Proceedings of the Human Factors and Ergonomics Society Annual Meeting
2023, Vol. 67(1) 650–656

DOI: 10.1177/21695067231192615
journals.sagepub.com/home/pro
Sage

Effects of Color Commonality of Overlay Clutter and Information Access Effort on Tasks Requiring Visual Search

Amelia C. Warden[1], Christopher D. Wickens[1], Benjamin A. Clegg[1], and Francisco R. Ortega[2]

Abstract

The amount and color of overlay clutter can impede focusing attention on one layer of information from multiple sources, and influence dividing attention when comparing across layers. The current experiment examined the effect of information access effort and color commonality of overlay clutter on performance. Participants viewed maps with two domains of information that were either overlaid, adjacent, or separated, and answered questions about either both domains (integration) or one domain (focused attention). The overlaid information was either similar (green) or dissimilar (red) in color relative to the background. Overlaid displays benefited integration tasks but imposed a cost to focused attention tasks for accuracy but not response time. Increased display separation did not impose performance costs. Computational models account for some costs of clutter in overlay displays but the color similarity between databases also contributes to cost and benefits of overlay in a fashion not currently considered by such models.

Keywords

Information Access Effort, Clutter, Attention, Visual Search, Color, Information Processing, Working Memory, Computational Modeling

Introduction

Data-rich displays are prevalent in professions that require using geospatial maps, such as aviation, weather forecasting, or search-and-rescue operations. Such displays often require visual search, whether it is searching for a specific landmark or deciding what direction a storm is moving. Visual search is the process of looking for a specific target among a set of distractors that may vary in similarity relative to the target (Sternberg, 1969; Wolfe & Horowitz, 2017; Drury, 1990). Such searches are directly impacted by the number of items in the display competing for attentional resources. The search time to find a target increases linearly as the number of items increases, consistent with a serial search model (Sternberg, 1969; Kroft & Wickens, 2003; Wolfe & Horowitz, 2017). Other factors that impact visual search performance include the attentional requirements of the specific task, display format, display clutter, and the color similarity of objects in the search field (Wickens et al., 2022; Nagy & Sanchez, 1990; Wolfe & Horowitz, 2017).

Additional map tasks that require visual search, but different forms of attention include information readout and information integration. Information readout tasks require searching for and finding a target to *extract specific information* (e.g., mountain elevation) via *focused attention*. Integration tasks require *dividing attention* between two or more sources of information within a display to combine the information, such as whether the altitude of an aircraft is higher than the elevation of a nearby mountain. Depending on the attentional demands of a task, there is a tradeoff between the amount of scanning required to access information and the clutter in the display.

According to the *proximity compatibility principle* (PCP, Wickens & Carswell, 1995), two or more sources of information should be in close display proximity (e.g., overlaid) when they require mental integration. For example, head-up displays overlay information onto the real world creating close display proximity. Information requiring focused attention can be presented with decreased display proximity (e.g., separated), where displays of information are adjacent or separated (e.g., head-down display or a tablet).

Overlaying domains of information from different geospatial databases (e.g., weather and terrain) decreases display

---

[1] Departments of Psychology, Colorado State University. Fort Collins, CO, USA
[2] Departments of Computer Science, Colorado State University, Fort Collins, CO, USA

Corresponding Author:
Amelia C. Warden, Departments of Psychology, Colorado State University, 1876 Campus Delivery, Fort Collins, CO 80523-1019, USA.
Email: amelia.warden@colostate.edu

proximity but also increases the amount of overlay clutter. Overlay clutter refers to the superimposition of multiple objects or several databases of information onto the same display, thereby obscuring information in one domain with information in another. Clutter can be quantified by the number or density of objects in a display, and directly impacts overall search time (Moacdieh & Sarter, 2015). Consistent with the proximity compatibility principle, overlaid displays hurt performance for tasks requiring focused attention due to the costs of overlay clutter but benefit performance for tasks requiring integration (Jang et al., 2012; Kroft & Wickens, 2003; Fadden et al., 2000). While overlaying displays increases the amount of clutter, it also minimizes the amount of information access effort required when two sources of information need to be integrated. This is referred to as the *scan-clutter* tradeoff.

While several factors can influence information access effort (IAE), in the current context we operationally define IAE as the visual angle of separation between information in two domains: a view of the geographical terrain and a separate schematic map. There are some performance costs associated with increased visual angle separation between sources of information (Draschkow et al., 2021; Murata et al., 2018; Warden et al, 2022) but these costs may be small depending on the display and whether the head-field is invoked (a situation in which neck rotation is required to augment visual scanning; Kim et al., 2010, Warden et al., 2022).

In the current study, we examine IAE as it is increased from overlay to a condition where the two domains are adjacent, and as it is increased from adjacent to further separated. In the first comparison, the displays are confounded with the removal of overlay clutter, whereas the second comparison only looks at IAE unconfounded with clutter.

Another way to increase display proximity for tasks is to use similar (high proximity) rather than dissimilar (low proximity) color coding. In the case of cluttered displays, color acts as an attentional filter to help guide search tasks (Yeh & Wickens, 2001; Remington et al., 2001; Wolfe & Horowitz, 2017). Prior work suggests that similar color proximity disrupts focused attention tasks but assists integration tasks (Yeh & Wickens, 2001). Summarizing visual search studies that have examined the similarity between targets and distractors reveals that: dissimilar colors create a pop-out effect, capturing attention faster and decreasing search time for focused attention tasks regardless of clutter level; and color similarity imposes a cost to focused attention tasks but facilitates mental integration because the searcher can filter out irrelevant information via grouping (Wickens & Andre, 1990; Nagy & Sanchez, 1990).

Color similarity is also a factor to consider when quantifying display clutter. The Rosenholtz (2007) feature congestion (FC) clutter metric quantifies clutter based on features such as orientation, color, and luminance contrast. This model suggests that adding more items to the display, such as more diverse colors, increases clutter by reducing room in feature space to add more salient items that may guide attention. However, clutter metrics fail to account for the effects of overlaying clutter onto background information that is similar in color.

In the current experiment, we examined the cost of increasing display clutter and spatial separation for a map task consisting of focused attention and integration questions. Participants viewed a series of maps containing information about two different domains: the far domain (ground domain) consisting of geographical terrain information, and the near domain (air domain) consisting of flight path and weather information. The amount of clutter in each domain varied and was determined by two metrics: the Rosenholtz (2007) feature congestion (FC) metric and an object count metric. The domains of information were displayed as either directly overlaid, adjacent, or separated.

We hypothesized that: (H_1) overlay displays will increase response time more than adjacent displays due to the increased search time imposed by overlay clutter; (H_2) increasing display separation will impose a performance cost for focused attention and integration tasks, but this cost will be greater for the integration task with its working memory demands; and (H_3) color similarity will impose a cost to the focused attention task due to confusion and color dissimilarity will impose a cost to the integration task due to the more distant display proximity.

Method

Participants

Fifty-one students enrolled in an introductory psychology course at Colorado State University received course credit for completing the experiment. All participants had normal or correct-to-normal vision and were screened for colorblindness. The experiment was approved by the Institutional Review Board.

Stimuli and Apparatus

All stimuli were presented on a BenQ IPS 27-inch monitor with a screen resolution of 1920 x 1080 pixels and a 60 Hz refresh rate. All map stimuli were created using CalTopo, an online application used to create topographic maps. The ground terrain (far domain) imagery was generated from hybrid maps and included a mountain icon with identification and elevation data, and contour lines. An important feature of this domain is that the overall color was predominately shades of green (see Figure 1). The air domain (near domain) imagery consisted of aircraft icons with aircraft identification and altitude information, flight paths, and weather icons (i.e., warm and cold fronts, rain, wind). Near domain icons, except for the warm and cold fronts, were either green or red. Thus, near domain imagery was either similar or different,

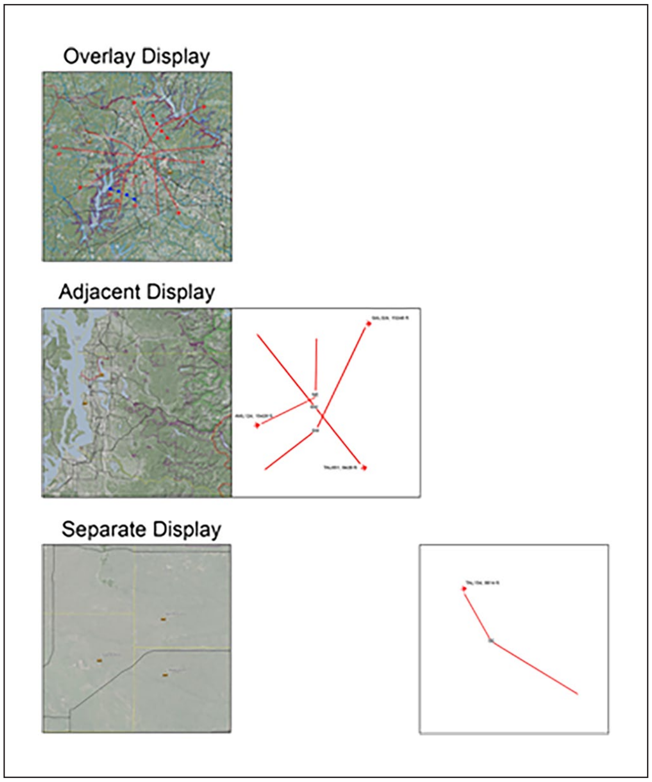

Figure 1. An illustration of each display condition for a single trial. The overlay display (top) shows high far and high near domain clutter. The adjacent display (middle) shows medium far and medium near domain clutter. The separate display (bottom) shows low far and low near domain clutter. All possible clutter combinations were presented. Near domain clutter was either red (shown here) or green.

respectively, from the dominant color of the far domain. For the overlay condition, the near domain was superimposed onto the far domain. For the adjacent and separate conditions, the near and far domains were either side-by-side with 10° or 32° of separation, respectively, measured from the center of each image. All images were 1000 x 1000 pixels. For the separate condition, participants sat approximately 73 centimeters away from the display to ensure a visual angle of approximately 32°.

The magnitude of feature congestion (FC) clutter values for each domain and the total FC value for overlay displays were verified using the Rosenholtz's (2007) computational clutter model. For the integration task, questions required integrating across both domains. For example, *"Is the altitude of aircraft UAL232 higher than the elevation of Black Mountain?"* For the focused attention task, each question required focusing on only one domain. For example, *"Is the flight path of aircraft UAL232 heading SW?"* requires focusing attention on the near domain only. Questions were presented adjacent to the images for overlay and adjacent displays. For separate displays, questions were presented above the images and horizontally centered on the screen.

Design

Participants completed six blocks that were counterbalanced based on task (focused attention and integration) and display condition (overlaid, adjacent and separated). Each block consisted of 36 randomized experimental trials for a total of 216 experimental trials. For the adjacent and separate display conditions, half of the trials asked questions about the near domain and the remaining half asked questions about the far domain. Integration questions for all three display conditions pertained to both domains. The entire experiment lasted approximately 40 minutes.

Procedure

Participants gave consent to participate after reviewing the consent documentation. Before the experiment, participants received instructions about each task and were trained to read the maps and icons. Then they completed four practice trials with feedback for each display condition before the test trials. Participants were asked to maintain proper posture during the experiment. Participants were instructed to answer yes/no questions as accurately and rapidly as possible by making a key press ('Y' for yes and 'N' for no).

Results

One participant was removed from the data based on Grubbs' and Rosner's (1983) tests for outliers. Five additional participants were excluded from the data because they did not finish the experiment. The data in all analyses were collapsed across each type of focused attention task.

Information Access Effort: Effect of Display Separation

Using two separate 2 (display condition) x 2 (task type) repeated measures ANOVAs, we examined whether performance differed with spatial separation (overlay *vs* adjacent, and adjacent *vs* separate) as a function of task (focused attention, and integration). For the sake of space, we include all display conditions in the figures even though two separate ANOVAs were used to analyze the effects of clutter and the effects of information access effort unconfounded by clutter.

Response Time. The first ANOVA examined H_1: whether overlay displays would increase response time more than adjacent displays. As shown in Figure 2, the ANOVA (overlay *vs* adjacent, the 4 left points in the figure) revealed that response time was significantly slower for overlay (9.21 s) than adjacent (8.24 s) displays, $F(1, 44) = 20.63, p < .001, \eta_p^2 = 0.32$. Response time was significantly slower for integration (10.12 s) than focused attention (7.33 s) questions, $F(1, 44) = 368.70, p < .001, \eta_p^2 = 0.89$. The interaction was

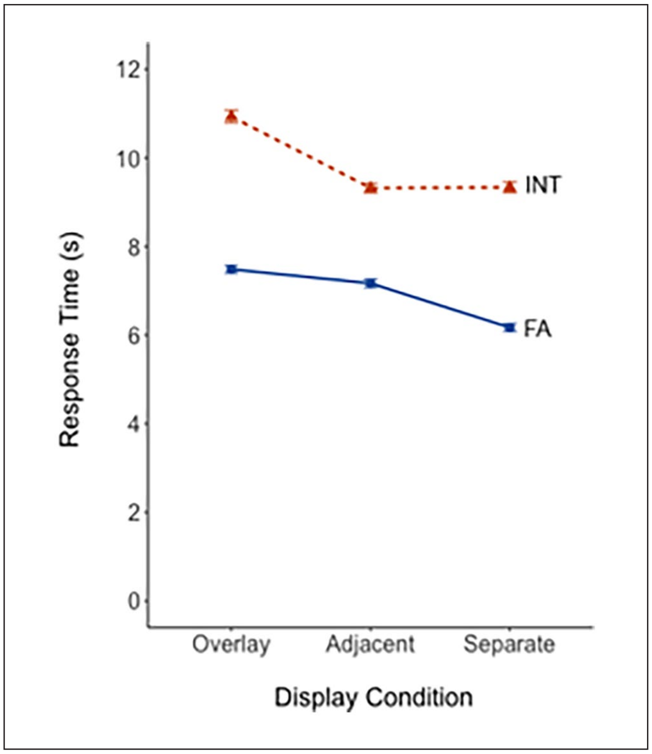

Figure 2. Mean response time (seconds) plotted as a function of display condition and question type. Solid blue and dashed orange lines represent integration and focused attention questions, respectively. Error bars represent one standard error of the mean.

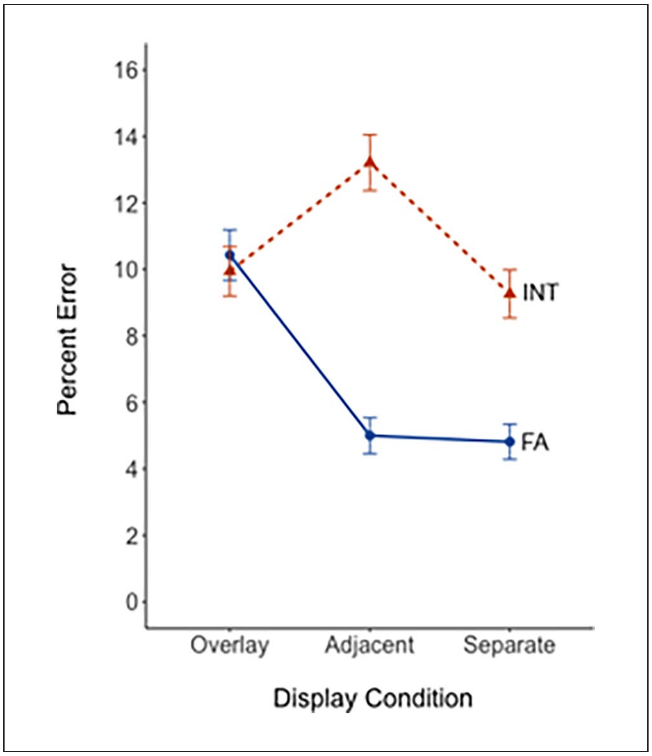

Figure 3. Mean percent error plotted as a function of display condition and question type. Solid blue and dashed orange lines represent integration and focused attention questions, respectively. Error bars represent one standard error of the mean.

significant, $F(1, 44) = 27.46, p < .001, \eta_p^2 = 0.38$. Planned comparisons show that, for integration questions (dashed line), overlay displays slowed response time more than adjacent ($t(44) = 5.51, p < .001, d = 0.73$), but there were no differences in response time for the focused attention questions ($p = .10$).

The second ANOVA (adjacent *vs* separate displays, the right 4 points) examined H_2: that increasing IAE will impose a cost to both tasks, but the costs will be greatest for the integration task. The ANOVA revealed a main effect of display condition on response time, $F(1, 44) = 5.12, p = .03, \eta_p^2 = 0.10$. There was a main effect of question type, $F(1, 44) = 367.60, p < .001, \eta_p^2 = 0.89$. Both effects are best interpreted in the context of the significant interaction, $F(1, 44) = 31.91, p < .001, \eta_p^2 = 0.42$. For focused attention questions, adjacent displays slowed response time compared to more separated displays ($t(44) = 5.14, p < .001, d = 0.66$) but for integration questions, display separation had no effect ($p = .95$).

Percent Error. As depicted in Figure 3, the first ANOVA (overlay *vs* adjacent, left four points) revealed that percent error was significantly greater for the overlay (10%) than the adjacent (9%) displays, $F(1, 44) = 4.31, p = .04, \eta_p^2 = 0.09$. Percent error for integration questions (12%) was significantly greater than focused attention questions (8%), $F(1, 44) = 40.44, p < .001, \eta_p^2 = 0.48$. The interaction was significant, $F(1, 44) = 54.73, p < .001, \eta_p^2 = 0.55$. Planned comparisons showed that, for integration questions (dashed line), overlay displays helped performance, $t(44) = -3.88, p < .001, d = -0.72$, but for focused attention, overlay hurt performance by a large amount, $t(44) = 7.52, p < .001, d = 1.29$.

The second ANOVA (adjacent *vs* separate) testing H_2 revealed a main effect of display condition on percent error, $F(1, 44) = 8.79, p = .005, \eta_p^2 = 0.17$. There was a main effect of question type, $F(1, 44) = 132.90, p < .001, \eta_p^2 = 0.75$. The main effects are better interpreted in the context of the significant interaction, $F(1, 44) = 8.74, p = .005, \eta_p^2 = 0.17$. For the integration task, increasing spatial separation of the displays helped performance by reducing error ($t(44) = 4.44, p < .001, d = 0.92$), but for focused attention questions there was no difference ($p = .85$).

Effect of Overlay Clutter and Color

We used linear mixed models and binomial generalized linear mixed models to examine the effect that feature congestion

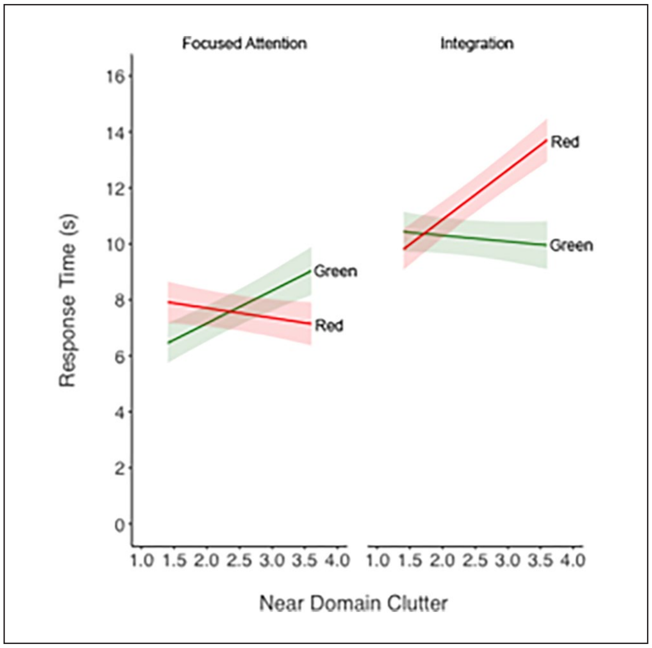

Figure 4. Response time (seconds) as a function of near domain clutter, question type, and color predicted from the model for the overlay display condition. Near domain clutter is based on the feature congestion metric. Shading represents 95% CIs as calculated from the model.

clutter had on performance only for the overlay displays as a function of question type (focused attention and integration questions) and near domain color (green, red). These analyses examine H_3: the effects of color similarity and dissimilarity on performance for each task.

Response Time. As shown in Figure 4, response time significantly increased as near domain clutter increased, estimate = 1.17 s, $SE = 0.22$, t = 5.40, $p = .02$, $d = 0.25$.

Response time was significantly faster with green (8.8 s) than red (9.6 s) near domain information, estimate = 3.59, $SE = 0.71$, $t = 5.08$, $p < .001$, $d = 0.77$. There was a significant three-way interaction between near domain clutter, question type, and color, estimate = 3.52, $SE = 0.41$, $t = 8.54$, $p < .001$, $d = 0.76$. Contrasts revealed that, for focused attention (left panel), increasing green overlay clutter degrades focused attention (slope = 0.477, $SE = 0.152$, 95% CIs [0.18, 0.78]), whereas increasing red overlay clutter does not (slope = 0.71, $SE = 0.14$, 95% CIs [0.44, 0.98]). However, for the integration task (right panel), as near domain clutter increases, red overlaid information increases response time more than green overlaid information, estimate = -1.20, $t = 0\text{-}5.65$, $p < .0001$).

Percent Error. The data depicted in Figure 5, reveals that there was a significant effect of near domain clutter on percent error, estimate = 0.48, $SE = 0.22$, $t = 2.18$, $p = .03$. Also, errors were significantly more frequent with red (more discriminable) than green (more confusable) color on the display, estimate = 5.05, $SE = 0.68$, $z = 7.48$, $p < .001$. Critically, there was a significant three-way interaction between near domain clutter, question type, and color, estimate = 2.28, $SE = 0.40$, $z = 5.77$, $p < .001$. For focused attention questions (left panel), as near domain clutter increased, error increased when information was green but decreased when information was red, estimate = -0.99, $SE = 0.20$, $z = -5.07$, $p < .001$. In sharp contrast, for integration questions (right panel), error increased as near domain clutter increased when information was red, ($t = -1.20$, $SE = 0.20$, $z = -6.01$, $p < .001$) but not when it was green (confusable).

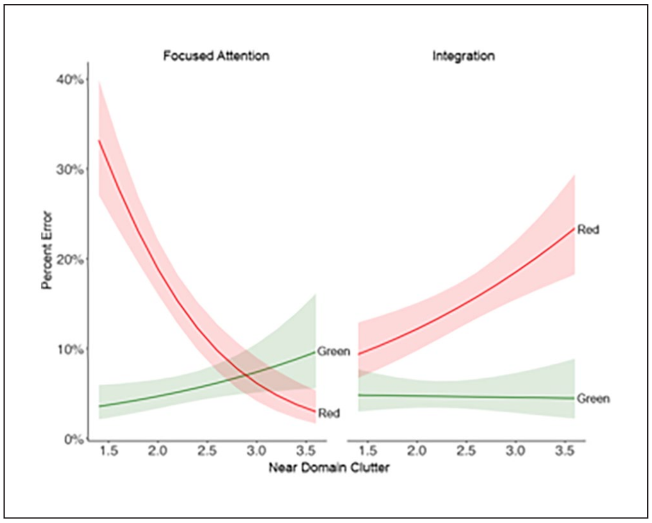

Figure 5. Percent error as a function of near domain clutter, question type, and color predicted from the model for the overlay display condition. Near domain clutter is based on the feature congestion metric. Shading represents 95% CIs as calculated from the model.

Discussion

The current experiment examined the effect of increasing information access effort and overlay clutter varying in color similarity relative to the background on tasks requiring focused attention or integration. When comparing overlay vs adjacent displays (i.e., the scan-clutter tradeoff), the proximity compatibility principle (PCP) predicts an interaction. In the current work, overlay relative to adjacent displays were found to improve performance (or hurt performance less) for integration tasks than for focused attention tasks.

For response time, overlay was generally costly for both tasks (i.e., clutter hurts more than spatial proximity helps), but overlay increased response time more for the integration task than the focused attention task. However, reflecting a sort of speed-accuracy tradeoff, overlay helped accuracy (i.e., reducing error) for the integration task, but harmed it for the focused attention task. Thus, accuracy supported predictions of the PCP; but response time did not.

Our second hypothesis predicted that the costs of increasing display separation will be greater for the integration task than the focused attention task. This hypothesis was only partially supported. For the integration task, increasing separation did not impact response time and improved accuracy. For focused attention, increasing separation slightly shortened response time but did not effect accuracy.

Our third hypothesis predicted that color similarity would impose a cost to the focused attention task and color dissimilarity would impose a cost to the integration task. This hypothesis was partially supported. For the focused attention tasks, which required pure visual search, when more green objects are overlaid onto the green far domain, search time slows down consistent with Nagy & Sanchez (1990). Conversely, more red objects that contrast with the far domain actually speeds the search and greatly increases accuracy.

However, for the integration task, an unexpected opposite effect was revealed. Increasing near domain clutter in the color similar (green) to the far domain had little effect on performance, while increasing near domain clutter in the color that was different (red) from the far domain actually hurt both response time and accuracy (Figure 5). This effect may be accounted for by another dimension of display proximity besides space; that is, *color proximity* (Wickens et al., 2022). Common color between two relevant items can improve integration of the two just as common color hinders focused attention on one while ignoring the other.

Based on the current findings, color similarity plays an important role in predicting the effects of overlay clutter that are not accounted for by current computational models, like the FC metric (Rosenholtz, 2007). To better capture display clutter and its effects on performance, in overlay conditions, such models should incorporate color commonality as a penalty for focused attention tasks and a reward for integration tasks, and should include the effects of overlay clutter.

Conclusions

In this experiment we used a realistic military scenario to examine the scan-clutter tradeoff in the context of a design decision to overlay geospatial databases or present them separately. The proximity compatibility principle recommends that overlay serves information integration tasks between the databases, but not focused attention on each. This relation was confirmed for task accuracy, but not speed. Further separation of displays produced only muted costs to speed and accuracy.

We also found that the quantification of clutter by computational metrics could account for some of the costs of increasing clutter when there was overlay, but that the similarity of color between databases also accounted for both benefits of overlay (for task integration) and costs (for focused attention) in a way that was not accounted for by the computational model. While our study was conducted on a 2D flat panel display, the results should generalize to use of a head-mounted display.

Acknowledgments

This research was supported by the Office of Naval Research under grant numbers N00014-21-1-2949 and N00014-21-1-2580. Dr. Peter Squire was the scientific/technical monitor.

References

Draschkow, D., Kallmayer, M., & Nobre, A. C. (2021). When Natural Behavior Engages Working Memory. *Current Biology*, *31*(4), 869-874.e5. https://doi.org/10.1016/j.cub.2020.11.013

Drury, C. G. (1990). Visual search in industrial inspection. *Visual search*, 263-276.

Fadden, S., Wickens, C. D., & Ververs, P. (2000). Costs and Benefits of Head up Displays: An Attentional Perspective and a Meta Analysis. SAE Technical Paper 2000-01-5542. https://doi.org/10.4271/2000-01-5542

Jang, J., Trickett, S., Schunn, C., & Trafton, G. (2012) Unpacking the temporal advantage of distributing complex visual displays. *International Journal of human Computer Studies.70*. 812-827.

Kim, K. H., Reed, M. P., & Martin, B. J. (2010). A model of head movement contribution for gaze transitions. *Ergonomics*, *53*(4), 447–457. https://doi.org/10.1080/00140130903483713

Kroft, P., & Wickens, C. D. (2003). Displaying multi-domain graphical database information. *Information Design Journal*, *11*(1), 44–52. https://doi.org/10.1075/idj.11.1.06kro

Nagy, A. L., & Sanchez, R. R. (1990). Critical color differences determined with a visual search task. *JOSA A*, *7*(7), 1209-1217

Moacdieh, N., & Sarter, N. (2015). Display clutter: A review of definitions and measurement techniques. *Human factors*, *57*(1), 61-100

Murata, A., & Kohno, Y. (2018). Effectiveness of replacement of automotive side mirrors by in-vehicle LCD–Effect of location and size of LCD on safety and efficiency–. *International Journal of Industrial Ergonomics*, *66*, 177–186. https://doi.org/10.1016/j.ergon.2018.03.010

Remington*, et al.* (2001): Rosenholtz, R., Li, Y., & Nakano, L. (2007). Measuring visual clutter. *Journal of vision*, *7*(2), 17-17. https://doi.org/10.1167/7.2.17

Rosenholtz, R., Li, Y., & Nakano, L. (2007). Measuring visual clutter. *Journal of vision*, *7*(2), 1-22. https://doi.org/10.1167/7.2.17

Rosner, B. 1983. Percentage points for a generalized ESD many-outlier procedure. *Technometrics 25*:165-172.

Sternberg, S. (1969). Memory-scanning: Mental processes revealed by reaction-time experiments. *American scientist*, *57*(4), 421-457.

Warden, A. C., Wickens, C. D., Rehberg, D., Clegg, B. A., & Ortega, F. R. (2022). Information Access Effort: Are Head Movements "Cheap" or Even "Free"?. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *66*(1), 2203-2207. https://doi.org/10.1177/1071181322661127

Wickens, C. D., & Carswell, C. M. (1995). The proximity compatibility principle: Its psychological foundation and relevance to display design. *Human Factors*, *37*(3), 473-494. https://doi.org/10.1518/001872095779049408

Wickens, C. D., & Andre, A. D. (1990). Proximity compatibility and information display: Effects of color, space, and objectness on information integration. *Human factors*, *32*(1), 61-77.

Wickens, C.D., McCarley, J., & Gutzwiller, R. (2022). *Applied Attention Theory* (2nd Ed). Taylor & Francis: Rutledge.

Wolfe, J. M., & Horowitz, T. S. (2017). Five factors that guide attention in visual search. *Nature Human Behaviour*, *1*(3), 0058.

Yeh, M., & Wickens, C. D. (2001). Attentional filtering in the design of electronic map displays: A comparison of color coding, intensity coding, and decluttering techniques. *Human Factors*. *43*(4), 543–562. https://doi.org/10.1518/001872001775870359