures, summary statistics, and hypothesis tests. Jay Neitz (University of Washington) kindly donated a set of his color-vision testing sheets.

Table of Contents

Figures

Tables

Acronyms and Abbreviations

Acronym/Abbreviation	Definition
ANOVA	Analysis of Variance
ATC	Air Traffic Control
ATCOV	Air Traffic Color Vision Test
ATOP	Advanced Technologies for Oceanic Procedures
CIE	Commission Internationale de l'Eclairage
CVD	Color Vision Deficient
D_{65}	CIE standard illuminant D_{65}
ERAM	En-Route Automation Modernization
FAA	Federal Aviation Administration
FLSD	Fisher's Least Significant Difference
GES	Generalized eta^2
LED	Light-Emitting Diode
NCV	Normal Color Vision
PC	Personal Computer
RGB	Red, green, and blue
SD	Standard deviation
STARS	Standard Terminal Automation Replacement System
TDM	Tower Display Monitor
WJHTC	William J. Hughes Technical Center
Wx	Weather

Executive summary

This project demonstrated that the color palette specified in FAA-HF-STD-010A (2020) can be used unaltered in air traffic control (ATC) towers during bright daytime conditions. The FAA palette introduced standard colors for coding ATC displays. It was developed to accommodate controllers who have mild-to-moderate color-vision deficiencies without compromising usability for color normals. The test conditions were dark, which represents the majority of ATC environments, but that choice left the palette's suitability for use in towers during daytime unknown. We have shown that, under worst-case ambient illumination, controllers with normal and deficient color-vision can be expected to perform almost perfectly with the new Standard Terminal Automation Replacement System (STARS) Tower Display Monitors if they are allowed to adjust the display's luminance. We provide an equation that predicts the peak luminance requirement for any candidate display that uses the FAA standard palette.

1 Introduction

This project is a follow-on to Gildea et al. (2018, 2020), which are the research projects that produced the Federal Aviation Administration (FAA) color palette contained in FAA-HF-STD-010A (2020). Gildea et al.'s (2018) project was motivated by the FAA Air Traffic Organization Program Management Organization's research requirement to address the need to accommodate air traffic controllers who have mild-to-moderate color-vision deficiencies (CVDs). The objective is to make it easier, and allow more of them, to perform their duties using the latest air traffic control (ATC) color displays without sacrificing usability for color normals. The color palettes used on contemporary ATC systems were not designed for color-vision deficient (CVD) controllers. Gildea et al. (2018) developed a palette that reasonably accommodates CVD personnel. More specifically, they showed that the colors are discriminable, recognizable, conspicuous (i.e., easy to locate), and legible for normal color-vision (NCV) and many CVD viewers. Gildea et al. (2020) refined Red slightly and the change was incorporated into the FAA-HF-STD-010A (2020) palette, which is shown in Tables 1 and 2, below.

Table 1. FAA standard color palette: Foreground colors

	Color name	**u'**	**v'**	**%Y**	**sR**	**sG**	**sB**	**Hex**
	White	0.1978	0.4683	100	255	255	255	FFFFFF
	Pink	0.266	0.418	41	246	132	216	F684D8
	Gray	0.1978	0.4683	45	179	179	179	B3B3B3
	Blue	0.17	0.348	28	94	141	246	5E8DF6
	Orange	0.294	0.541	42	254	147	13	FE930D
	Red	0.44	0.518	21.8	255	19	32	FF1320
	Green	0.13	0.54	55	35	225	98	23E162
	Yellow	0.193	0.55	80	223	243	52	DFF334
	Magenta	0.276	0.304	23	216	34	255	D822FF
	Aqua	0.142	0.428	50	7	205	237	07CDED
	Brown	0.241	0.519	34	197	149	91	C5955B

Table 2. FAA standard color palette: Weather (Wx) colors

	Color name	Severity	u'	v'	%Y	sR	sG	sB	Hex
	Black	0	---	---	0.0	0	0	0	000000
	Wx-Green	1 & 2	0.15	0.5	3.2	23	57	40	173928
	Wx-Yellow	3 & 4	0.23	0.54	7.1	90	74	20	5A4A14
	Wx-Red	5 & 6	0.26	0.4	5.0	93	46	89	5D2E59

The FAA developed the FAA-HF-STD-010A (2020) palette for the relatively dim ambient lighting conditions common to FAA indoor and nighttime tower viewing conditions. The objective for this new study was to determine whether it is possible to adapt the palette to yield the same usability for NCV and CVD controllers under daytime ambient illumination levels and, if so, how. Increasing the display's luminance while keeping all other aspects of the palette (i.e., chromaticity coordinates and luminance ratios) unchanged is a simple and attractive solution because Gildea et al. (2018) established those parameters using data from 103 NCV and 52 CVD participants, including a tritan (i.e., the rare, blue-weak type of deficiency).

This remedy can be shown to work mathematically in that, although the original luminance contrasts and chromaticity coordinates cannot be restored exactly, one can come arbitrarily close by increasing the display luminance sufficiently. Our first objective, therefore, was to determine the display luminance needed to maintain acceptable controller performance for a representative illuminance and display screen reflectance. Wilson, et al. (2007, p. 13) measured the illumination striking display screens in their survey of ATC towers under operational daytime conditions. They found that worst-case was approximately 55,000 lux (5130 fc), so we adopted 55,000 lux as the illuminance to use on our display screens.

Gildea et al. (2018) used four search tasks to evaluate the palette. The Weather-Color search required identifying each of the three luminous (i.e., non-Black) weather colors the palette provides with all four weather colors as the background. The Foreground-Color search required identifying all 11 foreground colors in the palette against all four weather colors. The Shape search required identifying short character strings that appeared in all 11 foreground colors against all four weather colors. The Redundant search coded targets using color and shape so they were identifiable either way. FAA-HF-STD-010A (2020, Section 4.2) requires coding critical information redundantly to enhance safety, so the Redundant search was the most realistic test of the palette.

Figure 1 shows the main results that Gildea et al. (2018) obtained. All the differences among the color vision types are significant for the Foreground-Color search, $p < 0.05$; further, the protan and deutan Scores are reliably lower for the Foreground-Color search than for the other three search types, $p < 0.05$. The CVDs were clearly disadvantaged for the pure color search, as one would expect, but their mean Scores of 81.2 and 85.5 show that their ability to identify the colors was impressive. By contrast, none of the differences among the other search types are significant, implying that the palette is suitable as a replacement for existing FAA color sets that accommodates color-deficient controllers better.

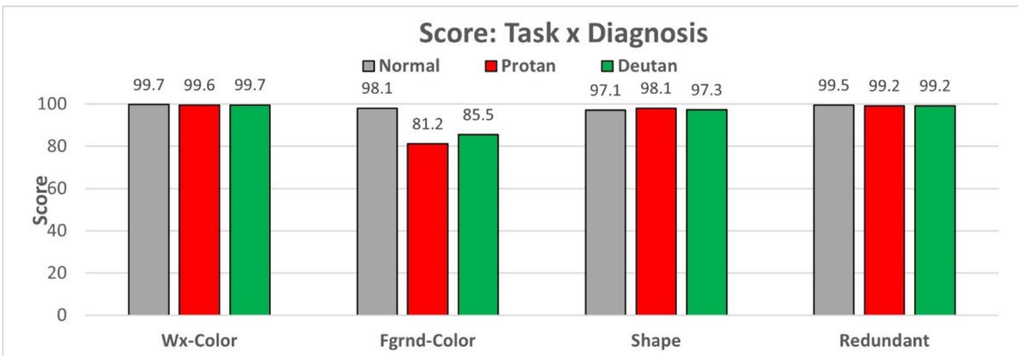

Figure 1. Main results from Gildea et al. (2018)

2 Method

2.1 Participants

The participants were William J. Hughes Technical Center (WJHTC) federal and contract employees, supplemented by people recruited from the local population. None had experience as controllers. We paid all non-federal participants an hourly rate. The only requirement for participation was 20/40 or better near visual acuity, corrected or uncorrected, in accordance with the FAA Guide for Aviation Examiners (2022; Item 51.a). All participants were under age 55, which is the age at which controllers transition to tasks that do not require directing traffic.

The pilot study involved 11 NCV individuals who were federal employees from the WJHTC. The main experiment involved a total of 39 individuals (19 CVDs and 20 NCVs; we ran 3 extra for reasons that are explained in Section 3.5). We used Faul et al.'s (2007) *G*Power* program to estimate the number of participants needed, based on the pilot study, and to monitor power as the main experiment progressed to ensure the number of participants was adequate.

2.2 Apparatus

2.2.1 Displays

The stimuli were presented on the latest Standard Terminal Automation Replacement System (STARS) Tower Display Monitors (TDMs), connected to personal computers (PCs) running Windows (see Figure 2). The TDMs were EIZO Raptor model RP-2124-02 liquid crystal displays that are intended for environments with high ambient illumination. They have 1600 x1200-pixel resolution (4:3 aspect ratio) on a 21.3-inch diagonal screen, 24-bit color, 600 cd/m^2 default maximum luminance with the backlight limiter turned on, and a 60-Hz refresh rate. We used two displays so we could test two participants simultaneously, thereby reducing data-collection time. We placed the displays on a workstation table in the Tower laboratory in the Research Development and Human Factors Laboratory (Technical Center Building 28). We used foam core with a matt finish to reduce the potential for one participant to affect the other and to stabilize their state of luminance adaptation.

Figure 2. TDM configuration with foam core blinders

We measured the TDMs using a calibrated Photo Research model 740 spectroradiometer. We began by determining the best TDM settings for our purposes. It became evident that the backlight controller should be turned off and Contrast, which ranges from 0 to 100, should be 50. The backlight controller limits the display's peak luminance to roughly 600 cd/m^2 to extend the life of the backlight's light-emitting diodes (LEDs); turning it off boosts the peak luminance to approximately 1000 cd/m^2.

Our next step was to identify the best correlated color-temperature setting to use from the TDM's three presets: 5600, 7300, and 9300 K. The FAA standard palette's White has the chromaticity coordinates of CIE standard illuminant D_{65} (i.e., u′ = 0.1978 and v′ = 0.4683) and we wanted to maximize its luminance. With the Backlight-Brightness set to 50, the 7300-K setting yielded peak luminances of 171, 736, and 93 cd/m^2 for the red, green, and blue (RGB) channels, respectively, all of which are greater than the corresponding peak luminances produced by the other two presets. That outcome showed that the 7300-K setting would produce the most luminous D_{65} white.

Next, we determined the Backlight-Brightness setting that would produce that maximum D_{65} white on the two displays. We found that setting it to 80 gave us a bit more than 800 cd/m^2 on both TDMs with a u′v′-chromaticity error ≤ 0.0004. Higher luminances would require more luminance than one of the channels can produce, meaning the peak white's chromaticity would drift away from D_{65}.

We used custom software to determine the RGB values to use for the 15 FAA standard palette colors on the two displays. That software controlled the TDMs and recorded the spectroradiometer's measurements. It incorporates Post and Calhoun's (1989, 2000) measure-and-adjust algorithm to ensure that all colors' luminances are accurate within ± 2.5% and the chromaticity error is ≤ 0.0025 on the CIE 1976 u′v′-chromaticity diagram.

Finally, we iterated adjusting each TDM's Backlight-Brightness setting and measuring the resulting peak D_{65} luminance to produce Table 3. It shows the settings needed to produce TDM peak D_{65} luminances ranging from 450 to 700 cd/m^2. The other colors' luminances scaled proportionally, so their luminance ratios remained constant. Spot checks showed that the colors' luminance and chromaticity accuracy was stable over time.

Table 3. Lookup table relating TDM Backlight-Brightness setting to peak D_{65} luminance

	TDM 1					TDM 2			
Target Luminance cd/m^2	Backlight Brightness Setting	Measured Luminance cd/m^2	u′	v′		Backlight Brightness Setting	Measured Luminance cd/m^2	u′	v′
700	44	701.5	0.2033	0.4695		47	702.7	0.1986	0.4685
650	41	647.2	0.2040	0.4698		44	649.9	0.1990	0.4685
600	38	599.4	0.2043	0.4702		41	598	0.1993	0.4686
550	35	550.2	0.2053	0.4706		38	551.5	0.1997	0.4688
500	32	503.5	0.2063	0.4710		34	491.4	0.2006	0.4689
450	29	458.6	0.2072	0.4713		31	447.9	0.2013	0.4689

2.2.2 Illuminants

Six GVM-SD300D LED spotlights illuminated the displays (see Figure 3). We attached the spotlights to a ceiling track and adjusted them to illuminate each TDM with 55,000 lux (see Figure 4). We set them to a correlated color temperature of 5600 K, which is representative of daylight in the early morning and late afternoon, when the sun is low in the sky and, hence, ambient illumination in a tower is greatest. We allocated three spotlights to each TDM, which allowed us to improve the illumination's spatial uniformity. We measured the illuminances using a calibrated Photo Research model 520 illumination meter.

Once we the positioned the spotlights, we measured the luminances they produced on the TDM screens with the PR-740 spectroradiometer positioned level with each screen's center. We measured three TDMs because we had a spare and wanted the largest sample size possible. The measurements were within 5% of each other and averaged 60 cd/m^2. If we make the simplifying assumption that the screen reflectances are Lambertian, that measure implies that their reflectance for the spotlights is (60 cd/m^2 * π / 55,000 lm/m^2 =) 0.00343.

Figure 3. GVM-SD300D LED spotlight

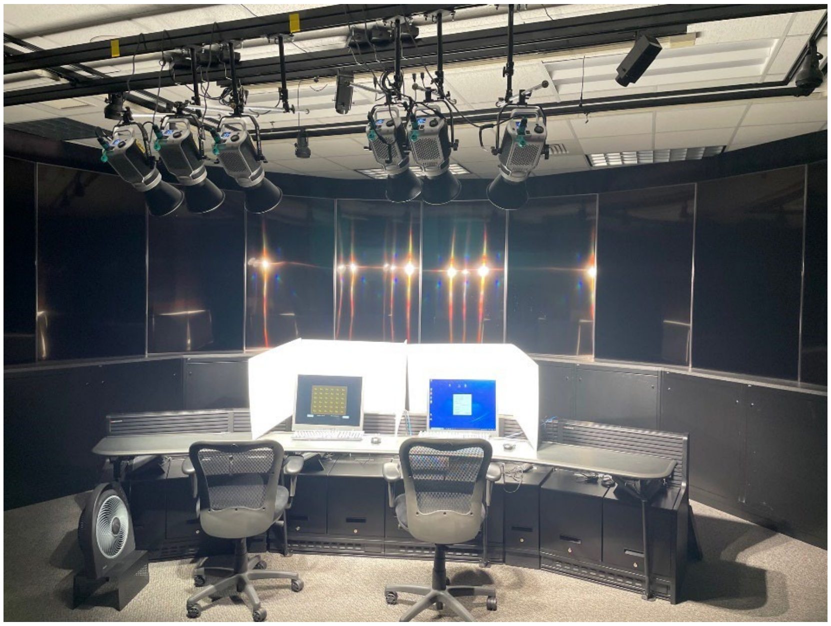

Figure 4. Overhead LED spotlight configuration

We used a Good-Lite True Daylight Illuminator (see Figure 5) to provide appropriate lighting for the color vision tests. The Good-Lite uses LEDs to produce an approximation to CIE standard illuminant D_{65} that is suitable for illuminating reflective vision-testing materials.

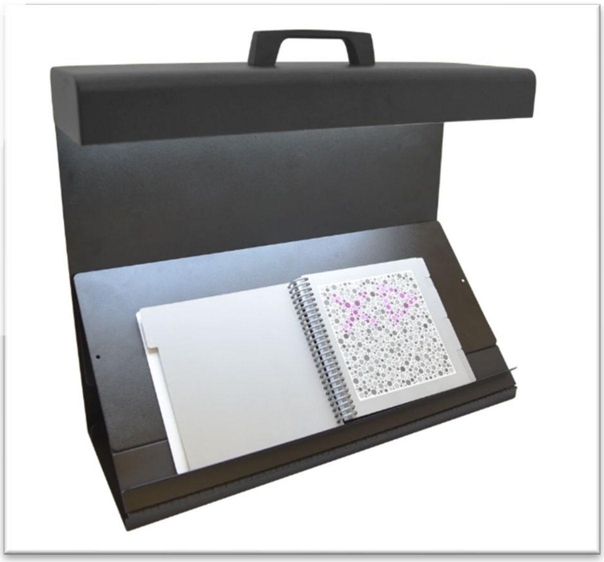

Figure 5. Good-Lite True Daylight Illuminator

2.2.3 Visual acuity test

To evaluate near visual acuity, we used the Smart Optometry Eye Test for Professionals version 4.6 (Smart Optometry, 2022), administered on an iPhone (see Figure 6). We used the Good-Lite Illuminator to provide the recommended illumination and as a stand for the iPhone with a 40-cm viewing distance. The test requires the examinee to swipe the screen in the direction of an optotype opening. The optotype symbols become smaller with each successfully completed row until the examinee commits errors. All participants met our 20/40 near visual-acuity requirement.

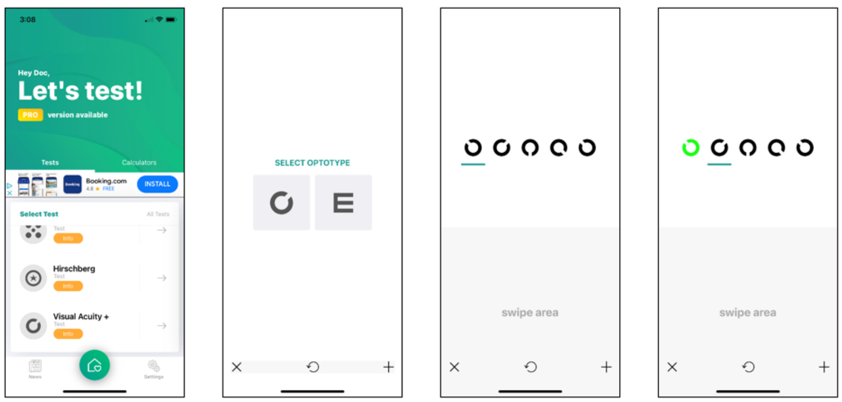

Figure 6. Smart Optometry Eye Test

2.2.4 Neitz test of color vision

We used the Neitz Test of Color Vision (NTCV; Neitz et al., 2023) to screen for color vision deficiencies. The NTCV is an accurate, easy-to-administer pseudoisochromatic-plate test that identifies the probable type and severity of color vision deficiencies, including tritan deficiencies. The participants examined nine colored shapes within patterns of grey dots on a piece of paper (see Figure 7) that rested on the Good-Lite Illuminator's easel and read as many as they could.

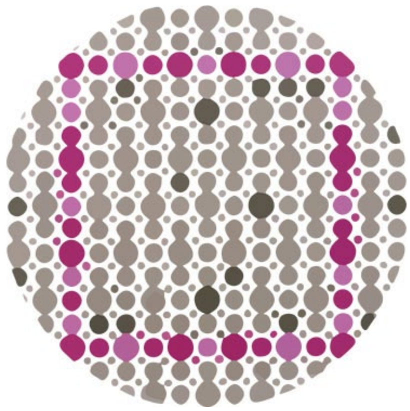

Figure 7. Example of a Neitz test plate

We administered the NTCV to all participants. The 18 NCV participants in the main experiment averaged 8.2 (min/max 7-9) correct responses; the 18 CVD participants averaged 3.5 (min/max 3-5). Most of the CVD participants made errors that indicated more than one type of deficiency, which is not unusual, so we combined them into one group for purposes of analysis.

2.2.5 Ishihara pseudoisochromatic plates

The Ishihara 24-plate set is a commonly used test for CVDs and is approved for screening air traffic controllers in the FAA *Guide for Aviation Medical Examiners* (2022; Item 52). Figure 8 shows an example of the Ishihara test plates. We administered the Ishihara test to those participants who did not score a perfect 9 on the Neitz test, using the Good-Lite Illuminator. Plates 1-15 determine whether the examinee's color vision is normal or defective; 18-24 are for examinees who are unable to read numbers. We used the set of 15, each of which contains an array of colored dots forming a number within a circle. If 13 or more plates are read correctly, the examinee's color vision is regarded as normal. If nine or fewer plates are read correctly, the examinee's color vision is considered deficient. Values in between indicate anomalous trichromacy (e.g., deutan rather than deuteranope, etc.).

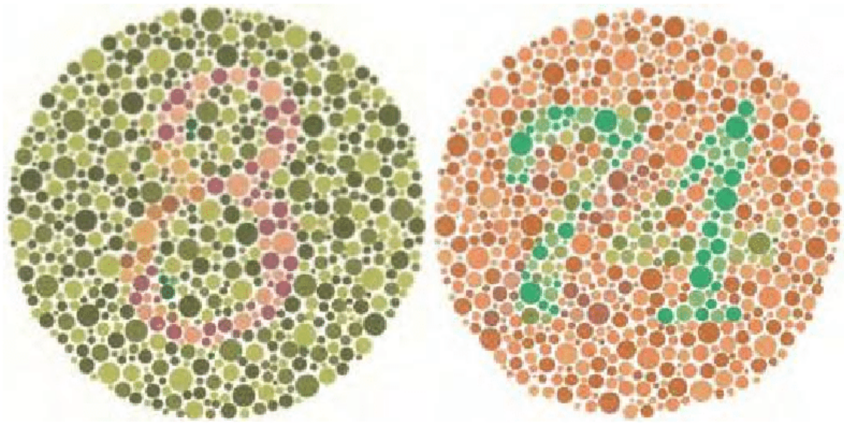

Figure 8. Examples of Ishihara test plates

We administered the Ishihara test to all 18 CVD participants. Most of them (12) scored correctly only on the demonstration plate, which is designed to be visible to all people. Five other CVDs scored two or three correctly. The remaining CVD scored 11 correct responses, which still falls in the color deficient category.

2.2.6 ATCOV

The Air Traffic Color Vision Test (ATCOV; Chidester et al., 2011; Chidester et al., 2013) is an occupational color-vision test, used to determine whether ATC candidates who have a color-vision deficiency can nonetheless perform adequately with the color sets used by the En Route Automation Modernization (ERAM) system, STARS, and the Advanced Technologies for Oceanic Procedures (ATOP) system. The ATCOV consists of four subtests that evaluate an individual's ability to discriminate among critical data elements in each of the Air Traffic Control environments. CVD candidates must pass the ATCOV for at least one ATC system to become controllers. Each subtest is pass/fail. We administered the test to our CVD participants using an ATCOV test station and calibrator that was provided by the FAA Civil Aerospace Medical Institute's Human Factors Research Division (AAM-500). We followed the instruction manual's procedure, including use of the prescribed illuminant.

We administered the ATCOV to 17 CVD participants (one declined). If a participant failed a subtest, they had the option to take it a second time, as prescribed by the manual. Table 4 presents a summary of the subtest results. We believe, therefore, that our group of CVDs represented the range of deficiency that would be expected in the controller population.

Table 4. ATCOV subtest results

	Subtest 1	*Subtest 2*	*Subtest 3*	*Subtest 4*
Pass	12	9	15	17
Fail	5	8	2	0

3 Procedures

3.1 Pre-experiment

On arrival, we provided participants with the Informed Consent Statement shown in Appendix A. It described the study, foreseeable risks, their rights and responsibilities, informed them that their participation was voluntary, and that they could withdraw at any time without penalty. We protected the information they provided, including Personally Identifiable Information. The WJHTC Institutional Review Board determined that this study met the criteria for exempt status and approved it to proceed. We obtained the participants' verbal consent, indicating they understood their rights and wished to participate. Then, they read a briefing that described the study procedures.

3.2 Dependent measures

Our pilot study and main experiment evaluated NCV and CVD viewers when performing the Gildea et al. (2018) visual search tasks under 55,000 lux of ambient illumination with their TDM's peak D_{65} luminance set to one of six possible levels: 450, 500, 550, 600, 650, and 700 cd/m^2. For each trial, participants responded to stimuli using the PC's mouse and the software recorded two dependent measures:

(1) An accuracy score ("Score"): The percentage of correct target identifications minus the percentage of incorrect identifications (i.e., false alarms); it ranges therefore from -100 to 100.

(2) Response speed: The inverse of the time between each target click, measured in seconds. Gildea et al. (2018) and others have used this transform of response time because it reduces the positive skew that is common to response-time distributions and yields intuitively clear units, that is, responses/second.

Our main concern was Score because the ATCOV does not fail anyone based on their response speed. Score is the sole arbiter of success. Score exhibits a strong ceiling effect, though, so response speed provides a second measure that can reflect task difficulty even when the Score is 100.

3.3 Search tasks

We used the same tasks and software (with some improvements to the output data-file formatting and other refinements) that Gildea et al. (2018) used to assess the palette's suitability as an FAA standard.

3.3.1 Weather search

The Weather search tests the ability to find targets that are the smallest weather stimuli that appear on ATC displays. A target cue, consisting of a small square in a larger square, precedes each trial. That cueing presentation identifies the target and background colors for the upcoming search (see Figure 9). After the participant clicks the "Go" button, a 6 x 8 array of stimuli is presented (see Figure 10).

Figure 9. Example of a Weather search cueing presentation with a Wx-Yellow target and Wx-Green background

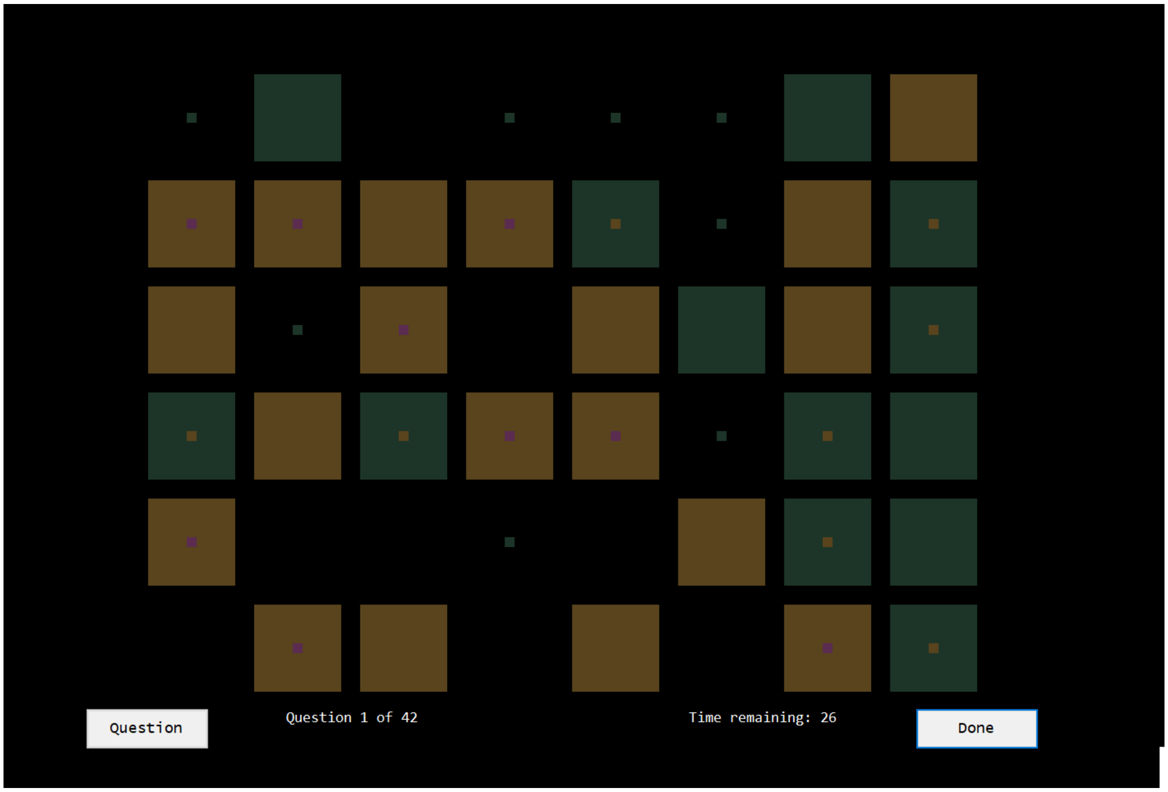

Figure 10. Example of a Weather search trial with Wx-Yellow targets on Wx-Green backgrounds

We assigned each stimulus one of the three luminous (i.e., non-Black) weather colors as the foreground color and another weather color as the background color. These pairings were subject to the same constraint that Gildea et al. (2018) imposed: Each weather color must appear against a background that denotes the next lower severity. This bit of realism meant the targets could be identified by their background colors, which were larger and easier to see, so we added a spoiler: Each background square might or might not contain a target. We therefore had six possible stimuli (see Figure 11). The participant's task was to click each square that contained the target color as quickly as possible, using the computer's mouse, and then click the "Done" button.

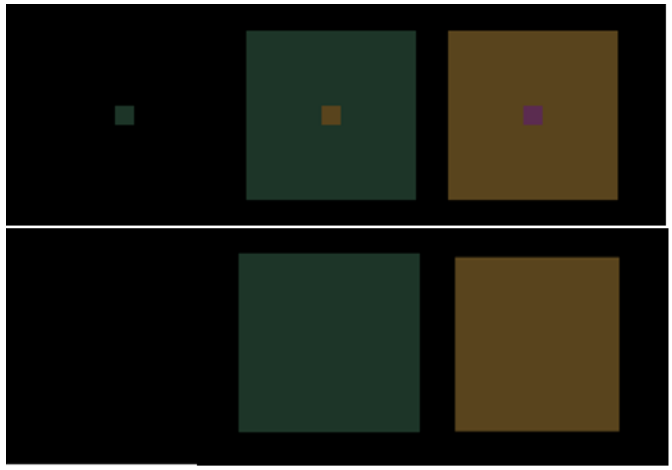

Figure 11. Weather search target and background color combinations

We used each of the three possible target/background color combinations 14 times, yielding a total of 42 trials per session. The number of targets/trial varied randomly from 1 to 10. Participants could click a “Question” button that showed the cuing presentation again and stopped the timer if they clicked “Go” prematurely and then realized they were unsure of the target. (This happened sometimes to participants who were trying to finish quickly.) Each trial ended when the participant clicked the “Done” button or 30 seconds elapsed.

3.3.2 Foreground-Color search

The Foreground-Color search tested the ability to find targets based solely on their color. The cueing presentation was an ATC data block, using one of the 44 possible foreground/background color pairings (see Figure 12).

Figure 12. Example of a Foreground-Color search cueing presentation using an Orange target and Wx-Yellow background

We presented a 6 x 6 array of data blocks after the participant clicked the "Go" button (see Figure 13). Those data blocks used the same character strings as the cueing presentation and were assigned one of the foreground colors randomly, subject to the constraint that the number of targets/trial ranged from 1 to 10. Personal experience with the program convinced us that 30 seconds were enough to find and click 10 targets. Each of the 11 foreground colors was the target color once for each of the 4 weather colors, yielding a total of 44 trials per session. We rendered the data blocks in the Consolas font, which is a sans serif font that resembles the ones used on ERAM, STARS, and ATOP. The characters subtended 20 arc-minutes visually at a 50-cm viewing distance, as specified by FAA-HF-STD-010A (2020, Section 5.6.2). We did not control participant viewing distance, however. Instead, like the ATCOV, we allowed them to change their viewing distance freely, just as air traffic controllers do.

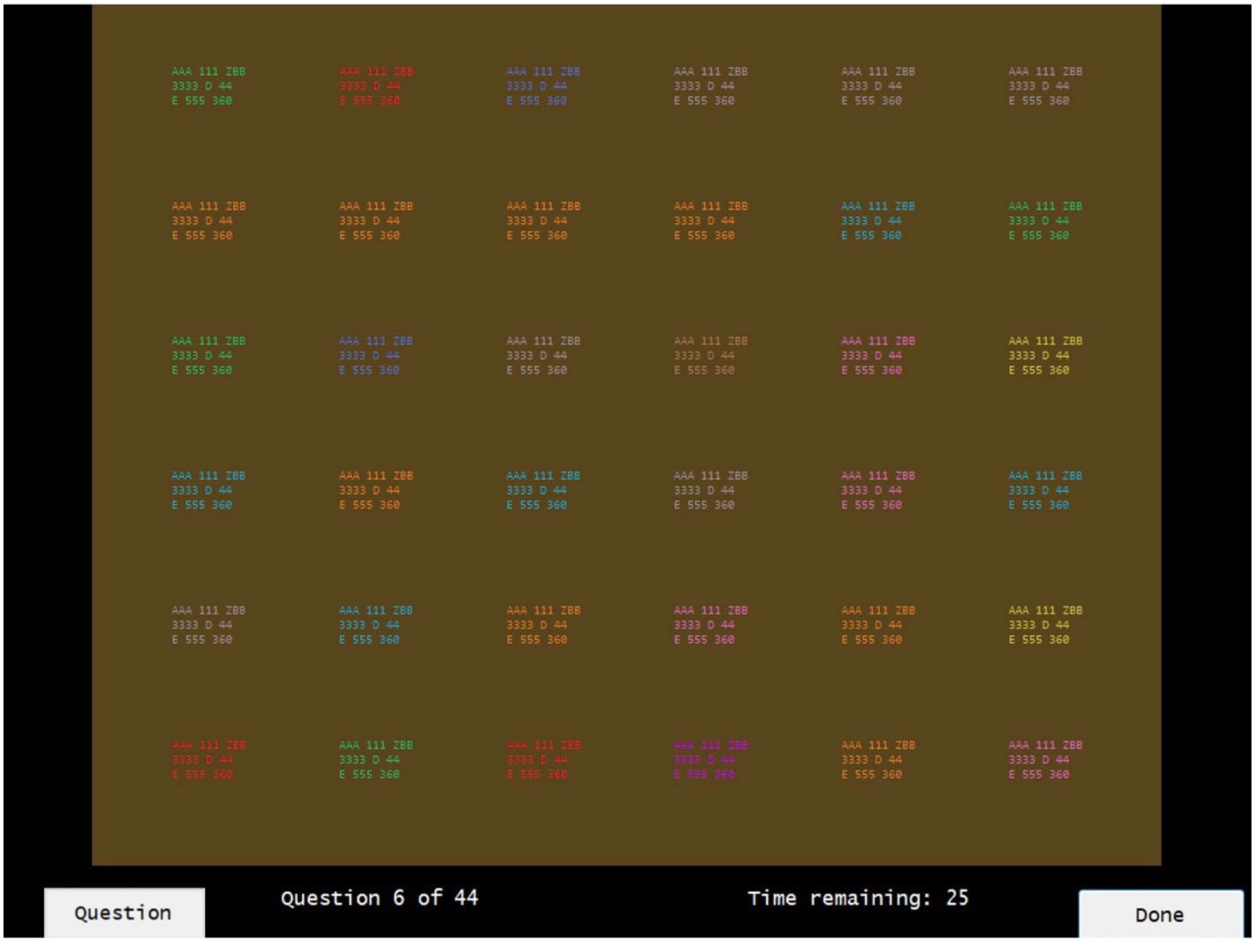

Figure 13 Example of a Foreground-Color search trial with Orange targets and Wx-Yellow background

3.3.3 Shape search

The Shape search tested the foreground colors' legibility. We preceded each trial with presentation of a three-character alphanumeric string. That cueing presentation identified the target string for the upcoming search (see Figure 14).

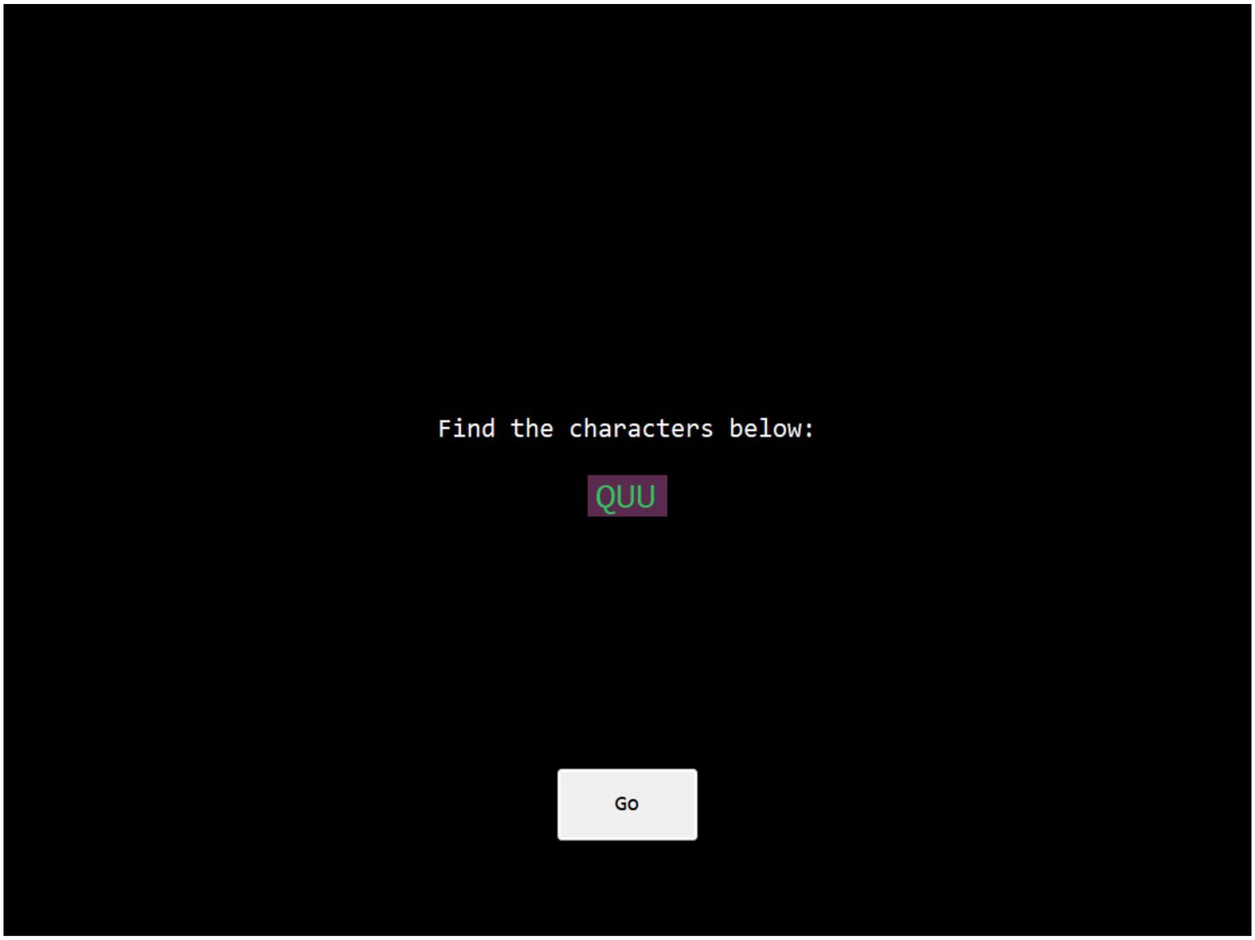

Figure 14. Example of a Shape search cueing presentation with Green "QUU" target, and Wx-Red background

After the participant clicked the "Go" button, we presented a 6 x 6 array of data blocks (see Figure 15). At least one began with the target string. Each of the 11 foreground colors was the data block color once for each of the 4 weather background colors, yielding a total of 44 trials per session.

Figure 15. Example of a Shape search trial with "QUU" targets, Green foreground, and Wx-Red background

3.3.4 Redundant-Coding search

The Redundant-Coding search used color and shape to encode targets, thereby satisfying an FAA-HF-STD-010A (2020, Section 4.2) requirement that critical information be coded redundantly. We preceded each trial with presentation of a three-character alphanumeric string using one of the 11 foreground colors (see Figure 16).

Figure 16. Example of a Redundant-Coding search task cue with “629” in Magenta target and Black background

After the participant clicked the “Go” button, we presented a 6 x 6 array of data blocks (see Figure 17). The number of targets varied randomly from 1 to 10 and could be identified by recognizing the character string or its color. Each of the 11 foreground colors was the target color once for each of the 4 weather background colors, yielding a total of 44 trials per session.

Figure 17. Example of a Redundant-Coding search trial with Magenta "629" targets and Black background

3.4 Pilot study

The pilot study's purpose was to identify the best combination of peak luminances and search tasks for the main study. Therefore, we evaluated participant performance as we progressed through the pilot study and made continual adjustments to our test conditions.

We ran a total of 11 NCV participants through all four search tasks at 55,000 lux of illumination. We began by running them with the peak D_{65} luminance set to 600 and 750 cd/m^2, alternating the order from one participant to the next, to determine where asymptotic performance occurred. Then, we looked for a lower luminance that should produce unacceptable performance. We settled on testing 450 to 700 cd/m^2 in 50-cd/m^2 increments. This sampling provides six data points from which the underlying relationship between peak display luminance and performance may be inferred.

We included the Shape search at 550 cd/m^2 as a spot-check because the pilot data led us to expect asymptotic performance for that task and luminance. We chose the Shape search because it tests legibility alone, without a redundant color cue, which put the NCVs and CVDs on equal footing.

3.5 Main experiment

The main experiment used a total of 39 participants. We made an initial estimate of the number required to achieve sufficient statistical power using data from the pilot study and refined it as data from the main experiment became available.

For the Weather search, we used a 6 x 6 balanced Latin square to reduce order effects caused by practice, fatigue, and the changing TDM peak luminances. We replicated the square six times, giving us a total of 36 participants. The experimental design was a three-way full-factorial fixed-effects analysis of variance (ANOVA) using Color-Vision Status (i.e., NCV vs. CVD), TDM Backlight Luminance, and Target Color as main effects and Score and Response Speed as dependent measures (two ANOVAs total).

The Shape search at 550 cd/m^2 followed the six Weather search tasks. Interim analysis (i.e., during data collection) caused us to wonder whether performance is truly asymptotic at that luminance, so we expanded this second part of the experiment for the remaining five participants, plus three more. This group consisted of four NCVs and four CVDs. The expanded second part involved performing Shape searches at 550, 600, 650, and 700 cd/m^2, with order effects reduced via use of a 4 x 4 balanced Latin square. We based the number of participants on a power analysis that showed eight should be sufficient to detect a reliable difference among the luminances, if one exists.

3.6 Power considerations

One purpose of our experiment was to determine whether CVD performance with the standard color palette compares acceptably with NCV performance when daytime ambient illumination is present. Another purpose was to learn whether the STARS TDMs produce enough luminance to overcome worst-case ambient illumination, meaning performance is indistinguishable from the low-ambient results reported by Gildea et al. (2018). If both criteria are met, the palette is suited for use as an FAA standard on STARS TDMs under daytime ambient illumination. Meeting those criteria implies failing to reject null hypotheses, though, which poses a hazard: If the statistical power for a comparison is low, we might take a resulting failure to reject the null as evidence the palette is acceptable when it is not; that is, we might commit a Type II error. To boost power and reduce the likelihood of Type II error, we set α equal to 0.1 and used Fisher's Least Significant Difference (FLSD) test for our post-hoc paired comparisons. The FLSD produces the smallest critical differences of all the post-hoc paired comparison tests, making it the most sensitive option and the best choice for our purposes. If the FLSD shows that a comparison is not significant, all the alternatives would, too.

4 Results

4.1 Weather Search

4.1.1 Data distributions

Figures 18 and 19 show boxplots of our participants' Weather search Scores at each peak D_{65} luminance setting. Figures 22 and 23 show corresponding boxplots of their Response Speeds. They are rotated 90 degrees to facilitate comparisons with the corresponding histograms.

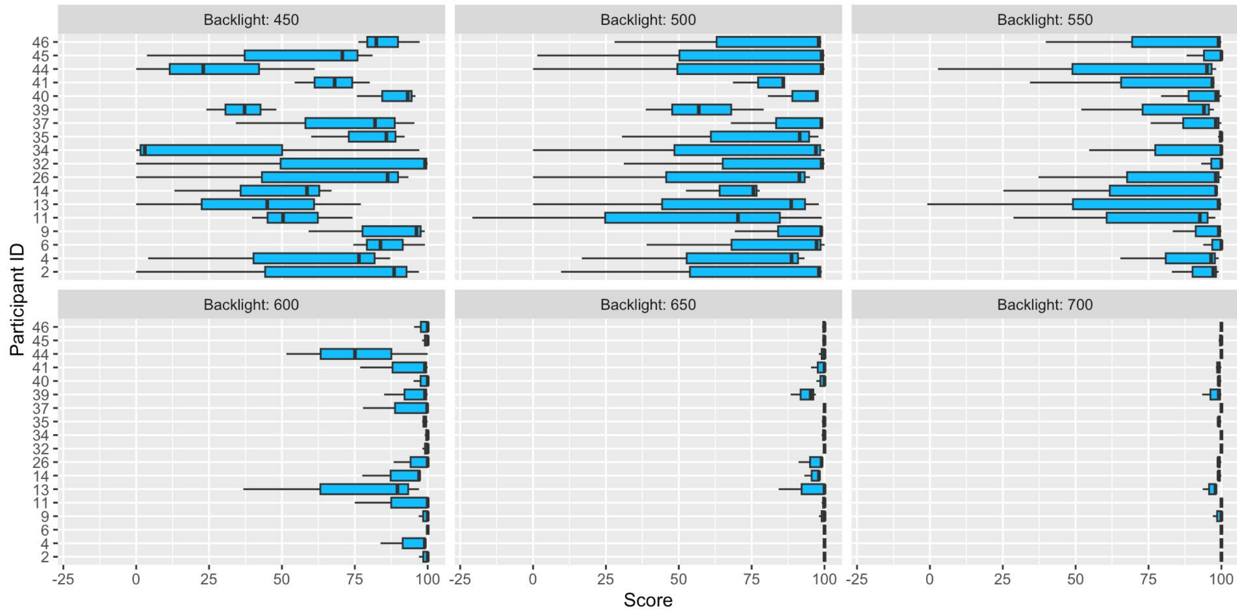

Figure 18. Boxplots of NCV Weather-search Score

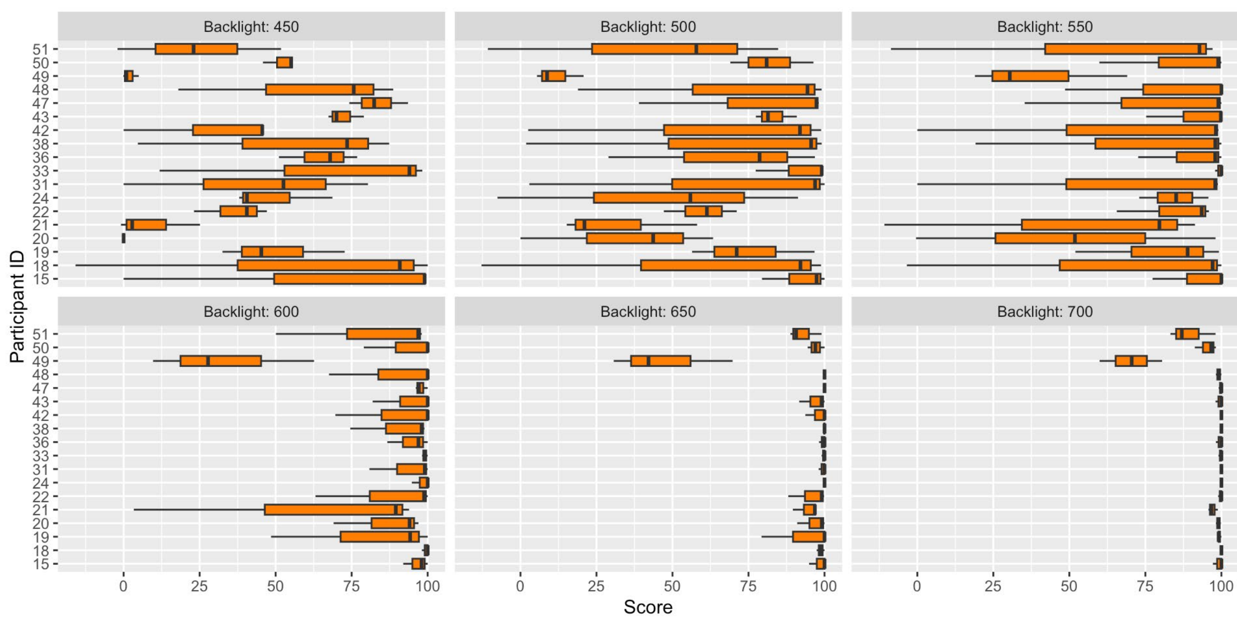

Figure 19. Boxplots of CVD Weather-search Score

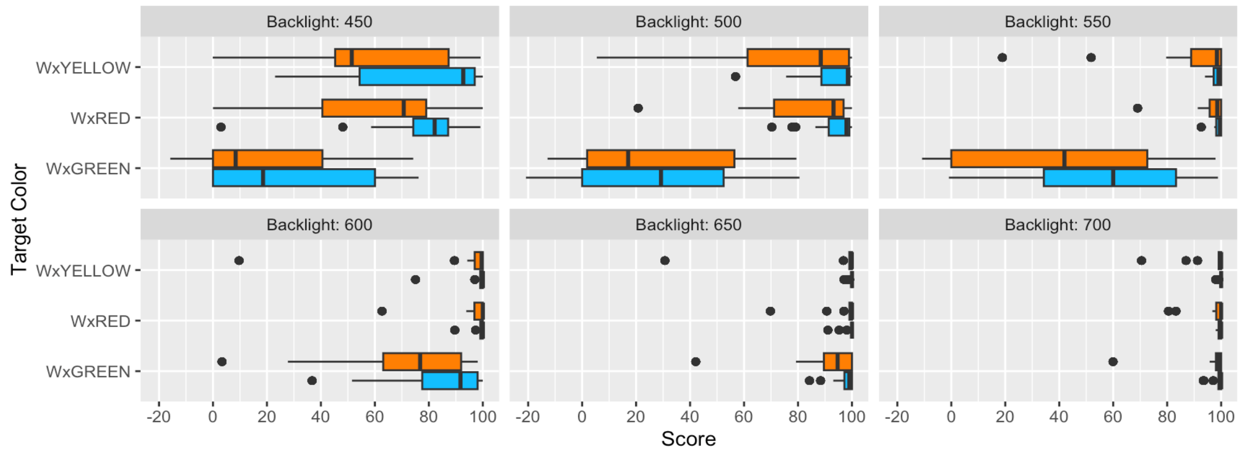

Figure 20. Boxplots of NCV and CVD Weather-search Score for each weather color.

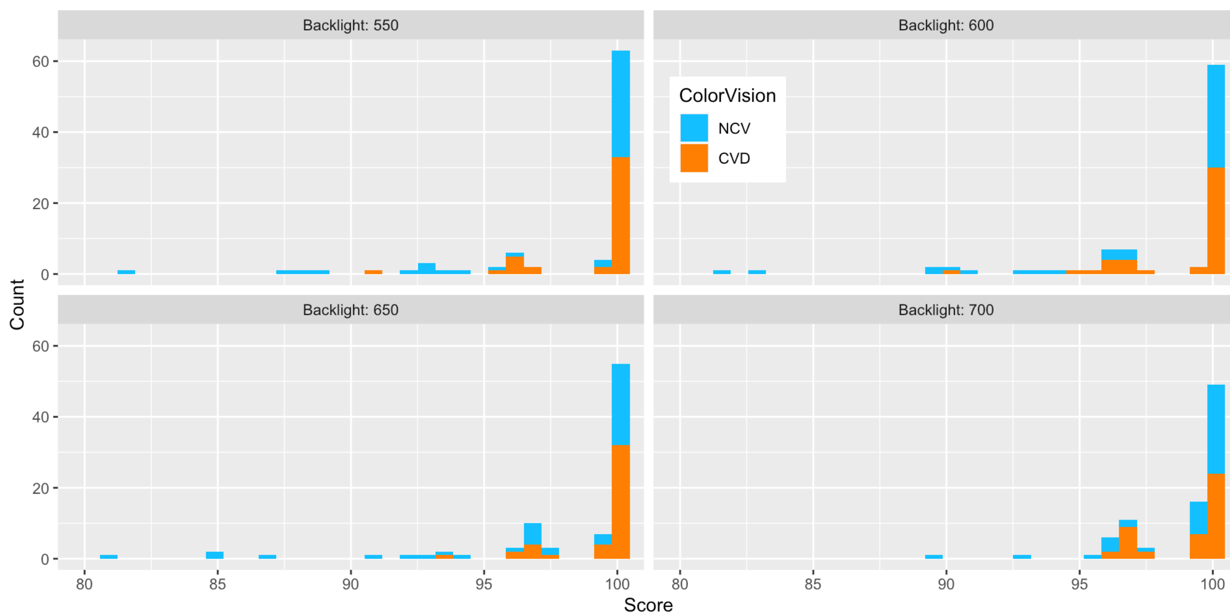

Figure 21. Histograms of NCV and CVD Score

Score exhibits a ceiling effect that produces negative skew in the Figure 21 histograms. This feature is apparent in the histograms shown by Gildea et al. (2018, Figures 13 to 16) and Chidester et al. (2011, Figures 4 to 7 and 2013, Figures 6 and 7), also. A consequence of the ceiling effect that is apparent in Figures 18 to 25 is that the variance in Score and Response Speed decreases as peak luminance increases. Figure 20 shows that Score for Wx-Green was lower than the other colors until the peak luminance reached 700 cd/m^2. Figure 24 shows no corresponding effect for Response Speed.

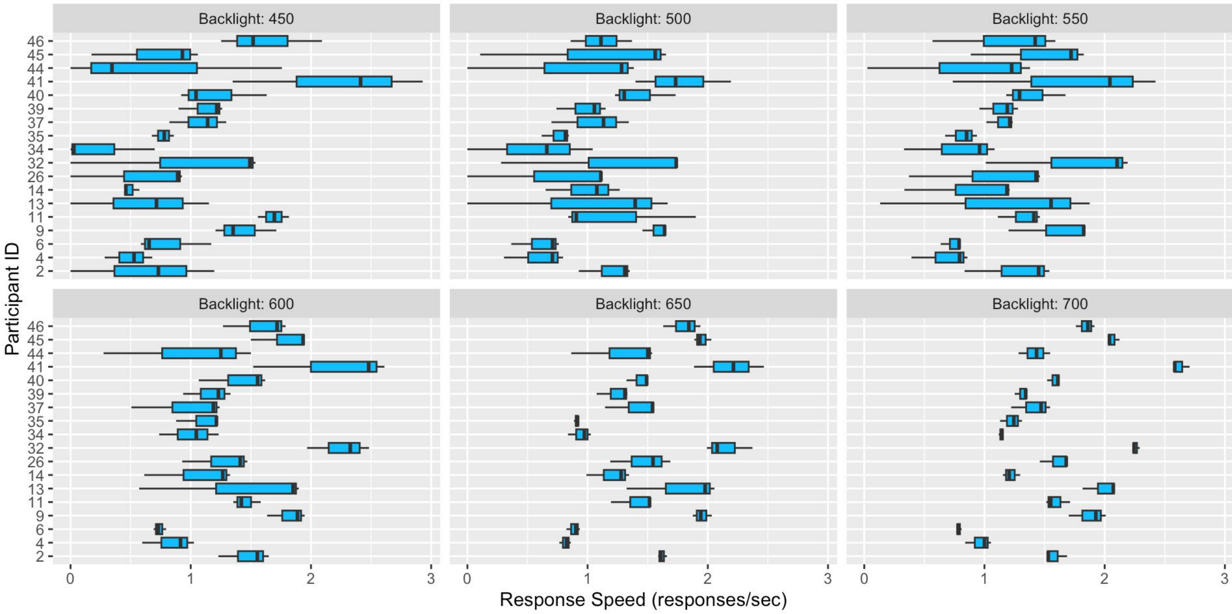

Figure 22. Boxplots of NCV Weather-search Response Speed

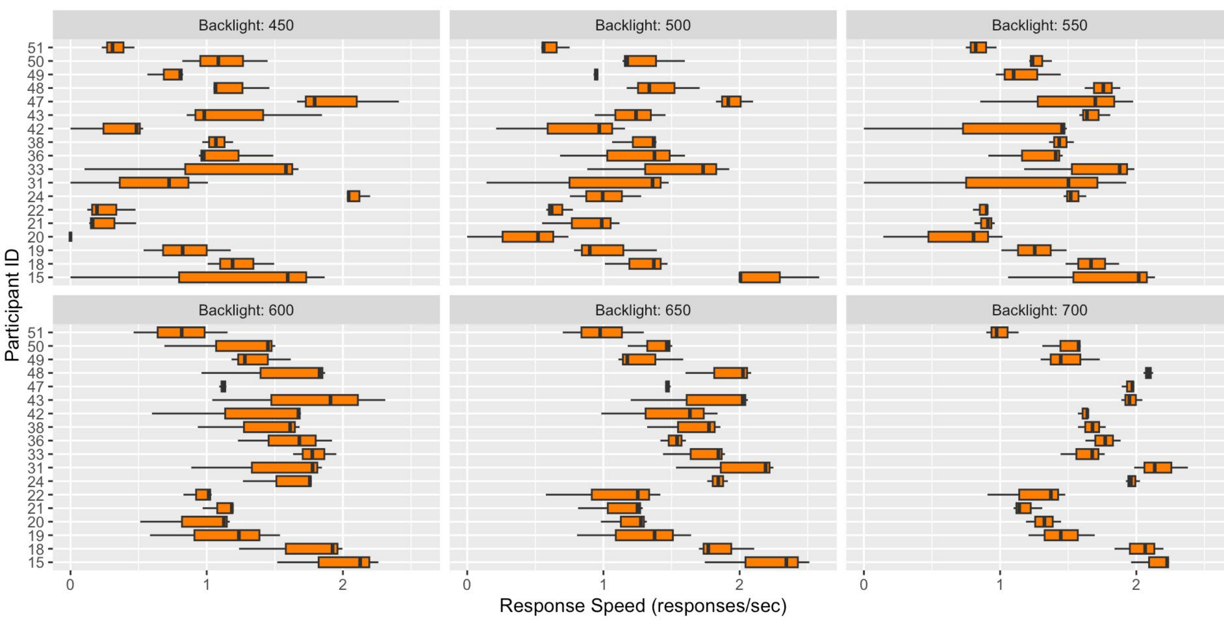

Figure 23. Boxplots of CVD Weather-search Response Speed

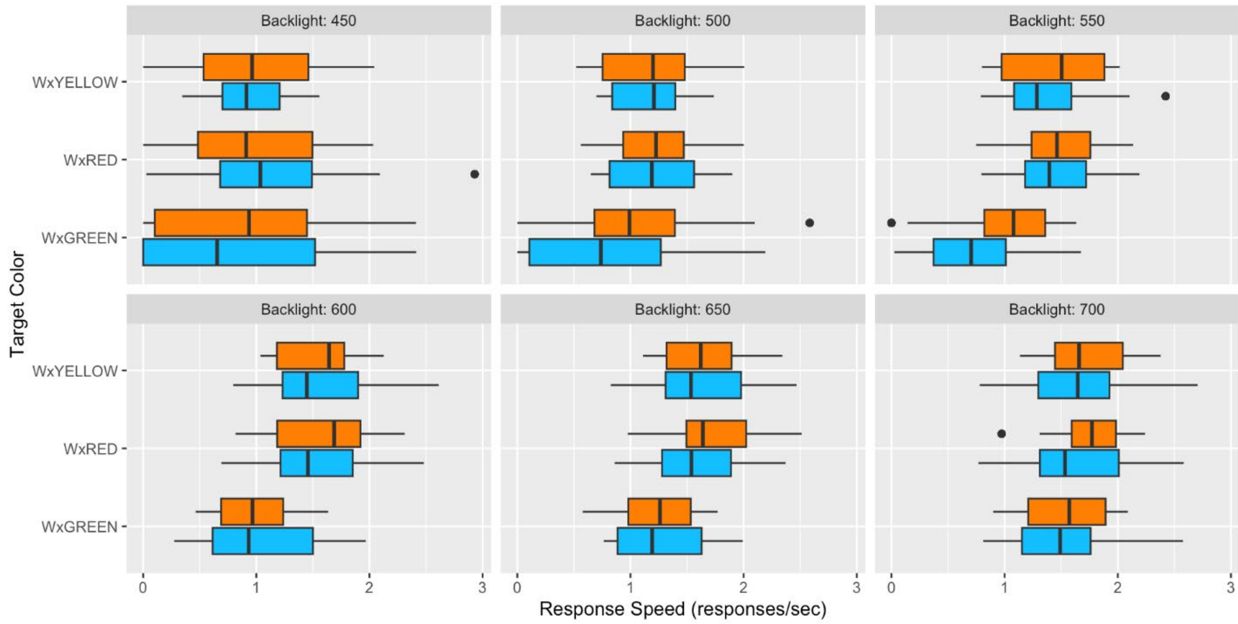

Figure 24. Boxplots of NCV and CVD Weather search Response Speed for each target color

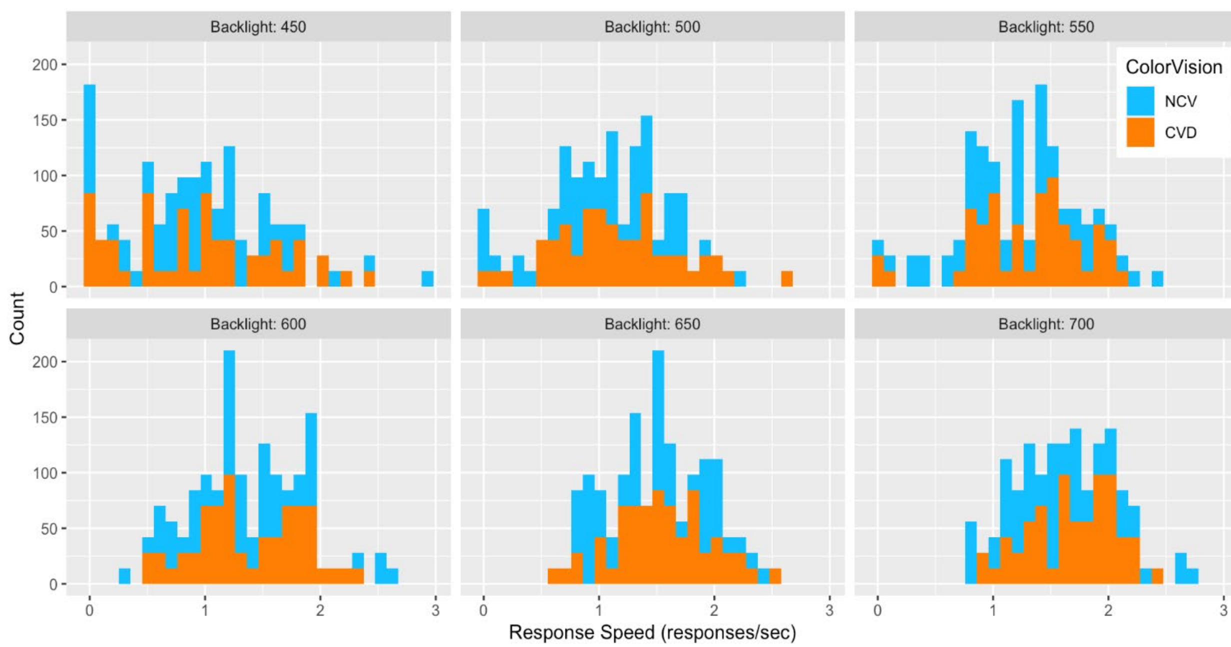

Figure 25. Histograms of NCV and CVD Weather search Response Speed

4.1.2 ANOVA for Weather-Search Score

We summarize the ANOVA results in Table 5. Using α = 0.1, the main effect of Color Vision Status is significant but its generalized eta[2] (GES) is small. The Backlight x Color-Vision Status interaction shown in Figure 26 is also significant but its GES is even smaller. (The bars in this and subsequent figures show the critical difference per the FLSD.) The FLSD critical difference indicates the CVDs' disadvantage is not reliable statistically at any of the luminances we tested. The trends in Figure 26 are obvious, nonetheless: The NCVs reached asymptote at a mean Score of 99 near 700 cd/m^2 whereas visual extrapolation indicates our CVDs' asymptote is nearer 750 cd/m^2. The cut score Gildea et al. (2018, Table 2) derived for the Weather search is 98, so the NCVs passed at 700 cd/m^2. The CVDs' mean Score at 700 cd/m^2 is 97, which falls a bit short and is due to Participants 49 and 51 (see Figure 19). Increasing the luminance to 750 cd/m^2 would probably cure that shortfall.

Table 5. Significant effects on Weather-search Score

Effect	***p***	**GES**
Color Vision Status	= 0.05	0.058
Backlight Luminance	<< 0.01	0.496
Target Color	<< 0.01	0.418
Backlight x Target-Color	<< 0.01	0.25
Backlight x Color-Vision Status	= 0.09	0.016

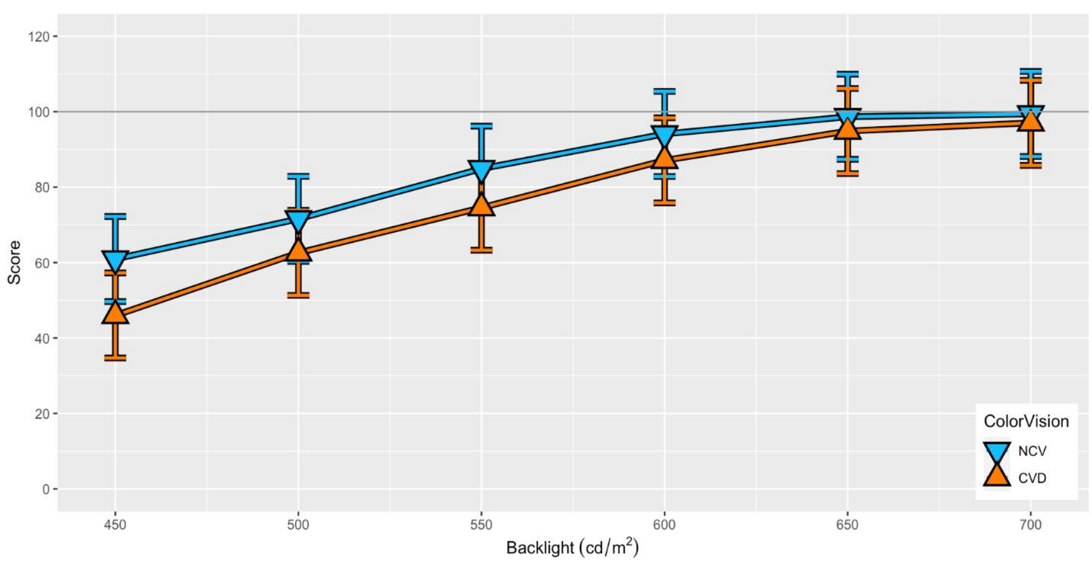

Figure 26. Backlight x Color-Vision Status interaction for Weather-search Score ($p = 0.09$, GES = 0.016)

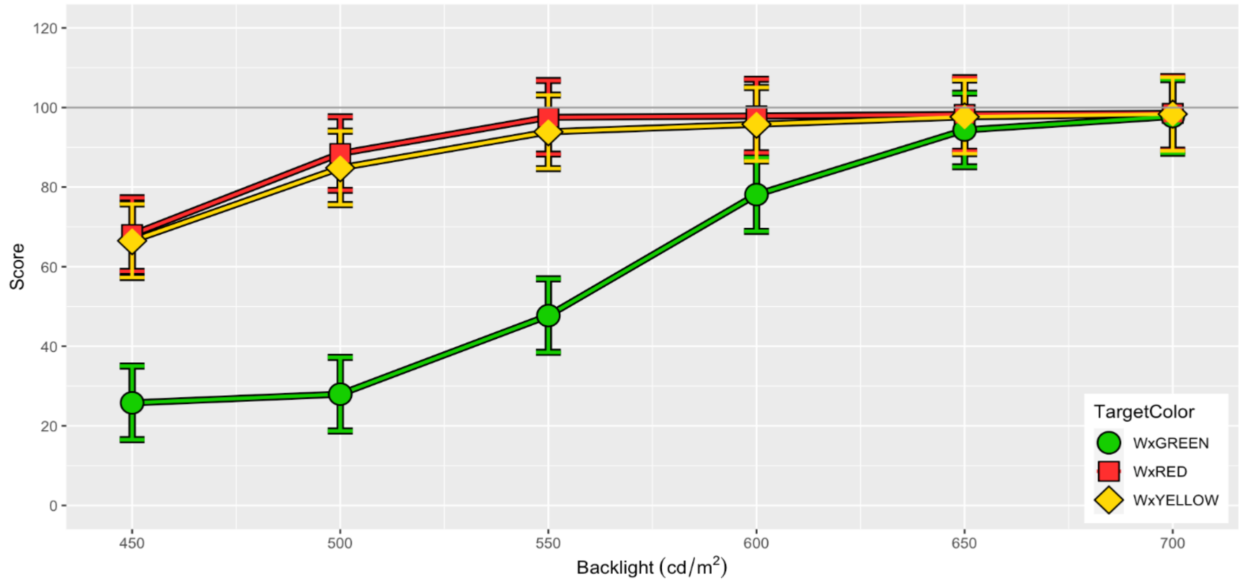

Figure 27. Backlight x Target-Color interaction for Weather-search Score ($p << 0.01$, GES = 0.25)

The effects that account for the most variance are the main effects of Backlight and Target Color and their interaction. Figure 27 shows that Wx-Yellow and Wx-Red Score benefit from increasing peak luminance up to 650 cd/m^2, at which point they asymptote at Scores of 98 and 99, respectively. Wx-Green graphs as a sigmoid that asymptotes at a Score of 98 at 650 cd/m^2. We might suppose that Wx-Yellow and Wx-Red would also graph as sigmoids if Backlight extended down to 300 cd/m^2. Comparison of these results with the Gildea et al. (2018) cut score and Figure 1 shows that our NCVs and CVDs performed the Weather search on TDMs under 55,000 lux as accurately as Gildea et al.'s (2018) participants did under 40 lux.

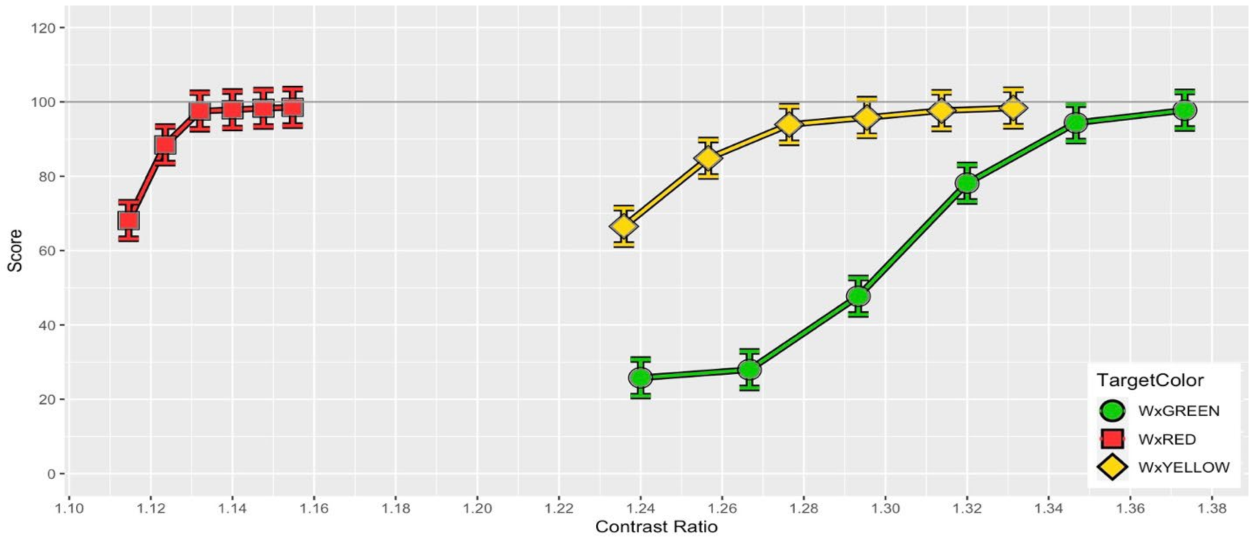

Figure 28. Weather-search Score as a function of contrast ratio

Figure 28 graphs mean Score as a function of contrast ratio, which is a function of Backlight, Target Color, and the ambient illumination. The Wx-Green graph is sigmoidal and, just as in Figure 27, we might suppose that Wx-Red and Wx-Yellow would be too if we had tested lower contrast ratios. The asymptotes are near 1.13, 1.32, and 1.37 for Wx-Red, Wx-Yellow, and Wx-Green, respectively.

Our asymptote estimates resemble values reported previously. Havig et al. (2003) fit Weibull (1951) functions to their data and estimated that contrast ratios of 1.2, 1.42, and 1.64 are needed to achieve 95%-correct color identification for their (much dimmer) red, yellow, and green targets, respectively.

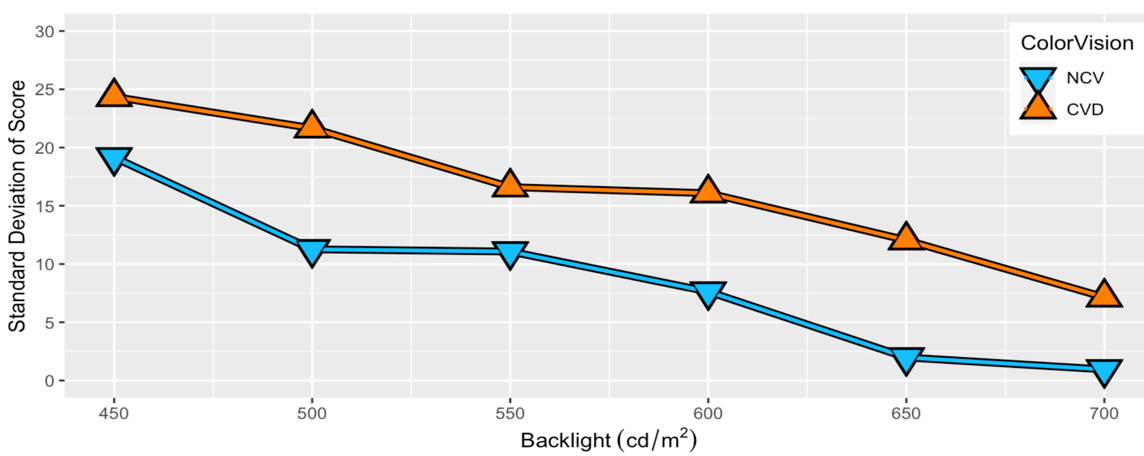

Figure 29. Standard deviation of Weather-search Score as a function of peak D_{65} luminance

Figure 29 shows that CVD Scores were more variable than NCVs' over the range of luminances we tested. Variance in NCV Score is near its minimum when peak luminance reaches 700 cdm^2. Variance for CVDs appears to reach its minimum at a peak luminance near 800 cd/m^2.

4.1.3 ANOVA for Weather-Search Response Speed

The main effects of Backlight and Target Color and the three-way interaction are significant. Figure 30, which graphs the interaction, shows the differences among the colors are significant only for the NCVs at 550 cd/m^2, unlike our results for Score. Figure 30 shows also that Response Speed increased with luminance for all three target colors, with no sign of an asymptote. Mean Response Speed at 700 cd/m^2 is 1.58 and 1.68 responses/sec for NCVs and CVDs, respectively. These speeds are notably slower than Gildea et al. (2018) obtained (2.74 for NCVs and 2.29 for CVDs). The most likely reason is the added uncertainty the spoiler caused.

Table 6. Significant effects on Weather-search Response Speed

Effect	***p***	**GES**
Backlight Luminance	<< 0.01	0.229
Target Color	<< 0.01	0.129
Color-Vision x Backlight x Target-Color	= 0.06	0.007

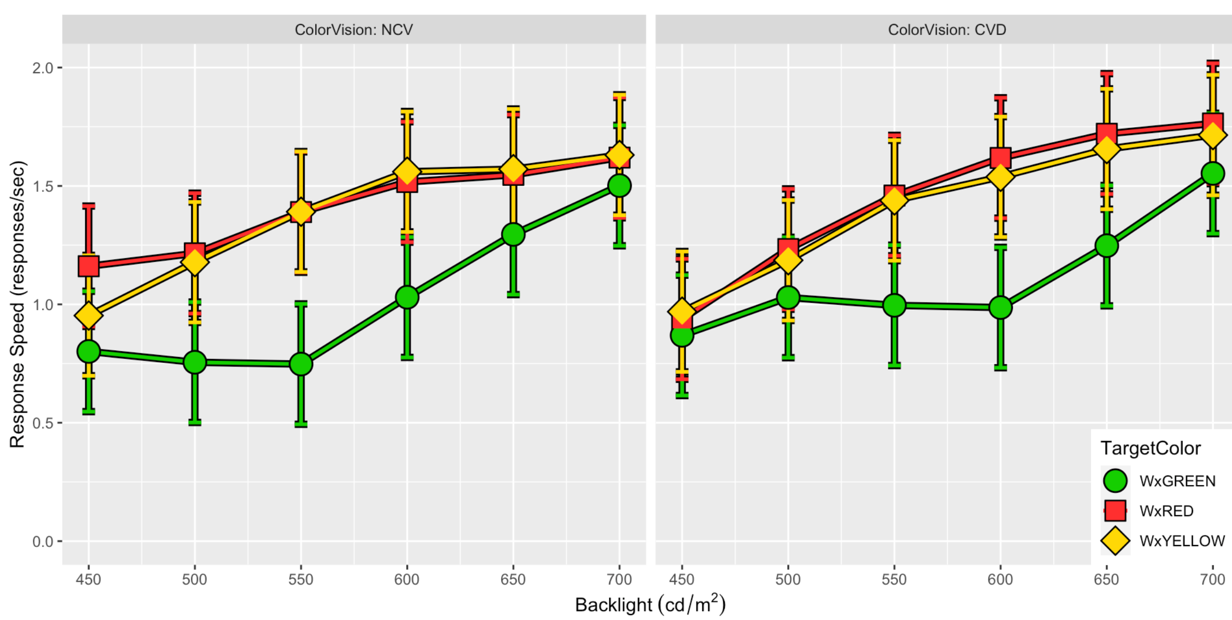

Figure 30. Three-way interaction for Weather-search Response Speed ($p = 0.06$, GES = 0.007)

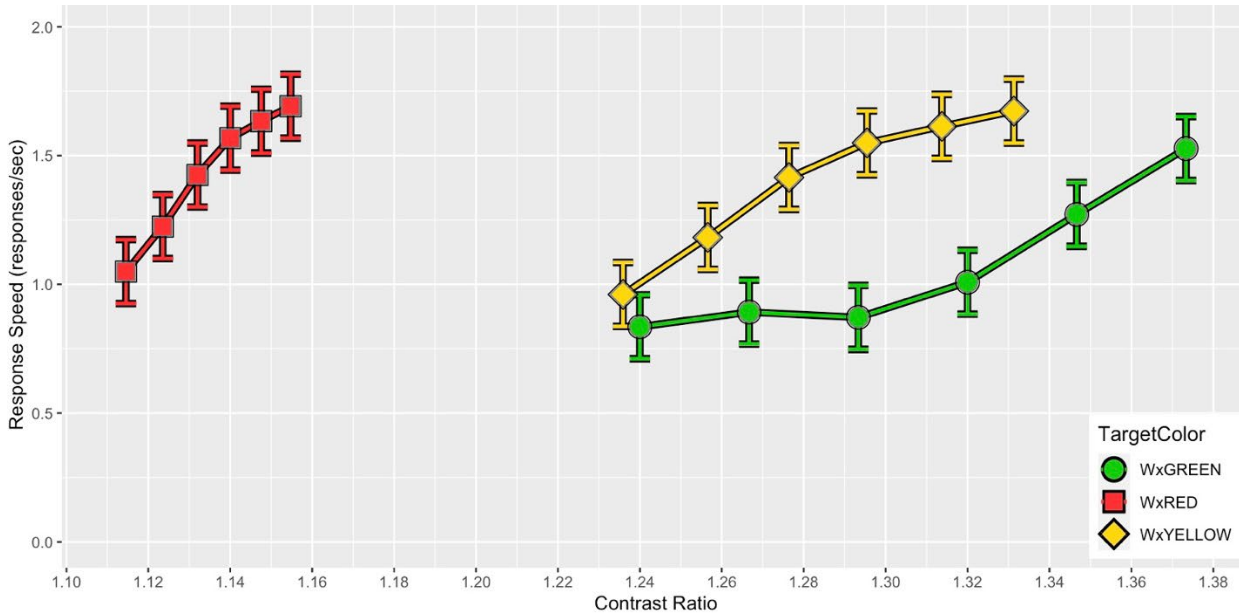

Figure 31. Weather-search Response Speed as a function of contrast ratio

The main effect of Color Vision Status (NCV mean = 1.27; CVD = 1.33) is not significant (p = 0.6) and the opposite of expectation. The benefit of increasing contrast ratio, shown in Figure 31, is monotonic up to 1.37 with no clear signs of where the asymptotes lie. Otherwise, Figure 31 resembles Figure 28's graph of Score vs. Contrast Ratio.

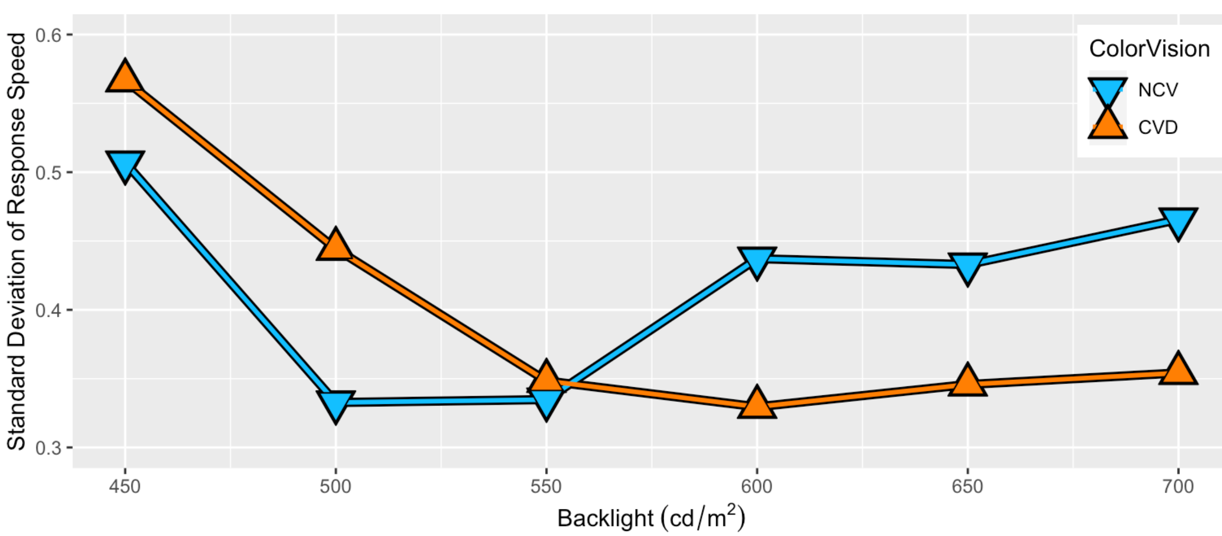

Figure 32. Standard deviation of Weather-search Response Speed as a function of peak D_{65} luminance

Figure 32 shows that the standard deviation (SD) of CVD Response Speed decreased as luminance increased and stabilized at a value near 0.35 at 650 cd/m^2. NCV SDs reached a minimum near 0.34 at 500 cd/m^2 but then increased and seemed to stabilize near 0.45. There is no apparent reason why increasing luminance should increase Response Speed variance or why NCV variance should be greater than CVD variance.

4.2 Shape Search

4.2.1 Data distributions

Figures 33 and 34 show boxplots of our participants' Shape search scores at each peak luminance. Figure 35 displays corresponding boxplots for each target color. Figure 33 shows that Participant 41's Score increased and became less variable at 700 cd/m^2. The histograms in Figure 36 show that the variance in NCV Score decreased as Backlight luminance increased whereas variance in CVD Score was consistently low at all Backlight luminances. Otherwise, there are no clear effects of Backlight or Target Color on Score. The distributions in Figures 33 to 36 show the same ceiling effect that is evident in the Weather search results. The boxplots and histograms for Response Speed in Figures 37 – 41 exhibit no noteworthy effects.

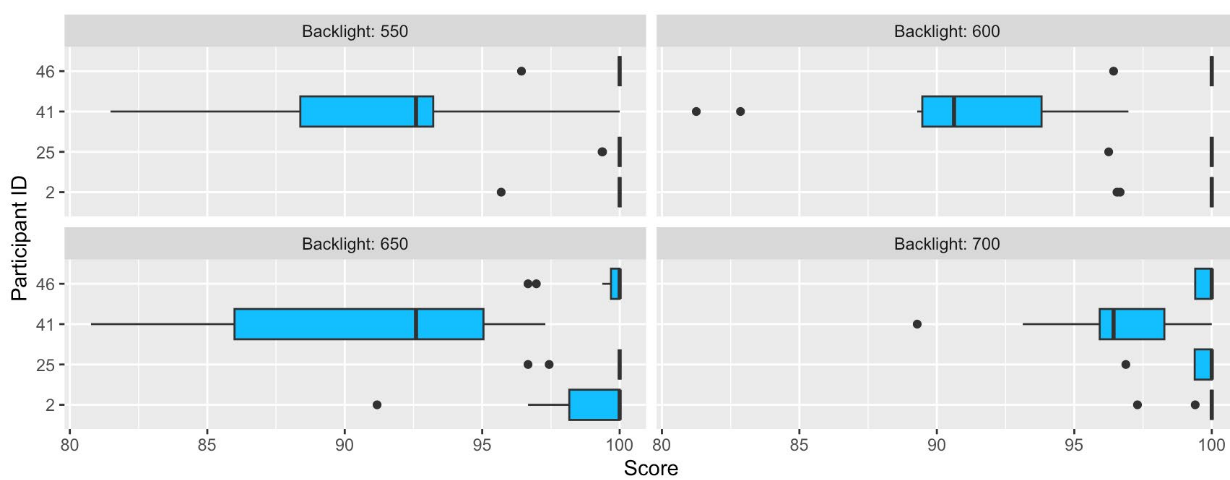

Figure 33. Boxplots of NCV Shape-search Score

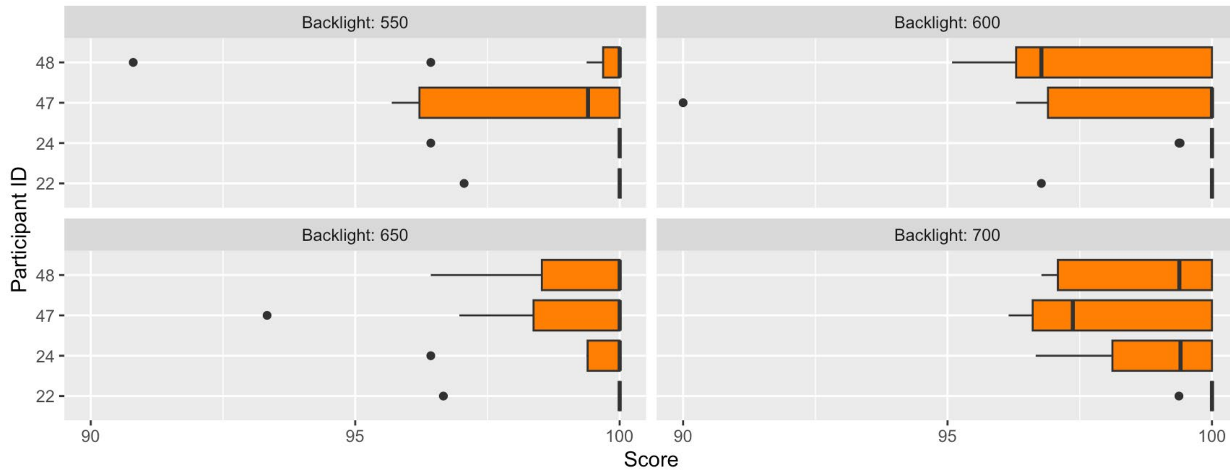

Figure 34. Boxplots of CVD Shape-search Score

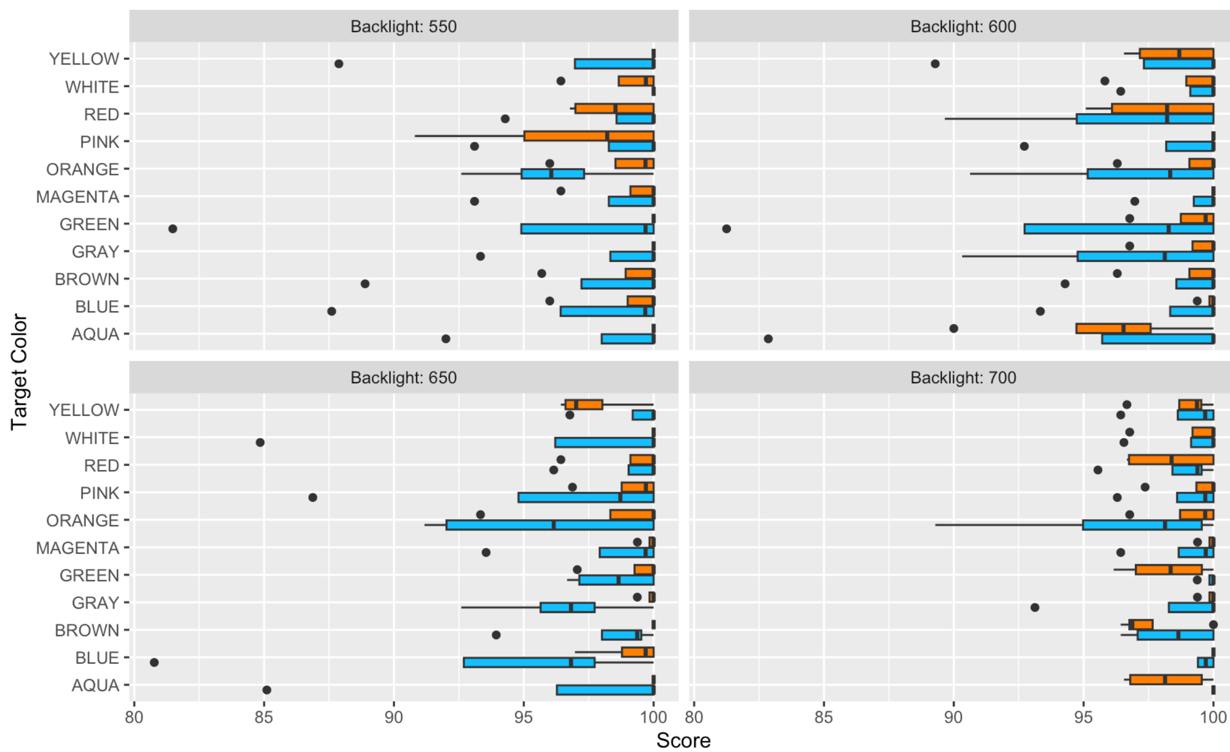

Figure 35. Boxplots of NCV and CVD Score for each target color

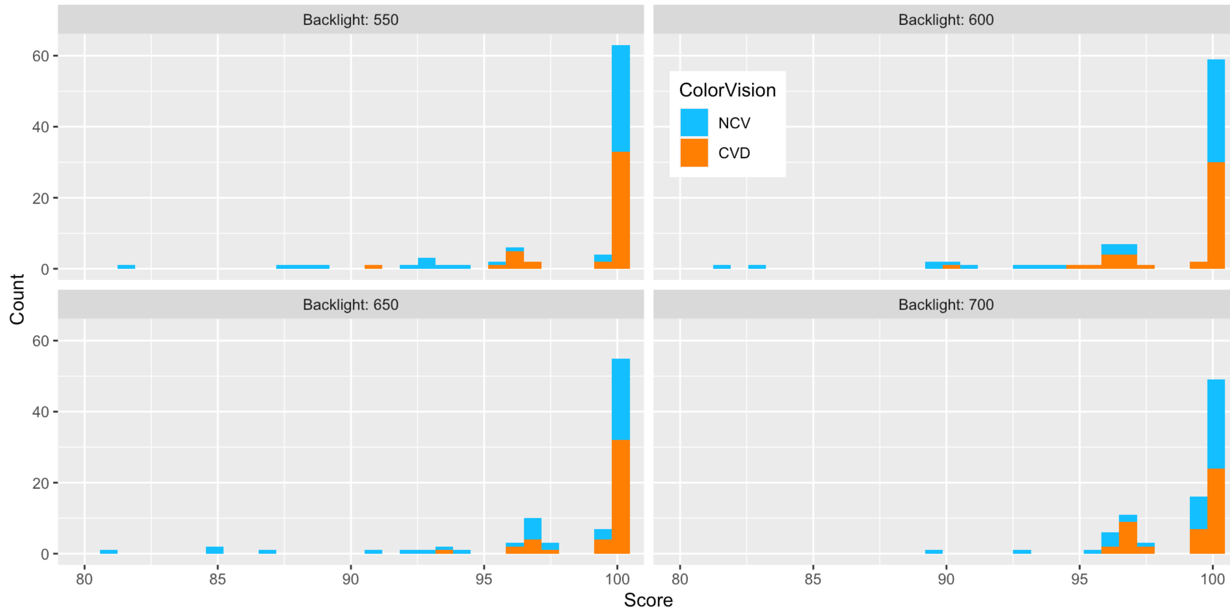

Figure 36. Histograms of NCV and CVD Shape-search Score

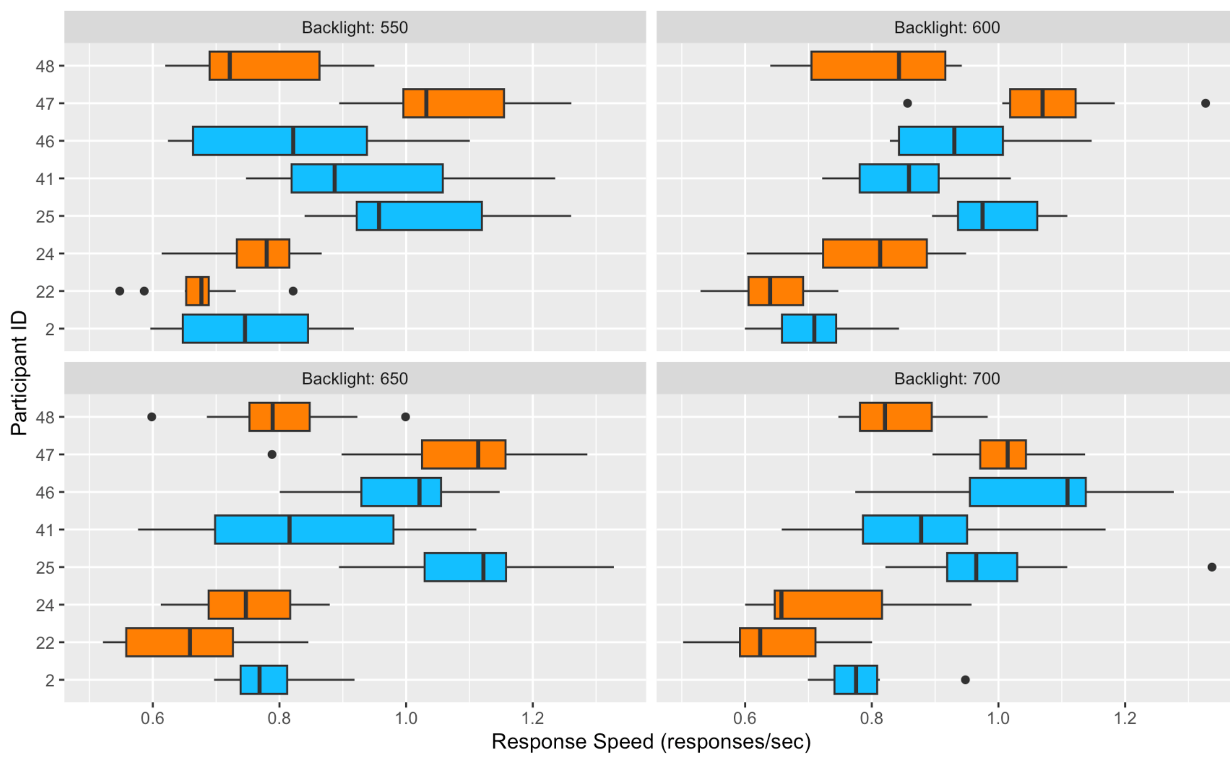

Figure 37. Boxplots of NCV and CVD Shape-search Response Speed

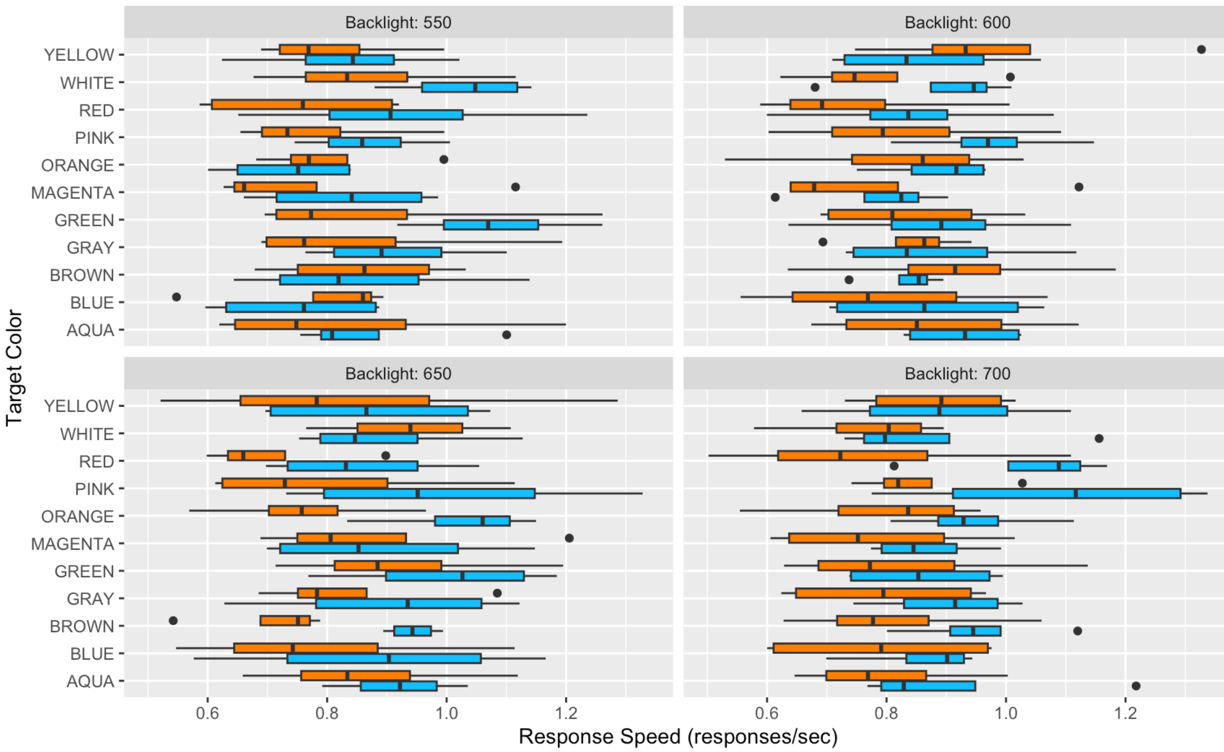

Figure 38. Boxplots of NCV and CVD Shape-search Response Speed for each target color

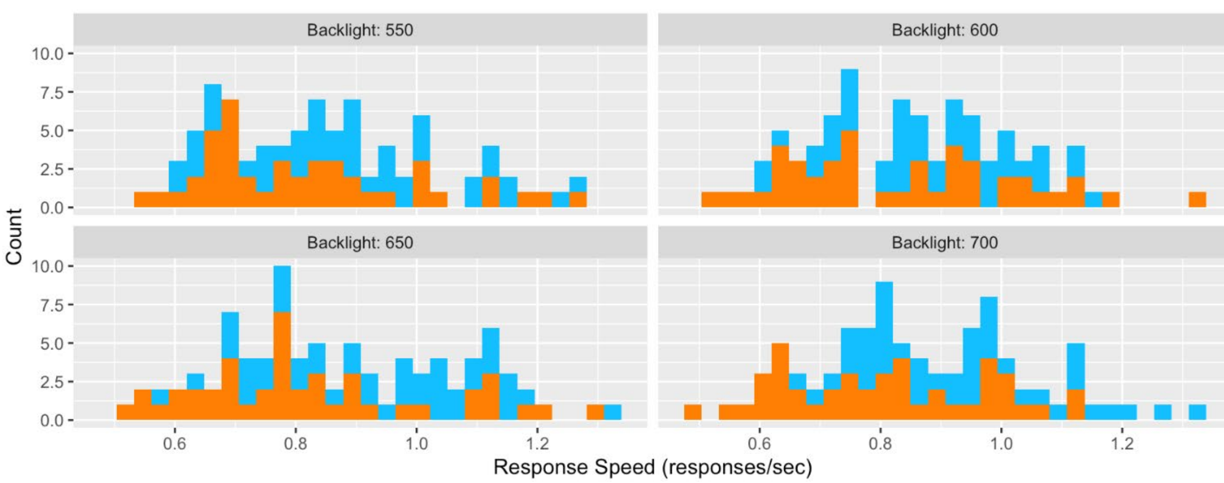

Figure 39. Histograms of NCV and CVD Shape-search Response Speed

4.2.2 ANOVA for Shape-Search Score

The ANOVA shows that the only significant effect is Target Color ($p = 0.07$) and its effect size is trivial (GES = 0.025). The effect is not evident in Figure 4. The effect of contrast ratio (Figure 43) appears to be nil and noisy. (Each datapoint is the average of one trial for each of the eight participants, i.e., the mean of only eight trials, so the variance in Figure 43 is large compared with our Weather-search figures.) The most likely reason for finding no noteworthy effects is that Score for this task asymptotes at a lower contrast ratio than the lowest one that we tested, i.e., 1.82. Figure 44 shows that the NCVs' Scores were more variable than the CVDs' at most luminances, but approached equality at 700 cd/m^2.

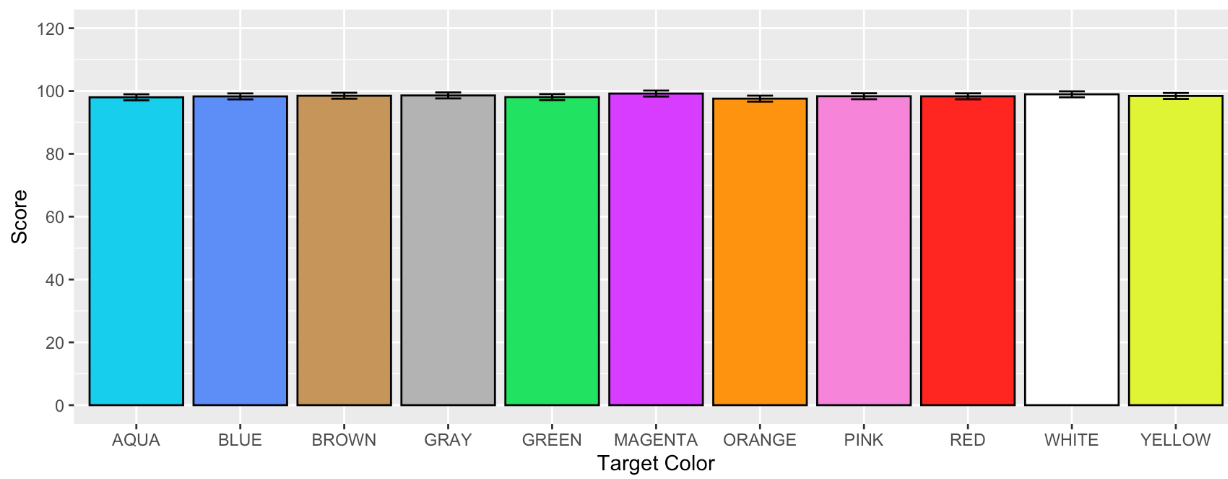

Figure 40. Main effect of Target-Color for Shape-search Score ($p = 0.07$, GES = 0.025)

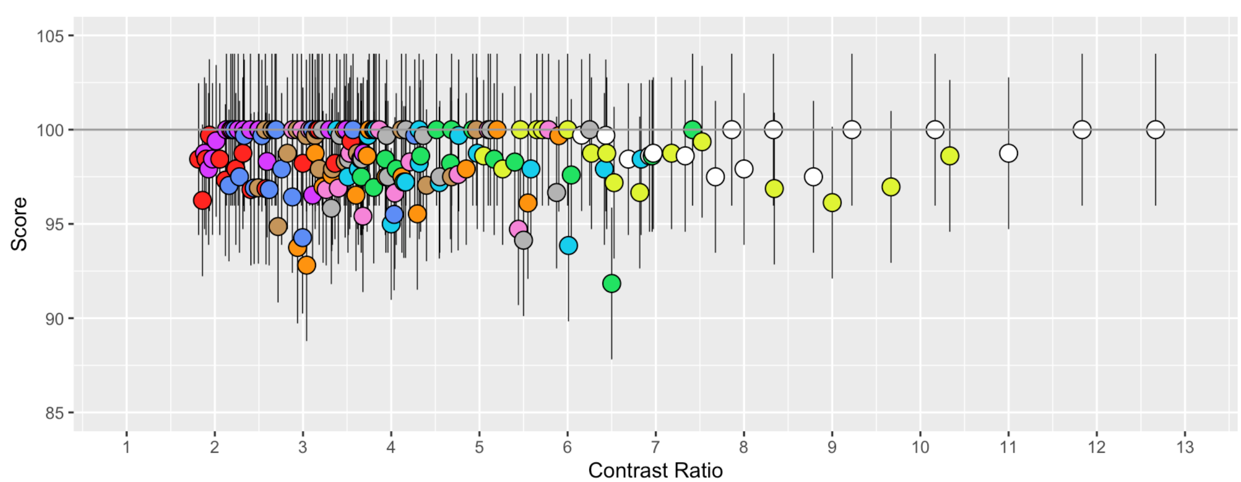

Figure 41. Shape-search Score as a function of contrast ratio

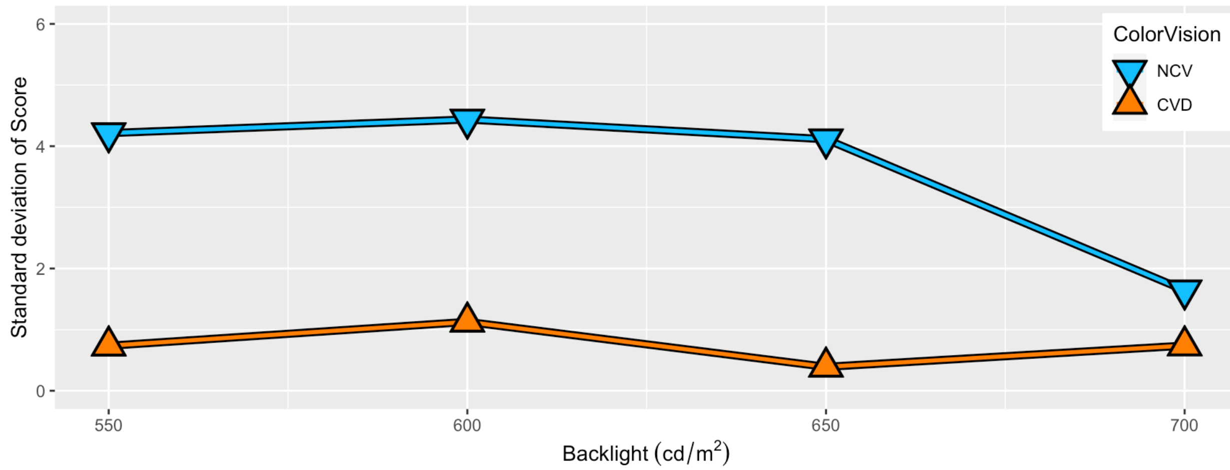

Figure 42. Standard deviation of Shape-search Score as a function of peak D_{65} luminance

4.2.3 ANOVA for Shape-Search Response Speed

The Response Speed ANOVA shows the main effect of Target Color and the Color-Vision x Target-Color and Backlight x Target-Color interactions are significant, but the GES values show that these effects account for little of the variance (see Table 8). The main effect of Target Color is more discernible in Figure 44 than in Figure 41, but it is difficult to say which colors differ reliably from others.

The only pattern that is evident in Figures 44 to 46 is that CVDs are a bit slower than NCVs for most target colors (see Figure 45). Our participants were faster than Gildea et al.'s (2018): 0.9 vs. 0.62 responses/sec for NCVs and 0.82 vs. 0.56 responses/sec for CVDs. These increases are probably due at least in part to our higher stimulus luminances because reaction time is known to decrease as luminance increases (Lit et al., 1971). Indeed, the effect is evident in Figure 30. There is no clear effect of contrast ratio (see Figure 47). These observations reinforce the idea that performance on the Shape search was asymptotic already at the lowest contrast ratio (1.82) and the lowest Backlight level (550 cd/m^2).

Table 7. Significant effects on Shape-search Response Speed

Effect	***p***	**GES**
Target Color	= 0.01	0.038
Target-Color x Color-Vision Status	= 0.05	0.03
Target-Color x Backlight Luminance	= 0.06	0.029

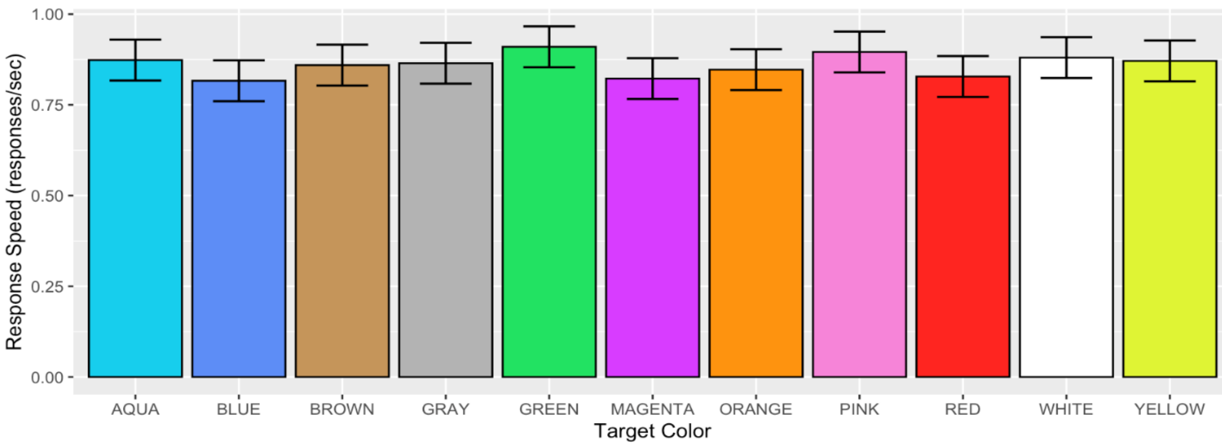

Figure 43. Effect of Target Color on Shape-search Response Speed (p = 0.01, GES = 0.038)

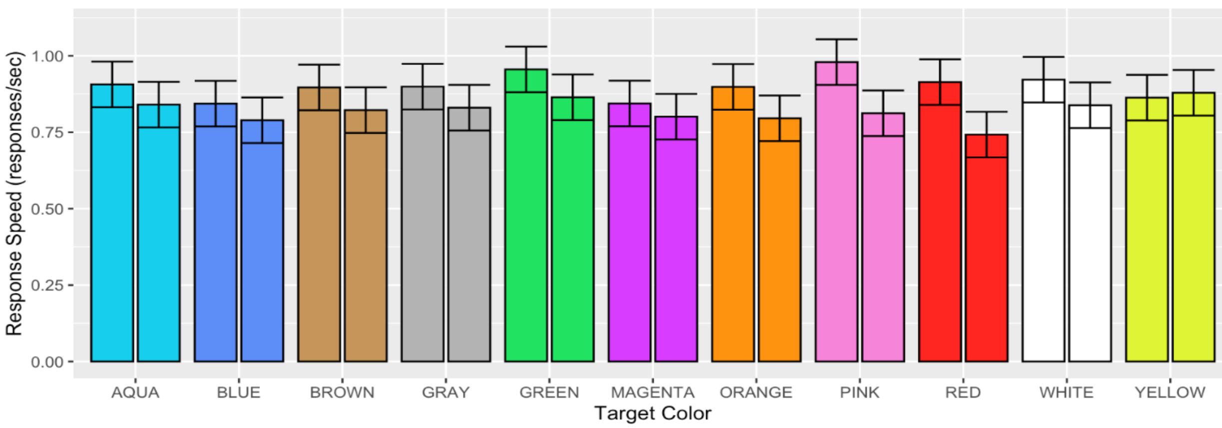

Figure 44. Target-Color x Color-Vision Status interaction for Shape-search Response Speed (p = 0.05, GES = 0.03)

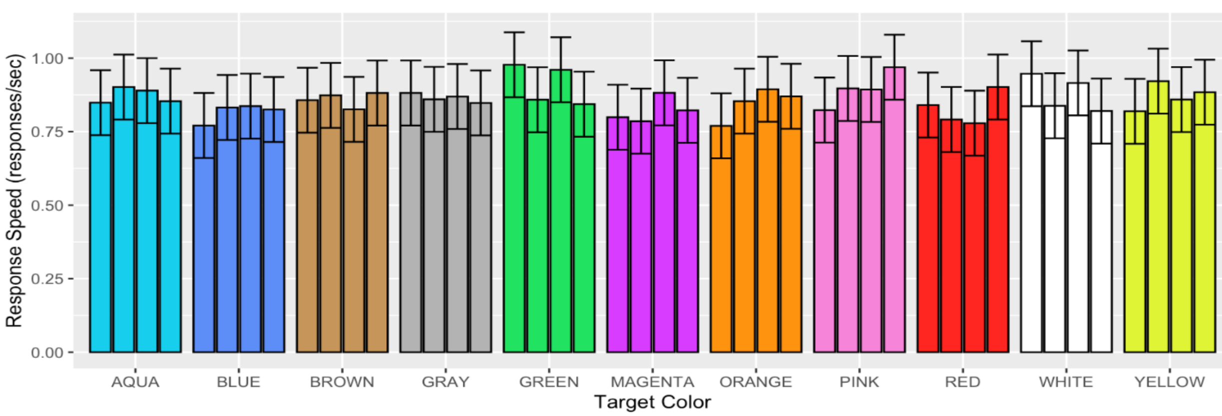

Figure 45. Target-Color x Backlight interaction for Shape-search Response Speed (p = 0.06, GES = 0.029)

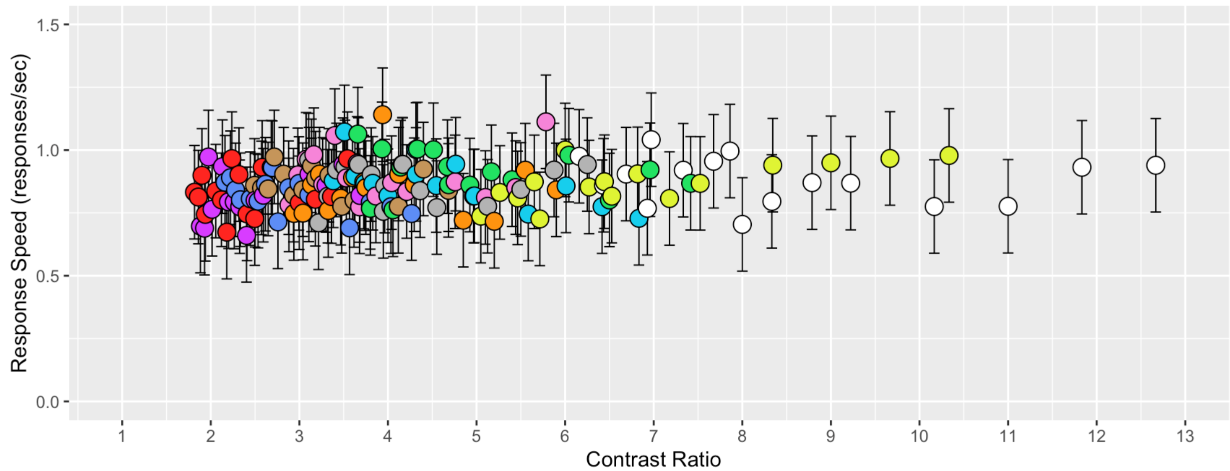

Figure 46. Shape-search Response Speed as a function of contrast ratio

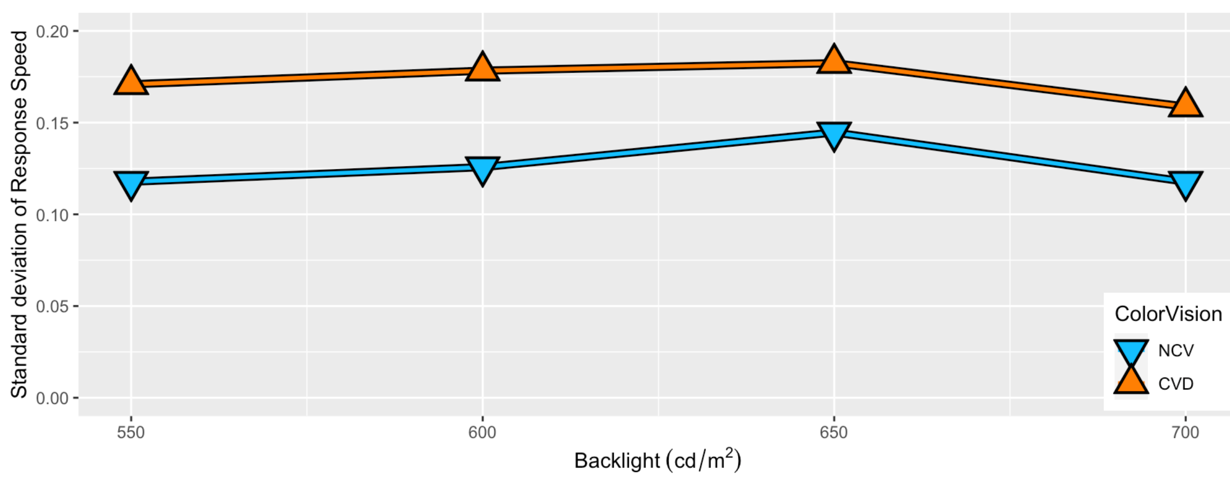

Figure 47. Standard deviation of Shape-search Response Speed as a function of peak D_{65} luminance

5 Discussion

The Weather search was clearly more difficult than the Shape search for NCVs and CVDs alike, making it the more sensitive test of the palette. This outcome matches expectations, considering the luminance-contrast ratios: The highest for the Weather search was 1.37, whereas the lowest for the Shape search was 1.82.

Figures 28 and 31 show that luminance contrast alone is a poor predictor of our participants' Weather search scores, though, because its effect depends on the target's color. This outcome implies the participants were not searching for luminance increments and then inferring the target's color from the (larger and easier to see) background color. In that case, the three curves in each figure would overlap. Our participants were looking for the colors, instead, and the differing luminance contrasts needed to make each color discernible can be understood ordinally by considering their chromatic differences from their backgrounds.

Table 8. Mixture of weather colors with illumination at 700 cd/m^2 peak D_{65} luminance

				Mixture		Distance from Background
Source	**u'**	**v'**	**Y (cd/m^2)**	**u'**	**v'**	
5600-K	0.1951	0.4802	60			
Wx-Red	0.26	0.4	14	0.2093	0.4627	**0.0348**
Wx-Yellow	0.23	0.54	26.8	0.205	0.4972	**0.0219**
Wx-Green	0.15	0.5	11.2	0.1882	0.4832	**0.0075**

The rightmost column of Table 9 shows the distance of each luminous weather color from its background color on the CIE 1976 u'v'-chromaticity diagram when they are displayed on a TDM under 55,000 lux of 5600-K illumination. (The peak D_{65} luminance is assumed to be 700 cd/m^2 in the table. At other luminances, the distances would scale proportionally.) Those distances predict that Wx-Red should be recognizable at a lower luminance contrast than Wx-Yellow, which should be recognizable at a lower luminance contrast than Wx-Green. This ordering matches our results and Havig et al.'s (2003). The probability of it occurring by chance is (1 / 3! =) 0.17.

Weather search Response Speed increased as luminance increased with no sign of leveling off at 700 cd/m^2. Wx-Green (on Black) yielded the slowest speeds at all luminances, although the difference ceases to be reliable statistically at 650 cd/m^2 according to the FLSD.

It surprises us that our participants' Shape-search performance shows no evidence that legibility suffered at contrast ratios below 3:1. FAA-HF-STD-010A (2020, Section 5.5) requires 3:1 for stimuli that must be legible. Even at lower contrasts, though, our Shape search data seem to reflect normal performance variations for stimuli that differ little in ways that affect that performance. Perhaps the freedom to vary the viewing distance enabled our participants to overcome contrasts that would have been too low at the 50-cm viewing distance FAA-HF-STD-010A (2020, Section 5.6.2) assumes.

We have shown that the FAA standard color palette's weather colors are recognizable and the foreground colors are legible on STARS TDMs under worst-case ambient illumination. To accomplish this, controllers will need to adjust the TDM's backlight setting sometimes and increase its peak D_{65} luminance to as much as 700 cd/m^2. (The new TDMs should have no difficulty doing this. Recall from Section 2.2.1 that the ones we measured are capable of 800 cd/m^2 peak D_{65} luminance.) The TDM's maximum with its backlight controller turned on is 600 cd/m^2, which should be adequate for ambient illuminances (A) up to

$$(600 \text{ cd/m}^2 * 0.032 + A * 0.00343 / \pi) / (A * 0.00343 / \pi) = 1.37 , \quad (1)$$

$$(600 \text{ cd/m}^2 * 0.032) / (A * 0.00343 / \pi) = 0.37 , \text{ and thus} \quad (2)$$

$$\boldsymbol{A} = (600 \text{ cd/m}^2 * 0.032) / (0.37 * 0.00343 / \pi) = \mathbf{47{,}529 \text{ lux}} , \quad (3)$$

where 0.00343 is the TDM screen's measured reflectance for our illuminant (see Section 2.2.2), 1.37 is the contrast ratio needed to obtain asymptotic performance for Wx-Green (see Figure 28), and 0.032 is Wx-Green's luminance factor from Table 2. Consequently, controllers at most towers should not need to exceed 600 cd/m^2 very often, if ever. Wilson et al.'s (2007) data suggest that the St. Louis Downtown Airport (CPS) and North Perry Airport in Hollywood FL (HWO) would be exceptions.

6 Limitations

- We have no data for people with tritan (i.e., yellow-blue) deficiencies.
- Our participants were not air traffic controllers.
- On the other hand, there is no reason to expect controllers to have either an advantage or disadvantage for the tasks we used.
- Furthermore, the color palette was new to our participants. Controllers would become familiar with it and error rates would, therefore, become even smaller – especially for color-deficient controllers.

7 Conclusions

Our results demonstrate that NCVs and viewers with mild-to-moderate color-vision deficiencies can perform as well with the FAA standard color palette under bright ambient illumination as they do under dim illumination if they increase the display's luminance sufficiently. For TDMs under worst-case illumination, a peak D_{65} luminance of 700 cd/m^2 should be enough. For lesser cases, the required TDM peak D_{65} luminance is

$$(L * 0.032 + A * 0.00343 / \pi) / (A * 0.00343 / \pi) = 1.37, \quad (4)$$

$$(L * 0.032) / (A * 0.00343 / \pi) = 0.37, \quad (5)$$

$$L = 0.37 * A * 0.00343 / \pi / 0.032\text{, and thus} \quad (6)$$

$$\boldsymbol{L = 0.0126 * A}\text{ ,} \quad (7)$$

where L = the required peak D_{65} luminance in cd/m^2 and A = the ambient illumination in lux.

For display screens that have a different reflectance, a substitution into Eqn. 6 suffices:

$$L = 0.37 * A * R / \pi / 0.032\text{, and so} \quad (8)$$

$$\boldsymbol{L = 3.68 * A * R}\text{ ,} \quad (9)$$

where R = the display screen's reflectance for the illuminant, measured at the user's eye position. Eqn. 9 can be used to predict the acceptability of candidate displays, given their reflectance and the expected ambient illumination.

8 Remaining Issues

Gildea et al. (2018) ensured that all colors in the FAA standard palette can be produced by the STARS, ERAM, and ATOP displays. This new project shows that the newest STARS TDMs can, too. FAA-HF-STD-010A (2020; section 5.16) includes display hardware specifications that ensure compatibility with the FAA palette. Future technical refreshes for FAA displays that must be compatible should incorporate those specifications.

Gildea et al. (2020) showed controllers dynamic depictions of STARS and ERAM images with the FAA standard colors substituted for the current ones. The controllers reported no problems and noted that the dark, desaturated weather colors provide a welcome increase in foreground-color contrast. The next step is interactive simulation, which should expose problems with implementation or user acceptance, if there are any. The results from that project will appear soon.

The next question to be addressed for the palette is implementation-related. The FAA will need a way to ensure that the colors that appear on the displays match the ones in Tables 1 and 2. Creating tables of RGB values that are customized for each display, as we did for the project, is impractical for the numerous displays in the field. It will be necessary to develop a method to calibrate the displays, so they respond the same way to RGB values. Calibration will ensure that one set of RGB values produces the same colors on them all. The method cannot require measuring the displays with expensive instruments like our spectroradiometer and the procedure must be straightforward for technicians to learn.

References

Chidester, T., Milburn, N., Lomangino, N., Baxter, N., Hughes, S., & Peterson, L. (2011). *Development, validation, and deployment of an occupational test of color vision for air traffic control specialists* (Report No. DOT/FAA/AM-11/8). Washington, DC: Federal Aviation Administration.

Chidester, T., Milburn N., Peterson, L. S., Gildea, K. Roberts, C., & Perry, D. (2013). *Development, validation, and deployment of a revised air traffic color vision test: Incorporating advanced technologies and oceanic procedures and en route automation modernization systems* (Report No. DOT/FAA/AM-13/15). Washington, DC: Federal Aviation Administration.

FAA *Guide for Aviation Medical Examiners* (2022). https://www.faa.gov/about/office_org/headquarters_offices/avs/offices/aam/ame/guide/media/guide.pdf

FAA-HF-STD-010A (2020). *Color use in air traffic control displays.* Washington, DC: US Federal Aviation Administration.

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*, 175-191. DOI: 10.3758/bf03193146

Gildea, K. M., Milburn, N., & Post, D. L. (2018). *Development of a standard palette for color coding ATC displays.* (Report No. DOT/FAA/AM-18/22). Washington DC: Office of Aerospace Medicine.

Gildea, K., Willems, B., Benincasa, S. Jack, D. G., & Post, D. L. (2020). *Evaluation of a new color palette for ATC displays.* (Report No. DOT/FAA/AM-20/08). Washington DC: Office of Aerospace Medicine.

Havig, P. R., Martinsen, G. L., Post, D. L., Reis, G. A., and Heft, E. L. (2003). Chromaticity and luminance requirements for colored symbology in night vision goggles. In Rash, C.E. and Reese, C.E. (Eds.), *Proceedings of SPIE: Helmet- and Head-Mounted Displays VIII, 5079*, 361-369.

Lit, A., Young, R. H., & Shaffer, M. (1971). Simple reaction time as a function of luminance for various wavelengths. *Perception and Psychophysics, 10,* 397-483.

Neitz, J., Summerfelt, P., & Neitz, M. (2023, May 31). Neitz test of color vision. http://www.neitzvision.com/img/neitztest/Neitztestad.pdf

Post, D. L., & Calhoun, C. S. (1989). An evaluation of methods for producing desired colors on CRT monitors. *Color Research and Application, 14,* 172-186. DOI:10.1002/col.5080140406

Post, D. L., & Calhoun, C. S. (2000). Further evaluation of methods for producing desired colors on CRT monitors. *Color Research and Application, 25,* 90-104. https://doi.org/10.1002/(SICI)1520-6378(200004)25:2<90::AID-COL4>3.0.CO;2-S

Smart Optometry D.O.O. (2022). Smart optometry eye tests for professionals, version 4.6. [Mobile application software]. Apple App Store. https://apps.apple.com/us/app/smart-optometry/id1053891266

Weibull, W. (1951). A statistical distribution function of wide applicability. *Journal of Applied Mechanics, 18,* 293-297.

Wilson, E., Wilson, D., & Jha, P. D. (2007). *Airport traffic control tower lighting and viewing measurements.* (Report No. DOT/FAA/TC-07/09). Washington, DC: Federal Aviation Administration.

A Informed consent statement

I, ____________________________, understand that this study, entitled "Adapting the FAA-HF-STD-010A Standard Color Palette to Daytime Illumination" is sponsored by the Federal Aviation Administration's (FAA's) William J. Hughes Technical Center (WJHTC) in Atlantic City, NJ. It is under the direct supervision of Dr. Randy Sollenberger of the WHJTC Human Factors Branch.

1. Purpose. The objective of this study is to assess a proposed color palette for air traffic control displays under bright lighting and ensure that it is usable by individuals with normal color vision and those with some types and degrees of color vision deficiency. Approximately 100 participants will be involved.

2. Procedures. I understand that the data-collection protocols used in this study involve tasks that require responses by computer mouse, keystroke, or speaking. The data gathered during the study will be used solely for analyses. In no way will my name be associated with the coded subject identification number for any presentation or publication. Research results will be reported as group data.

3. Description of study requirements. There are various types of color vision deficiencies and various degrees of those deficiencies. For this study, we need a certain number of each type and degree of color vision deficiency and a certain number of participants with normal color vision. They must all have good visual acuity.

I understand that, on my first visit, I will be asked to complete a visual acuity screening (10 minutes) to ensure that I have 20/40 or better near visual acuity, corrected or uncorrected. If I do not meet those criteria, I will be ineligible to participate in the study. Otherwise, I will be asked to complete a simple test for color vision deficiencies (10 minutes), which involves spotting digits in fields of colored dots. If I exhibit a deficiency, I will be asked to take a more elaborate color vision test (60 minutes). My first visit will therefore require 10 to 80 minutes or so.

I understand that, if my visual acuity meets the criteria, I will be asked to return several times. Each time, I will perform four visual search tasks under different room lighting conditions that will require no more than 2 hours to complete.

As a person who is in reasonably good health, I know of no medical or other conditions that would prevent me from operating a computer workstation attentively for up to 120 minutes in one sitting. I do not have any prior experience of seizure due to exposure to flicking lights, moving lights, flashing displays, etc. I do not have diabetes that is controlled by medicine (because the medicine may affect one's color vision).

I understand that:

a. I must perform to the best of my ability on all the tasks to allow accurate assessment of my visual performance, including accuracy of color identification and response time.
b. It is essential that I am attentive and follow the instructions during the course of this study.
c. Completing all the tasks will take up to 2 hours for each visit and I will be permitted a short break if I request it.
d. I will be asked to work using a computer display in a dim or brightly lit room.
e. If I have any questions, I may ask the experimenter.
f. I agree to report any injury or suspected adverse effects from this study immediately to Dr. Randy Sollenberger at (609) 485-7169.

4. Possible risks. The risks involved in this study are minimal. Sitting for approximately 120 minutes at a time performing the computer-based tasks may produce some fatigue. The overall fatigue experienced from a visit to the laboratory should be similar to that experienced during 2 hours of intensive work at a computer.

5. Benefits. The major benefit from this study for the research community will be a better understanding of how to adapt a color palette that is usable under dim lighting by people with normal color vision and those with some types of color vision deficiencies so it works well in bright lighting, too. This study will also aid in developing new methods of assessing the color vision of air traffic controller applicants. I understand that my participation is voluntary and I benefit by acquiring an understanding of my color vision. I understand that I will be paid for my time participating in this study.

6. Conditions of participation. I understand that I am free to withdraw my consent and to withdraw from the project at any time. I understand that my participation in this study may be terminated at any time at the experimenter's discretion. Examples of reasons for termination would include an intentional lack of cooperation or disregard of the experimental equipment or procedures.

7. Assurances. I understand that:

a. Dr. Randy Sollenberger and his associates will take every precaution to utilize proven research procedures.
b. If any new findings develop during the course of this study that may relate to my decision to continue participation, I will be informed.
c. By giving my consent, I have not waived any legal rights or released WJHTC or any individual from liability for negligence. I may revoke my consent and withdraw from the study at any time.
d. The results will be treated as confidential and will receive a code number so that they will be anonymous when filed with the Laboratory Manager. In no case will any use be made of these results other than the application of experimental analyses unless I provide explicit written permission.
e. I understand that Dr. Randy Sollenberger or his colleagues will be available to answer

my questions concerning procedures throughout this study.

8. Qualification for participating in the study. I understand the following requirements for qualifying me to participate this study:

a. Required age: 18 - 55 years.
   This requirement represents the eligible age range to work as an air traffic controller. The FAA requires that a controller applicant holds a high school diploma plus three years of post-high school experience. This results in the age of the youngest controllers in the current work force being roughly 20 years old. The official retirement age for controllers is 56 years.
b. Required gender: Both men and women will be allowed to participate.
c. Required health conditions: I do not have any prior experience of the following medical conditions:
   i. Seizure induced by optical stimuli due to exposure to flicking lights, moving lights, flashing displays, etc.
   ii. Diabetes that is controlled by medicine, as the medicine may affect my color vision.
d. I have no health condition that prevents me from working intensively for 2 hours in a day and I understand that some tasks will be administered in a dimly-lit room and others will be in a brightly lit one.
e. **Compensation / Injury (FOR CONTRACT SUBJECTS).** By signing, I understand that the contract company that recruited me is responsible for the compensation of my participation in this study and potential injury during the study.

9. Signatures. By signing, I understand that I am free to withdraw this consent and discontinue participation at any time, for any reason. If I do withdraw, I will be paid for my time worked.

I have read this consent document. I understand its contents and I consent freely to participate in this study under the conditions described here. I will receive a copy of this form upon request.

______________________________	______________________________
Investigator's Signature	Participant's Signature
______________________________	______________________________
Participant's Signature Witnessed by	Date