ERGONOMICS, 2018
VOL. 61, NO. 5, 682–696
https://doi.org/10.1080/00140139.2017.1401669

An integrated measure of display clutter based on feature content, user knowledge and attention allocation factors

Carl Pankok Jr. and David B. Kaber

Industrial & Systems Engineering, North Carolina State University, Raleigh, NC, USA

ABSTRACT
Existing measures of display clutter in the literature generally exhibit weak correlations with task performance, which limits their utility in safety-critical domains. A literature review led to formulation of an integrated display data- and user knowledge-driven measure of display clutter. A driving simulation experiment was conducted in which participants were asked to search 'high' and 'low' clutter displays for navigation information. Data-driven measures and subjective perceptions of clutter were collected along with patterns of visual attention allocation and driving performance responses during time periods in which participants searched the navigation display for information. The new integrated measure was more strongly correlated with driving performance than other, previously developed measures of clutter, particularly in the case of low-clutter displays. Integrating display data and user knowledge factors with patterns of visual attention allocation shows promise for measuring display clutter and correlation with task performance, particularly for low-clutter displays.

Practitioner Summary: A novel measure of display clutter was formulated, accounting for display data content, user knowledge states and patterns of visual attention allocation. The measure was evaluated in terms of correlations with driver performance in a safety-critical driving simulation study. The measure exhibited stronger correlations with task performance than previously defined measures.

ARTICLE HISTORY
Received 31 January 2017
Accepted 31 October 2017

KEYWORDS
Display clutter; eye tracking; driver behaviour; visual search; interface design

Introduction

As in-vehicle display technologies have advanced in content and format, user ability to perceive and interpret display information has become an increasingly important design objective. Related to this, a cluttered information display can have a negative impact on human visual search for information (Wickens, Kroft, and Yeh 2000). Clutter is defined as, 'an unintended effect of display imagery that obscures or confuses other information or that may not be relevant to the task at hand' (Kaber et al. 2008). There are many image properties that have been identified in the literature, which can obscure or confuse information on a display, including similarity of target and background/distractor colour (Rosenholtz et al. 2005; Rosenholtz, Li, and Nakano 2007; Lohrenz et al. 2009), image luminance levels (Kim et al. 2011; Kaber et al. 2013), orientation differences between targets and background/distractor images (Rosenholtz, Li, and Nakano 2007; Bravo and Farid 2008) and differences in motion between the target and background/distractors (Rosenholtz et al. 2005). These features make human visual search more difficult by decreasing the salience of a target relative to a background and/or distractors.

In the driving domain, an estimated 4.5% of drivers between the ages of 16 and 24 in the United States (and 2.1% of drivers across all ages) manipulate hand-held devices while driving (National Highway Traffic Safety Administration 2017). Furthermore, in 2015, the United States National Highway Traffic Safety Administration revealed that 3477 people were killed and an estimated 391,000 were injured in automobile accidents in the United States resulting from distracted driving, a category which includes in-car device use (National Center for Statistics and Analysis 2016). These statistics highlight the importance of display designs that minimise clutter and, consequently, driver off-road glance time.

A key step in managing display clutter is developing a reliable measure of clutter that has some relation with task performance. On this basis, we conducted a review of existing measures of display clutter and reported relationships with task performance measures. We also review measures of visual attention allocation and how they are affected by display clutter, as visual attention could serve

CONTACT David B. Kaber dbkaber@ncsu.edu

as a physiological measure that more accurately reflects the degree of clutter in a display. In the subsections that follow, we provide a brief review of clutter research, accompanied with implications for our development of a novel measure of display clutter that correlates with performance in a complex, safety-critical domain.

Measures of display clutter

Alexander et al. (2008) and Doyon-Poulin, Robert, and Ouellette (2012) identified the importance of both user knowledge factors as well as display data content factors in perceptions of clutter. Similarly, a recent review of the clutter literature identified four approaches to the measurement of clutter: (1) image-processing techniques, (2) performance evaluation techniques, (3) subjective evaluation and (4) eye tracking techniques (Moacdieh and Sarter 2015). Our review was constrained to only those subjective evaluation and image-processing techniques (which we term 'knowledge-driven' and 'data-driven', respectively) for which correlations with task performance have been observed; that is, we do not report on measures of clutter absent of relations with task performance.

Knowledge-driven measures

Knowledge-driven measures of clutter generally fall into one of two categories: simple ratings or rankings of cluttered images (McCrobie 2000; Neider and Zelinsky 2011) and images indexed based on subject matter expert identification of feature importance (St. John et al. 2005; O'Hara and Dwyer 2007). Knowledge-driven measures can be very simple and effective ways to obtain an estimate of display clutter. Table 1 summarises three knowledge-driven measures of clutter and how well they correlate with various performance measures, including the Clutter Rating (CR), Clutter Score (CS) and a 'clutter transition score' method. Except for Neider and Zelinsky's (2011) method, knowledge-driven measures have generally produced relatively weak linear associations with task performance.

Data-driven measures

Data-driven measures of display clutter focus on display properties, such as the salience of objects in specific display areas. Most data-driven measures have been developed in the computer science field and validated by comparison with subjective user ratings or rankings of clutter (Lohrenz et al. 2009; Yu, Samaras, and Zelinsky 2014). While some authors have reported strong correlations between data-driven clutter measures and task performance, other authors applying the measures have generally reported weaker correlations (refer back to Table 1). For example, Rosenholtz et al. (2005) and Rosenholtz, Li, and Nakano (2007) report strong correlations of Sub-band Entropy (SBE) and Feature Congestion (FC) with task performance (i.e. measures they developed); however, experiments by Asher et al. (2013) and Henderson, Chanceaux, and Smith (2009), making use of SBE and FC, report much weaker correlations. Generally, correlations between data-driven measures and task performance are relatively weak from a systems engineering perspective (i.e. linear model slope coefficients near 0).

There are several implications of this prior research on data-driven measures for formulation of a new integrated measure of clutter. A general theme throughout the literature on display clutter is the importance of accounting for both 'top-down' and 'bottom-up' influences on perceived clutter (Alexander et al. 2008; Doyon-Poulin, Robert, and Ouellette 2012). Unfortunately, the influence of each type of factor relative to the other has not been established. With this in mind, we speculated that a weighted summation of knowledge-driven and data-driven measures might be appropriate as a composite index, with the weighting factors reflecting the relative importance of each type of measure. Furthermore, we inferred that the large range of correlation values between clutter measures and performance outcomes might be due to varying experimental/testing contexts. On this basis, we also speculated that any weightings of knowledge-driven and data-driven measures might need to change based on external contexts and internal operator states.

Effects of clutter on visual attention allocation

Due to the generally weak associations between task performance and existing measures of clutter, which generally focus on either top-down or bottom-up influences, it is possible that visual attention allocation can be leveraged to increase the utility of a composite measure of display clutter. Attention allocation differs from knowledge- and data-driven measures in that it can be considered as an unbiased indicator of how one views a display given the occurrences of clutter. Knowledge-driven measures are subjective, and thus potentially biased by the display user's interpretation of clutter, experience using the display and/or mental workload at the time of display use. Data-driven measures are biased by the researcher interpretation of the construct of clutter and underlying dimensions.

Beck, Lohrenz, and Trafton (2010) reported an increase in the number of fixations and proportion of distractors fixated in a target search task for high-clutter aeronautical maps. Similarly, Lim, Tsimhoni, and Liu (2010) reported an increased number of glances and mean glance time with increased clutter in a driving task utilising night vision technologies to detect pedestrians. In a target search of various *Where's Waldo?* scenes, Moacdieh and Sarter

Table 1. Knowledge-driven and data-driven clutter measures and correlations with task performance.

Clutter measure	Description	Source	Performance measure	Correlation coefficient
Knowledge-driven measures				
Clutter Rating (CR)	Unidimensional rating of clutter	Alexander et al. (2012)	Flight Localiser RMSE Flight Glideslope RMSE	Not reported due to lack of significance
		Kaber et al. (2013)	Flight Glideslope Deviation	$r = -0.14$ to -0.10
Clutter Score (CS)	Structured, rank-weighted measure of clutter, accounting for subjective perceptions of multiple factors considered to contribute to display clutter	Alexander et al. (2012)	Flight Localiser RMSE	$\rho = -0.5$ to 0.12
			Flight Glideslope RMSE	$\rho = 0.03$ to 0.08 (NS)
		Kaber et al. (2013)	Flight Glideslope Deviation	$r = -0.27$
			Flight Groundspeed Control	$r = -0.14$
Order Images by Clutter + 'Clutter Transition Score'	Images are ordered by perceived clutter, then assigned a between-images 'Clutter Transition Score'	Neider and Zelinsky (2011)	Search Time	$r = 0.72$ to 0.73
Data-driven measures				
Dissimilarity-Based Clutter Measure	Determines the degree of dissimilarity of a target to background noise	Xu and Shi (2013)	Search Time	$\rho = 0.77$
			Detection Time	$\rho = 0.80$
			False Alarm Rate	$\rho = 0.82$
Edge Density	Calculates the proportion of pixels in an image that are considered 'edge points'	Asher et al. (2013)	Area Under ROC Curve	$r = -0.21$ (Local and Global)
			True Positive Rate	$r = -0.12$ (NS)
			False Positive Rate	$r = 0.19$
			Search Time	$r = 0.02$ (NS)
		Neider and Zelinsky (2011)	Search Time	$r = 0.31$ to 0.56
		Henderson, Chanceaux, and Smith (2009)	Search Time	$r = 0.53$
			Search Accuracy	$r = 0.54$
Feature Congestion	Quantifies the impact of adding a salient item to a display without decreasing the level of salience of other surrounding objects	Asher et al. (2013)	Area Under ROC Curve	$r = 0.14$ (NS)
			True Positive Rate	$r = -0.01$ (NS)
			False Positive Rate	$r = 0.18$
			Search Time	$r = 0.01$ (NS)
		Henderson, Chanceaux, and Smith (2009)	Search Time	$r = 0.27$
			Search Accuracy	$r = 0.16$
		Rosenholtz et al. (2005)	Search Time	$r = 0.74$ to 0.76
Multi Scale Segmentation	Formulates a power law function using a segmentation algorithm to determine how many distinct object parts appear in an image	Asher et al. (2013)	Area Under ROC Curve	$r = 0.14$
			True Positive Rate	$r = 0.03$ (NS)
			False Positive Rate	$r = -0.15$ (NS)
			Search Time	$r = -0.01$ (NS)
Sub-band Entropy	Measures the efficiency with which an image can be encoded via an algorithm (similar to a JPEG image coder), positing that less cluttered images are encoded more efficiently	Asher et al. (2013)	Area Under ROC Curve	$r = -0.21$ (Local) $r = -0.3$ (Global)
			True Positive Rate	$r = -0.11$ (NS)
			False Positive Rate	$r = 0.2$
			Search Time	$r = 0.05$
		Henderson, Chanceaux, and Smith (2009)	Search Time	$r = 0.42$
			Search Accuracy	$r = 0.53$
		Rosenholtz, Li, and Nakano (2007)	Search Time	$r = 0.75$ to 0.77

(2012) reported increases in number of fixations, fixation frequency and mean fixation duration as a function of increasing clutter, but no significant effect on mean glance duration to the target AOI. Generally, cluttered displays require operators to search and process more information, resulting in more visual attention being allocated to the display. Instances of conflicting results, for example the difference in the glance time effect of clutter reported in Lim, Tsimhoni, and Liu (2010) vs. Moacdieh and Sarter (2012), may be attributable to differences in task context. That is to say, visual attention allocation may be influenced by clutter differently in a dynamic driving task than in a static target search task, for example.

There are several implications of this prior research on visual attention allocation for formulation of a comprehensive measure of display clutter. The unbiased nature of attention allocation (compared to knowledge-driven and data-driven measures) suggests it could strengthen the utility of a composite measure for differentiating between levels of clutter. Since cluttered displays generally require more operator visual attention, attention patterns may be an effective means for weighting bottom-up and top-down clutter contributors in order to distinguish between displays of varying levels of clutter. In particular, measures reflecting both temporal and spatial aspects of visual attention to a display or scene might be useful for promoting the accuracy of clutter measures.

Local and global clutter effects

The majority of existing work has focused on effects of global clutter, defined as the level of clutter on an overall display. Lohrenz and Beck (2010), however, sought to uncover the effects of local clutter, defined as the level of clutter in a smaller AOI on a display, in a map search task. During the search, eye tracking responses revealed participants to initially search areas of low clutter and then continue to areas of progressively higher clutter. Furthermore, the authors reported participants tended to avoid the highest clutter areas, suggesting a threshold level of clutter on which display users are not willing to fixate. Similarly, Beck, Lohrenz, and Trafton (2010) reported interaction between local and global clutter such that there was an additive effect on the total number of fixations in non-salient target search; however, there was no interaction when the target was salient. In the salient target case, increasing both local and global clutter generally led to more fixations in the search task. Taken together with the results of Lohrenz and Beck's (2010) research, there may be differential effects of local and global clutter, and display users may tend to avoid areas of high local clutter.

Motivation for present research

The corpus of clutter literature reveals that existing measures generally do not exhibit reliably high correlations with task performance. For this reason, it is important to develop a measure that can reliably predict user performance in safety-critical domains and that is sensitive to potential negative effects of cluttered displays. It is possible that a measure capturing display data-content factors, user knowledge-states and visual attention allocation patterns might have greater utility for explaining task performance than simpler measures accounting for just one aspect of the clutter construct. With this in mind, we identified two objectives for the current research, including: (1) formulate a new integrated measure of display clutter accounting for both display data (objective factors) and user knowledge (subjective factors) in the display clutter response, weighted by visual attention allocation indicators; and (2) assess the validity of the new measure for predicting performance in the context of driver in-car display use in a complex driving simulation. In general, driving is a complex information processing task in which display clutter can be an issue from performance and safety perspectives. Furthermore, most clutter research has focused on the aviation domain (Kaber et al. 2008, 2013) and driving is under-represented in clutter research. Beyond this, clutter issues may actually be more safety-critical in driving due to the greater likelihood of vehicular collisions. As an ancillary objective, we also sought to identify any differences between the effects of local and global clutter on any relation of the new measure with driving performance responses.

Development of a comprehensive measure of display clutter

As mentioned in our review of the literature, both top-down and bottom-up influences must be accounted for in measuring clutter responses. We hypothesised that a weighted summation could account for both types of influences. In addition, since clutter has been demonstrated to alter display user patterns of visual attention, we hypothesised that leveraging attention allocation as an unbiased weighting factor (i.e. a multiplier of other indicators) could increase the utility of the composite measure of display clutter. With these trends in mind, we formulated an equation to capture data, knowledge and visual attention variables in quantifying and predicting clutter responses. We refer to the equation as a Comprehensive Measure (CM) of display clutter, as shown below with accompanying variable definitions in Table 2.

Table 2. Description of variables that comprise the comprehensive measure of clutter.

Variable	Description
Longest Glance	Longest glance to the display or AOI, where a glance is defined as the total time the focus of attention remains within an AOI or on the entire navigation display, encompassing both fixations and saccades
Fixation Frequency	Number of fixations to an AOI divided by the total number of fixations to all AOIs in a scene during a defined observation period. A fixation occurs whenever the focus of visual attention remains 'fixed' (gaze speed < 100 degrees/sec) for a minimum of 100 ms (Holmqvist et al. 2011)
Edge Density (ED)	A measure of the proportion of pixels classified as 'edge points', used as a data-driven measure of clutter in an image
w_{ED}	Weight reflecting the contribution of data-driven aspects of clutter (i.e. edge density), relative to knowledge-driven contributors, to any association of the measure with task performance
Reduced Clutter Score (RCS)	A knowledge-driven multidimensional measure of clutter, measured as a rank-weighted sum of perceptions of colourfulness, consistency/similarity and dynamics/variability
w_{RCS}	Weight reflecting the contribution of knowledge-driven aspects of clutter (i.e. RCS), relative to data-driven contributors, to any association of the measure with task performance

$$\text{Clutter} = \left(\frac{\text{Longest Glance}}{\text{Longest Glance}_{max}}\right) \times \text{Fixation Frequency} \times \left(w_{ED} \times \left(\frac{\text{Edge Density}}{\text{Edge Density}_{max}}\right) + w_{RCS} \times \left(\frac{\text{Reduced Clutter Score}}{\text{Reduced Clutter Score}_{max}}\right)\right) \quad (1)$$

The weighted sum of knowledge-driven and data-driven measures of clutter can account for (a) the relative importance of each type of factor on task performance, as well as (b) relative factor effects given varying external contexts (e.g. task conditions) and internal operator states (e.g. operator goals). We chose to use ED to represent the data-driven component of our measure since it is the most widely used objective metric of clutter in the existing literature (Rosenholtz, Li, and Nakano 2007). Furthermore, Marr and Hildreth (1980) reported that the ED metric replicates the human visual system's sensitivity to edge contrast, providing further justification for its inclusion in the CM from a visual information processing perspective. User subjective perceptions of clutter are represented in the equation in terms of the Reduced Clutter Score (RCS) approach, as proposed by Kaber et al. (2013). The RCS measure includes ratings of several contributors to perceived clutter and allows for quick administration relative to other multi-dimensional subjective indices of clutter (e.g. the CS by Kaber et al. (2008)). Kaber et al. (2013) presented a factor analysis identifying three latent variables in clutter, as identified in Table 3. These three subdimensions accounted for a similar level of variability in pilot perceptions of clutter as the more complex six-subdimension CS. Finally, we chose two commonly used visual attention allocation measures to reflect temporal and spatial aspects of attention required for an operator to process information on a display, including fixation frequency and longest glance duration, as defined in Table 2. Previous driving research has reported that longer glances away from the roadway, due to driver distraction, can lead to vehicle control uncertainty (Wierwille 1993) and potential safety issues. Other research has suggested fixation frequency to a target AOI to be predictive of perceived clutter (van Orden et al. 2001; Moacdieh and Sarter 2012).

All objective and subjective terms included in Equation (1) were normalised so that the magnitude of each term was between zero and one. Note that fixation frequency (by definition) is a proportion with values from zero-to-one and normalisation was not necessary. Since clutter is highly dependent on user domain and task goals (Alexander et al. 2008), the terms w_{ED} and w_{RCS} were used to weight the ED metric and RCS in the equation, respectively. These weighting factors were considered critical in terms of any correlation of the CM with task performance since the values could change dramatically depending on a user's task. Furthermore, since prior research demonstrated local and global clutter to have differential effects on task

Table 3. Definitions of reduced clutter score subdimensions.

Subdimension	Anchors	Definition
Colourfulness (Data-Driven)	Colourful (High Clutter)	Full of colour; having striking colours; characterised by rich variety (Synonyms: vivid; vibrant; striking)
	Monochromatic (Low Clutter)	Having or consisting of only one colour or hue; lacking variety, creativity, or excitement (Synonyms: colourless; neutral; plain)
Consistency/Similarity (Data-Driven)	Obscure (High Clutter)	Difficult to detect; characterised by similarity or repetition (Synonyms: inconspicuous; repetitive; unnoticeable)
	Distinct (Low Clutter)	Standing out conspicuously; of notable significance beyond its surroundings; composed of mutually exclusive elements (Synonyms: unique; prominent; striking)
Dynamics/Variability (Knowledge-Driven)	Changing (High Clutter)	Likely to vary; likely to change frequently, suddenly, or unexpectedly (Synonyms: unstable; altering; inconstant)
	Unchanging (Low Clutter)	Repetitiously dull or lacking in variety; showing little or no change, action, or progress (Synonyms: uniform; rigid; unvarying)

performance, the CM equation can be applied to an entire display (global) or specific AOIs within a display (local).

The CM was formulated to be used by designers during an iterative, user-centred design process. Initially, there is a need for empirical work to define appropriate weights, w_{ED} and w_{RCS}, for a given domain and operator goal state. Once weights have been established, a designer can apply the CM equation to prototype display analysis using human-in-the-loop task simulations. Post hoc performance and clutter response analysis can be used in conjunction with user feedback to enhance display design. The new CM can be used to: determine whether clutter is present in a display, quantify the relative degree of display clutter and support analysis of how clutter may be related to performance degradations.

Methodology

Having formulated the new CM, we sought to assess the utility of the equation for distinguishing among displays as well as predicting user performance in a safety-critical domain sensitive to display clutter issues, specifically automobile driving. A driving simulation and route navigation experiment was designed to assess the strength of any correlation of CM values with driving task performance measures, and how these correlations compare to correlations with other existing clutter measures (CR, CS and ED). The research complied with the American Psychological Association Code of Ethics and was approved by the Institutional Review Board. Informed consent was obtained from each participant.

(a)

(b)

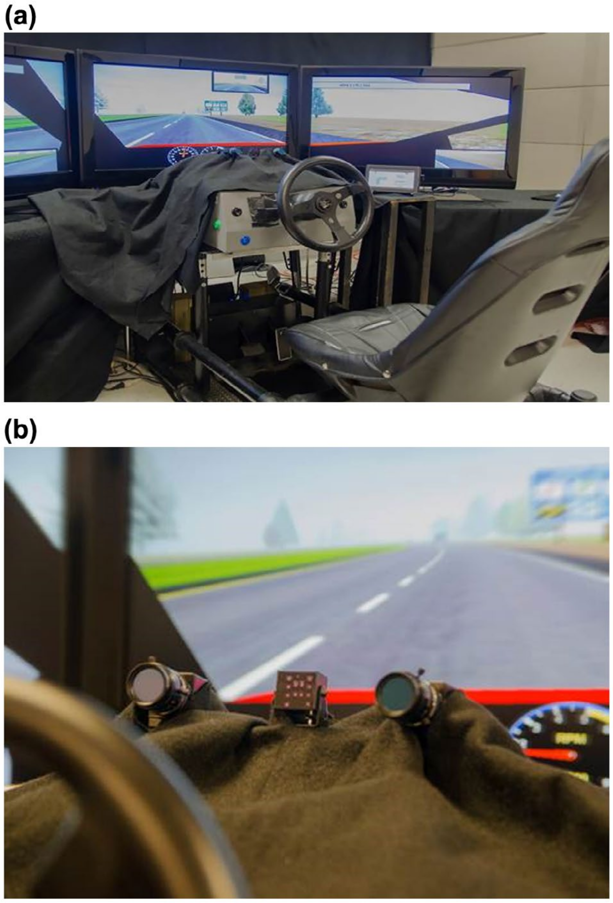

Figure 1. (a) Driving simulator hardware and (b) FaceLab eye tracking system hardware.

Participants

Twenty licensed drivers were recruited for participation in the experiment, with gender balanced within the sample. All participants were required to have 20/20 or corrected vision and full colour vision. The drivers in the sample had, on average, 10.05 ± 8.09 years of driving experience and were paid $20/h for their time. Participants were required to be between 18 and 60 years of age, as Chen et al. (2007) observed degradations in driving skills for persons 60 years and older.

Apparatus

A STISIM Drive M400 driving simulator (Systems Technology, Inc.) was used to present the driving task. The simulator included three 38-inch high-definition television monitors providing a 135° field of view of the driving environment (see Figure 1(a)). Similar systems have been used in prior studies with performance results validated against real-life driving behaviour (Wang et al. 2010). To the right of the forward view of the simulated roadway was a tablet computer (iPad 2, Apple Inc.), which presented a typical route navigation-aid display corresponding to the driving scenario. Driver visual attention allocation to the roadway or navigation aid was captured using a FaceLab 5 eye tracking system (Seeing Machines, Inc.; Figure 1(b)), which has a sampling rate of 60 Hz with an accuracy of 0.5° to 1° of rotational error.

The navigation-aid display included five navigation features that were identified as meeting Kaber et al.'s (2008) definition of clutter, when a driver's goal was to navigate to a destination since clutter is user goal dependent. These features included: (1) traffic density indication; (2) satellite imagery; (3) alternate route traffic density indication; (4) alternate route slower/faster tags; and (5) landmark labels. Figure 2(a) shows the 'low' clutter navigation display condition containing none of these features while the 'high' clutter display (Figure 2(b)) contained all of the five 'clutter' features. A preceding presentation-based experiment revealed the high-clutter display to result in longer query response time ($p = 0.001$) and higher RCS ($p < 0.001$) than the low-clutter display (Pankok 2015). Furthermore, ED measurements corroborated these results: high clutter

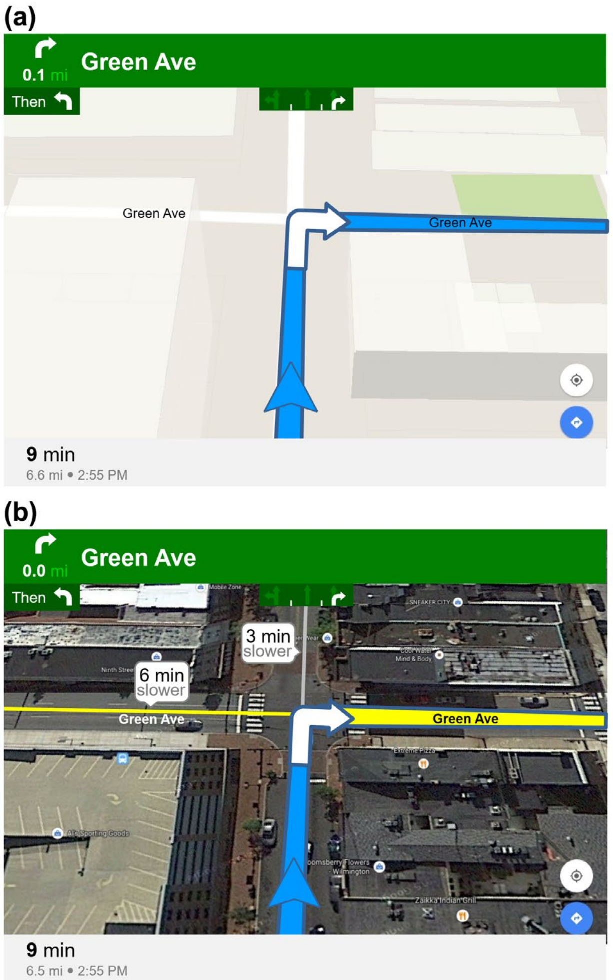

Figure 2. Example (a) low clutter display and (b) high clutter display combinations.

Figure 3. Screenshot of the driving simulator environment.

displays contained a mean of 10.6% edge pixels while the low clutter displays contained a mean of 3.4% edge pixels, suggesting that our approach successfully produced displays at different levels of clutter.

Driving task

In each simulation run, participants were required to drive an urban roadway, containing intersections and stoplights, to an identified destination under clear weather conditions. All roadways contained six lanes in total with three lanes in each direction of travel and had a 35-mph speed limit (see Figure 3). This type of roadway was used since it was expected to impose a higher cognitive load than normal freeway driving, but not so high as to substantially detract from driver capability to use the navigation aid and make assessments of perceived clutter. Traffic density on all roads was comparable to Level of Service C, approximately 6–8 vehicles per lane per minute (Transportation Research Board 2010).

At four predetermined points along the drive, participants were required to respond to a navigation query using information that could only be obtained from the navigation-aid display. All queries were presented via an automated voice in the simulator (e.g. 'What is your expected time of arrival at the destination?'). Immediately following a driver's verbal response to a query, the simulation was paused and (s)he was asked to provide global CRs and RCS subdimension ratings. The participant was asked to identify the AOI of the navigation-aid display that (s)he used to obtain the required information to answer the query, then provided local CRs and RCS subdimension ratings for the identified AOI. There were four AOIs the participant could identify, which were empirically verified using fixation locations recorded for participants in responding to navigation queries in the preceding presentation-based experiment (Pankok 2015, 72, 73). The navigation display was visible throughout the CR process, and definitions of all subdimension scale anchors were available to participants for each rating. After completing the CR and RCS subdimension ratings, the participant continued driving until the next query was posed or until (s)he arrived at the route's destination.

Experiment design

The experiment followed a combined Latin square design (Giesbrecht and Gumpertz 2004) with one of the two navigation display clutter levels assigned to each participant in each driving scenario. Testing of each level of display clutter was replicated once for each participant, resulting in four total driving scenarios (2 levels of clutter × 2 replicates). The 'scenario' blocking variable was defined by: (1) a driving route (denoted as 1 or 2); and (2) a set of queries (denoted as A or B) applied to the route. Each scenario contained different street names in order to prevent participants from learning a route and driving to the destination

Table 4. Latin square assignment of experiment combinations.

Scenario	Participant 1	Participant 2	Participant 3	Participant 4
1A	High	Low	Low	High
1B	Low	High	High	Low
2A	Low	High	Low	High
2B	High	Low	High	Low

without using the navigation display. Table 4 presents an example of the Latin square design assignments to four participants. Every set of four participants was assigned a similar Latin square (thus making it a combined Latin square design) and the order of presentation of the clutter conditions across scenarios was randomised for each participant.

Procedure

On initial contact, participants were asked to verify their age, driving experience, visual acuity, colour vision and status as a licensed U.S. driver. Upon arrival at the laboratory, participants were asked to complete an informed consent form, a demographic questionnaire and a baseline simulator sickness questionnaire (SSQ; Kennedy et al. (1993)). The SSQ was used to monitor potential motion sickness symptoms during the experiment. After completion of these items, driver Useful Field of View (UFOV) was measured using the Visual Awareness (Punta Gorda, FL) software package; Multivariate Analysis of Variance revealed UFOV responses to have no effect on any of the visual attention allocation or driving performance responses (all p-values > 0.17).

The initial training trial required participants to perform manoeuvres that were required during the experiment trials (e.g. changing lanes, making turns, etc.) in the absence of a navigation aid display. Following this, participants were trained on the use of the navigation-aid display, including definitions and explanations on how to interpret the features and information presented on the display. In the final training session, participants were required to respond to navigation queries while driving, replicating the procedure of the test trials (but utilising a route that was different from any experiment trial route). After training, the eye tracking system was calibrated for the individual driver.

The experiment commenced by asking the participant to complete the pre-experiment pairwise rankings of the RCS subdimensions, as part of the RCS procedure (Kaber et al. 2008). Subsequently, the participant began the first driving task. Participants were told that they were navigating to a friend's new house, whose address was programmed into the navigation system. They were instructed to drive at the posted speed limit and to stay in the centre lane of the roadway unless an upcoming turn required a lane change. Between experiment trials, participants were asked to exit the cab for at least 5 min and were administered a SSQ. If participants exhibited any simulator sickness symptoms (e.g. disorientation, nausea, oculomotor disturbances), they were allowed to rest for up to an additional 20 min. If symptoms persisted, despite the extra rest, the participant was dismissed from the experiment without penalty. No participants were dismissed from the experiment due to simulator sickness. The experiment lasted approximately 2 h in total for each participant.

Response variables

In addition to the measures included in the CM, driving performance measures were collected by STISIM with a sampling rate of 30 Hz, including lane deviation and speed deviation. Lane deviation was defined as the average absolute deviation of the centre of the vehicle from the centre of the lane, where larger values corresponded to degraded vehicle control. Speed deviation was defined as the average absolute deviation from the posted speed limit (56.3 kph/35 mph), where larger values correspond to degraded vehicle control. Lane maintenance and speed control were measured since they represent first-order (direct) measures or driver vehicle control and indicators of safe driving performance. Subjective perceptions of clutter were collected via Google (Mountain View, CA) electronic forms. Local and global versions of all measures were recorded. Since the CM was defined for a time 'window' in which the goal of the display user was known, all visual attention and driving performance measures were captured and summarised for the same time window. Each window started when a query was posed to the driver and ended when the driver answered the query. In this way, the CM values in this study represent clutter outcomes for specific user knowledge states.

Data analysis

A sensitivity analysis was performed on the w_{ED} and w_{RCS} parameter coefficients as part of the CM equation in order to identify weights yielding the strongest correlations of the CM outcomes with task performance. (w_{ED}, w_{RCS}) combinations of (0, 1), (0.25, 0.75), (0.5, 0.5), (0.75, 0.25) and (1, 0) were tested. It should be noted that the (1, 0) combination causes the CM to become a strictly display-data driven measure of clutter weighted by attention allocation, and the (0, 1) a strictly knowledge-driven measure weighted by attention allocation. The sensitivity analysis was performed on both the global and local versions of the CM. Correlation analyses were conducted to assess associations between the driving performance variables and three other clutter measures: (1) RCS, (2) ED and (3) CR. The objective of these analyses was to provide a comparison between the CM and existing measures of display clutter from the literature. None of the responses exhibited normality, so Spearman correlation coefficients were computed to assess any relations among the variables.

Results

An overview of the effect of the clutter manipulations on the various response measures is presented in Table 5.

Sensitivity analysis

Lane deviation

Figure 4 presents the correlations between the CM and lane deviation measures with the weights, w_{ED} and w_{RCS}, at the quartiles (as identified above). The global version of the CM exhibited significant correlations with lane deviation across both levels of display clutter and within the low clutter display condition, but not within the high clutter display condition. The local version of the CM (focused on AOIs), however, did not reveal any significant correlations with lane deviation. Across all displays, the global CM produced the strongest correlation with lane deviation when $w_{ED} = 0.25$ and $w_{RCS} = 0.75$; within the low clutter displays, the strongest correlation with lane deviation occurred when $w_{ED} = 1.00$ and $w_{RCS} = 0.00$.

Speed deviation

As presented in Figure 5, statistically significant correlations occurred only for the global version of the CM within the low clutter display condition. The strongest association of the response measures occurred when $w_{ED} = 0.75$ and $w_{RCS} = 0.25$.

Correlations among other clutter measures

Correlation analyses were also conducted on observations on the three other clutter measures collected during the study, including the CR, ED and RCS, as compared with driver lane and speed deviations. Figure 6 presents relations of clutter with lane deviations and Figure 7 presents clutter correlations with speed deviations. The local version of the ED measure was the only measure (outside of the CM) significantly correlated with lane deviations, occurring across all displays and within low clutter displays. Within high clutter displays, no measures were significantly correlated with lane deviation. There were no significant correlations with speed deviation. A comparison of the correlation coefficients between the driving performance responses and the clutter measures is presented in Table 6.

Table 5. Descriptive statistics on measures collected during the experiment.

Response	Clutter level	Mean ± SD
Lane Deviation (m)	High	0.262 ± 0.137
	Low	0.262 ± 0.157
Speed Deviation (kph)	High	2.425 ± 2.181
	Low	2.044 ± 1.790
Fixation Frequency	High	0.278 ± 0.155
	Low	0.267 ± 0.136
Longest Glance (ms)	High	719 ± 465
	Low	714 ± 465
Global Clutter Rating	High	2.377 ± 1.938
	Low	1.971 ± 1.329
Global RCS	High	5.321 ± 1.639
	Low	3.309 ± 1.625
Edge Density	High	0.106 ± 0.003
	Low	0.034 ± 0.002

Discussion

Four trends emerged from the correlation analyses between the clutter measures and simulated driving task performance with route navigation. First, the global version of the CM exhibited stronger correlations with driving performance than the local version. The local versions of all of the measures did not exhibit significant correlations with speed deviation; however, the local ED metric did show a significant association with lane deviation. The general lack of significant linear association between the local measures and driving performance may be attributed to the fact that the navigation-aid display was relatively small, so glances and fixations to AOIs were very short and participants were able to use parafoveal and peripheral vision to perceive and process display elements. It is possible that an experiment utilising a larger display with larger AOIs may not exhibit this trend.

The second general trend was that the CM generally exhibited stronger correlations with performance as the data-driven components of the measure were weighted more heavily, with the strongest correlations occurring when $w_{ED} = 0.75$ or $w_{ED} = 1.00$. As we identified in formulation of the CM, weights were expected to reflect the relative effect of top-down and bottom-up influences of clutter as related to operator internal and external states. This finding suggests that changes in driving performance resulting from display clutter were more dependent on display data factors (i.e. physical features) than on user knowledge factors (i.e. goal states, task information requirements, and awareness of display content). Metz, Schömig, and Krüger (2011) stated that drivers adapt eye movements to the demands of a situation, devoting more attentional resources to a task requiring a higher level of information processing. In our experiment, the high workload imposed by the driving task required participants to make very brief (714–719 ms, on average) glances to the display to obtain relevant information for responding to queries. These quick glances may have made it difficult for participants to consciously perceive and process the dynamics/variability contributor to clutter, as compared to

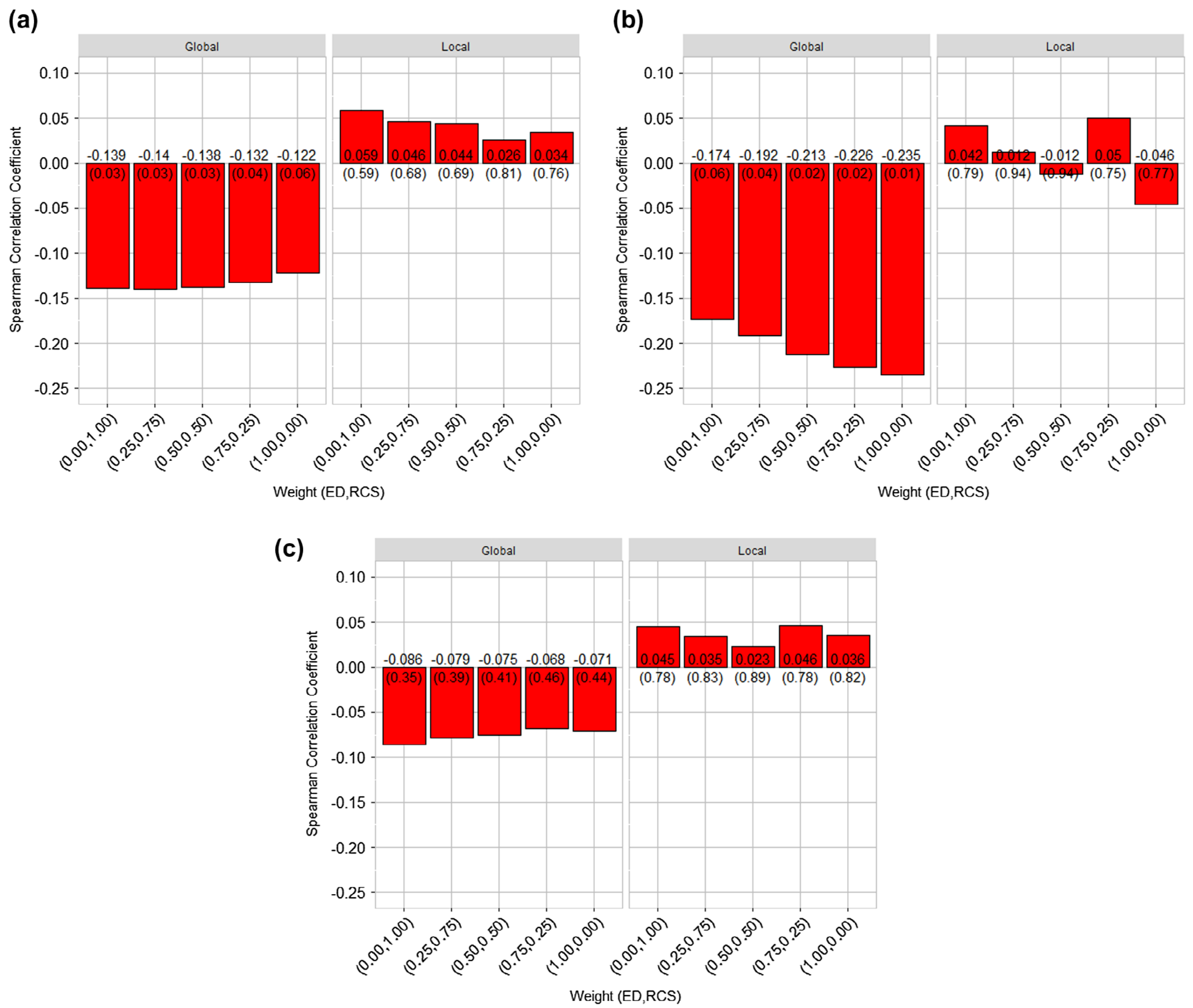

Figure 4. Spearman correlation coefficients between the CM and lane deviation for (a) all displays, (b) low clutter displays and (c) high clutter displays (*p*-values appear in parentheses).

situations permitting longer glances to the display. More extended glances to the display may be required to consciously process dynamic/variable features with inherent temporal properties (as opposed to consistency/similarity and colourfulness, which are not time-dependent). Therefore, participant gaze was influenced to a greater extent by the presence of display edge content as compared to the need for display data processing and interpretation, which exploits visual system pre-processing and image perception capabilities. Display data processing or interpretation has less of an influence on visual behaviour.

Counter to the trend of generally large weights for the ED metric in the CM equation, the association between the global CM and lane deviation for all displays was strongest when w_{ED} = 0.25 and w_{RCS} = 0.75. This suggests that ED was better able to distinguish levels of clutter from the group of low clutter displays than the RCS. As further evidence for this observation, the strength of association between ED and the RCS was significantly stronger when measured for all displays ($\rho = 0.476$, $p < 0.01$) than when it was measured only for the low clutter displays ($\rho = 0.018$, $p = 0.83$). In other words, the ED measure and the RCS were both able to distinguish among levels of clutter across all displays, but not for low clutter displays. This difference in trends highlights one of the potential issues with relying on knowledge-driven measures of clutter. In specific, subjective perceptions of clutter may not necessarily agree with performance indicators, particularly when display users need to differentiate levels of clutter for displays that are similar in appearance.

A third trend that emerged from the correlation analyses was that the CM exhibited stronger associations with performance measures when drivers were exposed to low clutter display conditions, as compared with high clutter

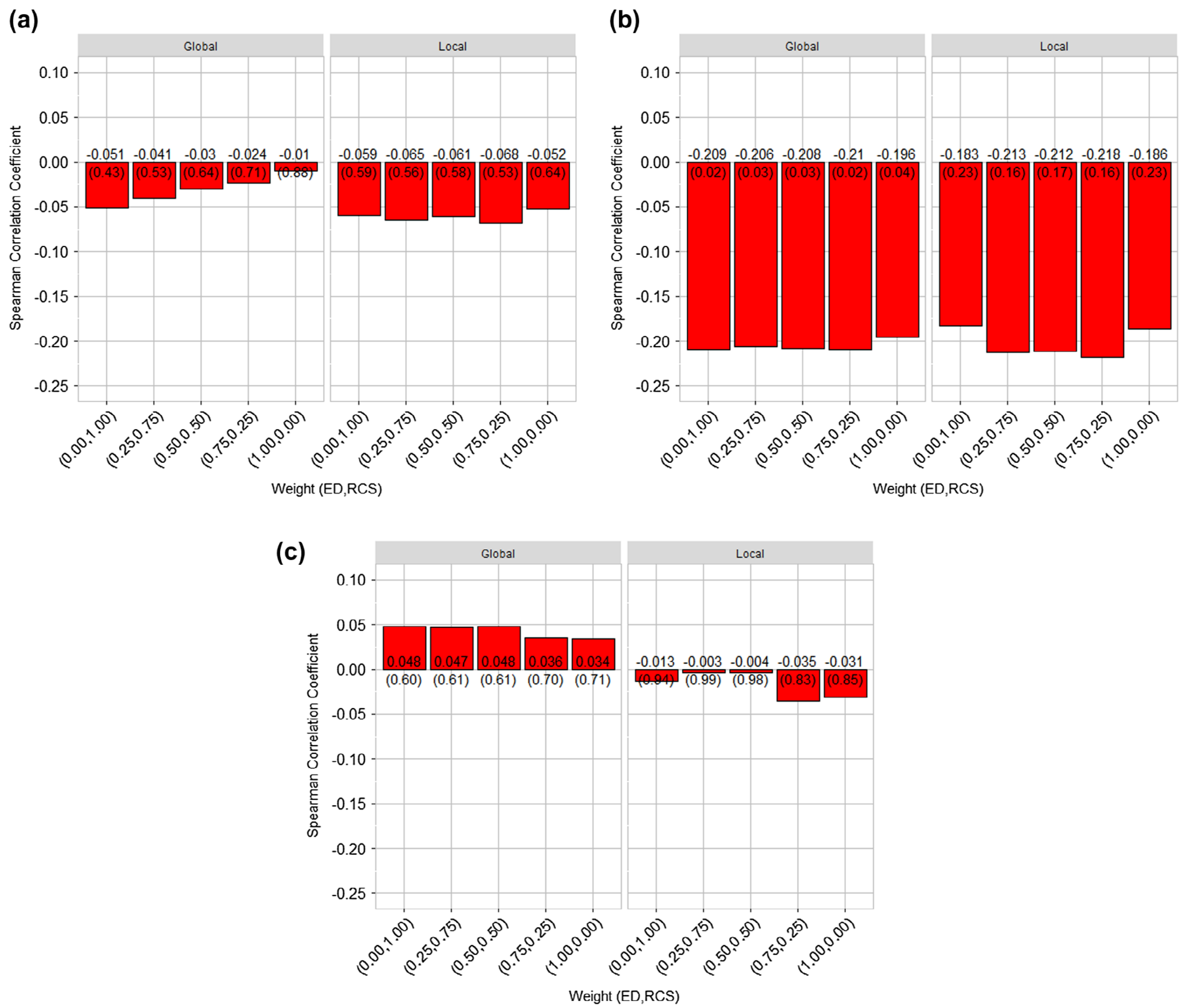

Figure 5. Spearman correlation coefficients between the CM and speed deviation for (a) all displays, (b) low clutter displays and (c) high clutter displays (*p*-values appear in parentheses).

conditions. This finding may be due to the fact that the high clutter displays contained a larger range of within-AOI (i.e. local) clutter than the low clutter displays. For example, participants generally rated the AOIs across the top and bottom of the high clutter display as being minimally cluttered, but rated the middle of the display as highly cluttered. This was not the case with low clutter displays, where all AOIs yielded similarly low perceptions of clutter.

The fourth trend, which was an unexpected result of the research, was the negative associations between the clutter measures (including the CM) and the driving performance responses. Both longest glance and fixation frequency were negatively correlated with the driving performance measures ($-0.26 < \rho < -0.03$), and since these measures were used as multipliers to the other components of the CM equation, the CM values were, ultimately, negatively correlated with driving performance. In particular, these negative correlations between visual attention allocation and driving performance are the reason that the local ED was positively correlated with lane deviation, but the global CM (which incorporates ED) was negatively correlated with lane deviation. This finding suggests that participants adopted a clutter avoidance search strategy; that is, the workload associated with the driving task caused participants to avoid searching high clutter displays. Clutter avoidance has also been reported by Lohrenz and Beck (2010), who asked participants to search aeronautical navigation maps of varying levels of clutter for a target symbol among distractors. Their results suggested that participants were able to quickly identify high clutter areas in the pre-processing phase to perception (using physical display properties) and to avoid areas containing high levels of clutter. However, their search task was not subject to competing demands of another task

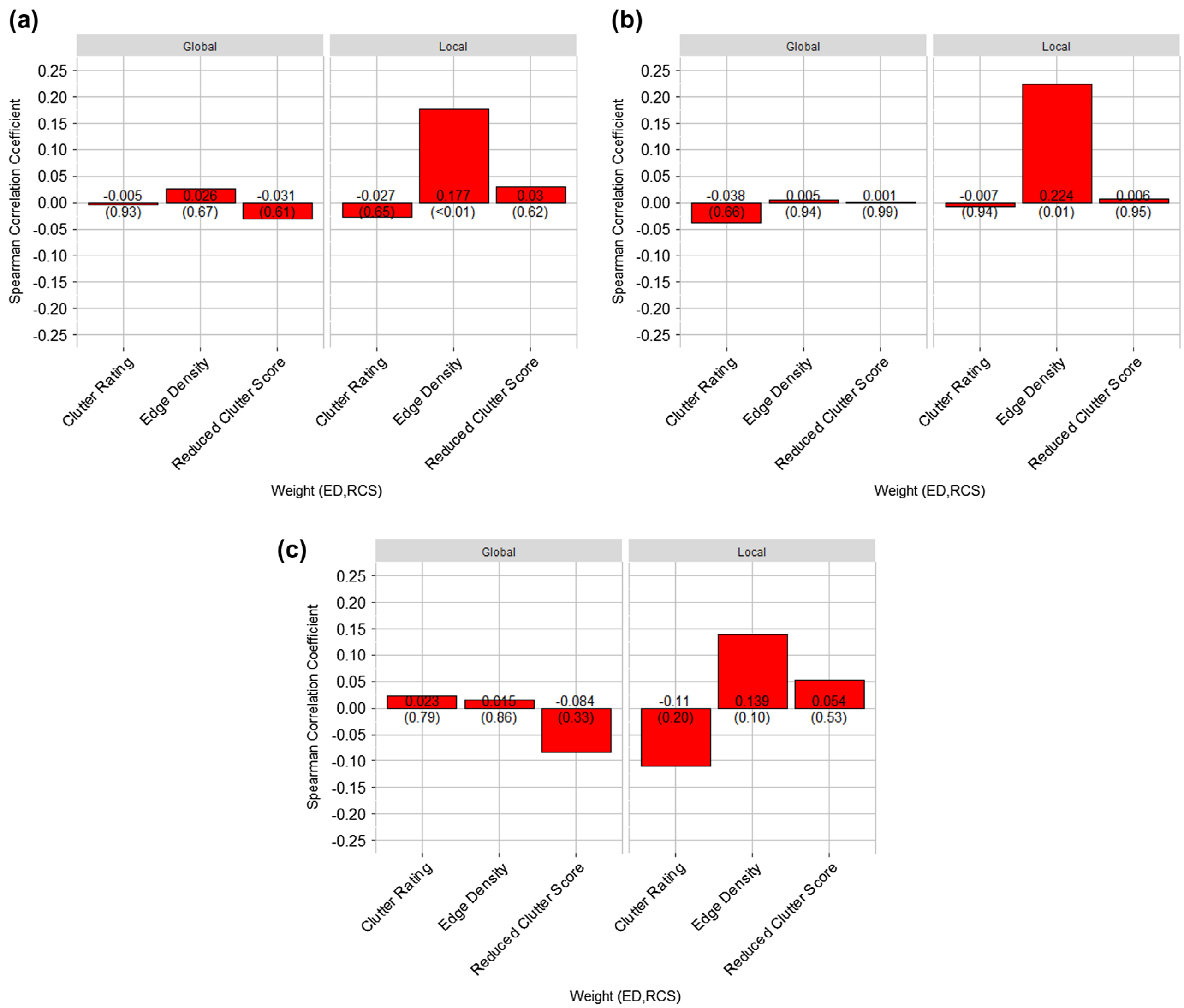

Figure 6. Spearman correlation coefficients between other clutter measures and lane deviation for (a) all displays, (b) low clutter displays and (c) high clutter displays (*p*-values appear in parentheses).

(like driving or flying), as was the case in our experiment. In the present study, drivers appeared to adopt a strategy of allocating less visual attention to the navigation display and more attention to the roadway when exposed to high clutter displays, resulting in superior vehicle control performance, similar to the attention allocation strategy adopted by drivers in Metz, Schömig, and Krüger (2011).

The correlations between the CM and driver vehicle control generally reflect the trends of the existing research on the aviation domain with correlation coefficients ranging from −0.27 to 0.08 (Alexander et al. 2012; Kaber et al. 2013). Correlations with search time are generally higher but also exhibit much larger variability. For example, Asher et al. (2013) reported a correlation coefficient of 0.2 with search time of static images of natural scenes, but Rosenholtz et al. (2005) reported correlation coefficients as high as 0.76 in target search of various maps. These disparate results highlight the importance of both internal operator states and external conditions in perceptions of clutter and its effects on task performance. The negative correlations reported in our research, along with the generally low and negative correlations of clutter and performance observed in the aviation domain suggest that clutter may generally not be a good predictor of task performance beyond visual search of static images or displays. However, the absolute associations between the CM and driving performance were generally stronger than for any of the other clutter measures examined in this study, showing promise for measures incorporating multiple contributors to display clutter.

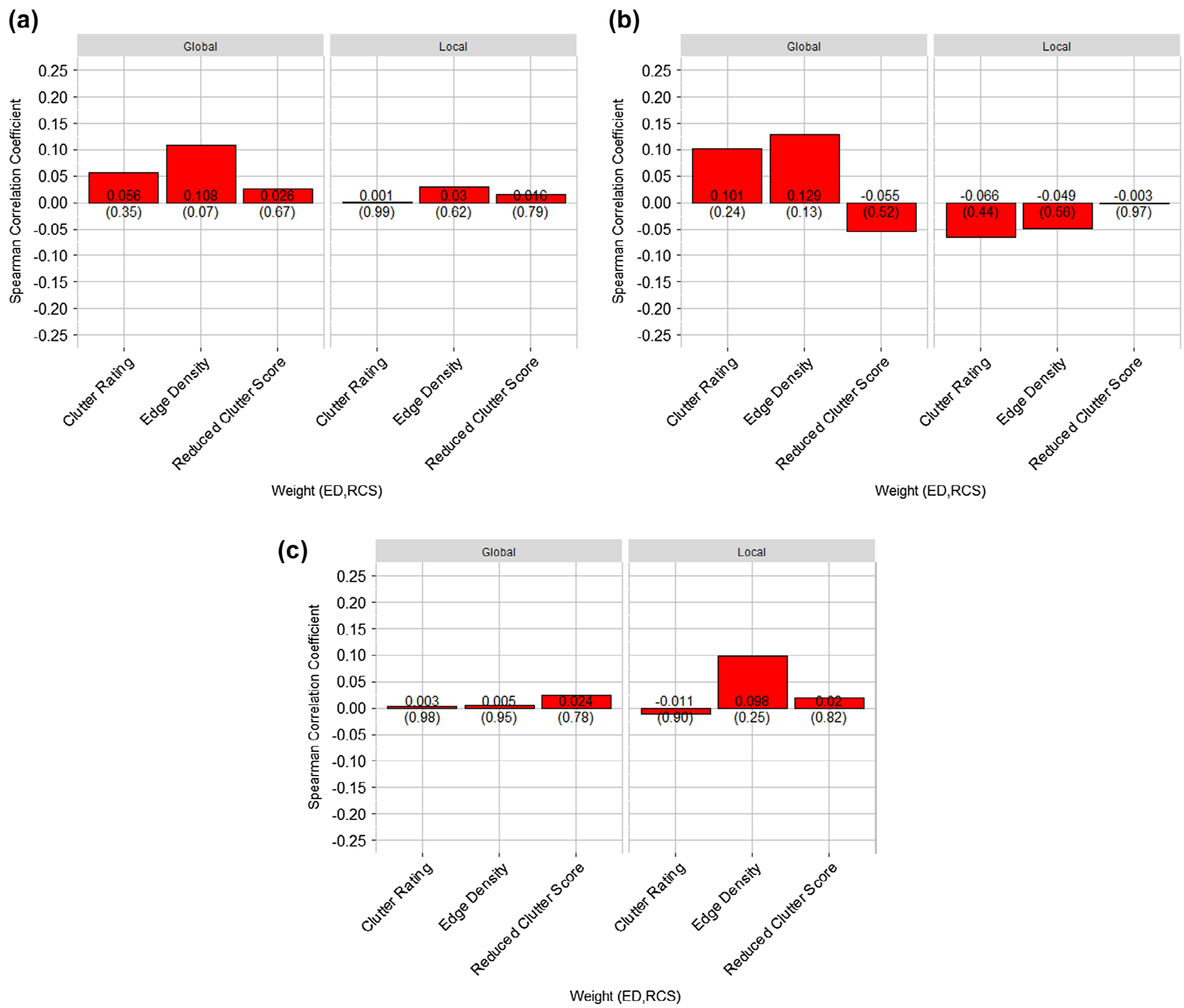

Figure 7. Spearman correlation coefficients between other clutter measures and speed deviation for (a) all displays, (b) low clutter displays and (c) high clutter displays (*p*-values appear in parentheses).

Table 6. Comparison of correlations of the Comprehensive Measure (CM) of clutter with performance outcomes and the strength of association of existing clutter measures (Clutter Rating (CR), edge density (ED) and Reduced Clutter Score (RCS)) with performance.

Driving response	Clutter level	Global			Local		
		CR	ED	RCS	CR	ED	RCS
Lane Deviation	All	+	+	+	0	–	0
	Low	+	+	+	0	–	0
	High	0	0	0	0	0	0
Speed Deviation	All	0	0	0	0	0	0
	Low	+	+	+	0	0	0
	High	0	0	0	0	0	0

Notes: A '+' indicates that the CM exhibited a stronger correlation with performance than the existing measure identified in the top row; a '–' indicates that the CM exhibited a weaker correlation than the measure at the top row; and a '0' indicates that both measures exhibited insignificant correlations.

Conclusions

The primary objectives of this research were to formulate and test a new comprehensive measure of display clutter accounting for display data features as well as user knowledge states, weighted by visual attention allocation patterns. Sensitivity analysis revealed stronger correlations of the CM with user task performance when greater weight was assigned to physical display factors vs. user perceptions of clutter. The CM generally exhibited stronger correlations with driving task performance than other existing measures of clutter, particularly for the low clutter displays. This finding indicates that the pattern of visual attention allocation, based on physical display features, has a greater influence on display user performance than the pattern of attention as a result of user task performance goals.

Given that the strongest correlations of the CM with display user performance were exhibited for the low clutter displays, the new measure could be a valid and useful tool for display designers using sound human factors principles to develop a series of low clutter prototypes, as it may be more sensitive to distinguishing between well-designed and very-well-designed displays. Use of the comprehensive measure in other domains may require different indicators of user knowledge, display feature content or visual attention allocation metrics. Furthermore, additional sensitivity analysis is needed to tune the weights associated with these variables, including the potential for other variables reflecting top-down and bottom-up contributors to maximise utility for explaining task performance.

Limitations and future work

There are several aspects of the experiment that may limit the generalisability of results to other task domains or other display user goal states. First, the driving simulation was paused during each test trial in order to collect driver subjective perceptions of AOI and global display clutter; it is possible that real-time collection of ratings could have made the simulation more realistic, potentially altering perceptions of display clutter, visual attention allocation and/or driving performance. Different simulated driving scenarios (e.g. rural freeway driving), presenting less demanding vehicle control requirements and allowing participants more time to scan the navigation display, might have made our results more sensitive to the display clutter manipulation. Finally, the comprehensive measure was formulated under the assumption that there is a linear relationship between clutter and performance, as reported by Wickens, Kroft, and Yeh (2000). Considering more display configurations (similar to the approach taken by Kim et al. [2011] and Kaber et al. [2013]) may reveal a higher order relationship.

Acknowledgements

The research was conducted while Carl Pankok, Jr. was enrolled as a graduate student in the Industrial & Systems Engineering Department at North Carolina State University; he is currently a research scientist at the Drexel University College of Computing and Informatics. The opinions expressed in this report are those of the authors and do not necessarily reflect the views of NIOSH.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

Carl Pankok, Jr. and David Kaber's efforts on this research were supported by a grant from the U.S. National Institute for Occupational Safety and Health [NIOSH: grant number 2 T42 OH008673-08].

References

Alexander, Amy L., David B. Kaber, Sang-Hwan Kim, Emily M. Stelzer, Karl Kaufmann, and Lawrence J. Prinzel, III. 2012. "Measurement and Modeling of Display Clutter in Advanced Flight Deck Technologies." *The International Journal of Aviation Psychology* 22 (4): 299–318. doi:10.1080/10508414.2012.718233.

Alexander, Amy L., Emily M. Stelzer, Sang-Hwan Kim, and David B. Kaber. 2008. "Bottom-up and Top-down Contributors to Pilot Perceptions of Display Clutter in Advanced Flight Deck Technologies." Paper presented at the Proceedings of the Human Factors and Ergonomics Society Annual Meeting, New York, NY, USA.

Asher, Matthew F., David J. Tolhurst, Tom Troscianko, and Iain D. Gilchrist. 2013. "Regional Effects of Clutter on Human Target Detection Performance." *Journal of Vision* 13 (5): 1–15. doi:10.1167/13.5.25.

Beck, Melissa R., Maura C. Lohrenz, and J. Gregory Trafton. 2010. "Measuring Search Efficiency in Complex Visual Search Tasks: Global and Local Clutter." *Journal of Experimental Psychology: Applied* 16 (3): 238–250. doi:10.1037/a0019633.

Bravo, Mary J., and Hany Farid. 2008. "A Scale Invariant Measure of Clutter." *Journal of Vision* 8 (1): 643–652. doi:10.1167/8.1.23.

Chen, Xiang-Mei, Liping Gao, Z. H. Wei, and Q. F. Li. 2007. "Correlation between Driver's Factors and Traffic Accident Rate." *Beijing Gongye Daxue Xuebao* 33 (7): 697–701.

Doyon-Poulin, Philippe, Jean-Marc Robert, and Benoit Ouellette. 2012. "Review of Visual Clutter and its Effects on Pilot Performance: A New Look at Past Research." Paper presented at the 2012 IEEE/AIAA 31st Digital Avionics Systems Conference (DASC), Williamsburg, VA, USA.

Giesbrecht, Francis G., and Marcia L. Gumpertz. 2004. *Planning, Construction, and Statistical Analysis of Comparative Experiments*. Hoboken, NJ: Wiley.

Henderson, John M., Myriam Chanceaux, and Tim J. Smith. 2009. "The Influence of Clutter on Real-World Scene Search: Evidence from Search Efficiency and Eye Movements." *Journal of Vision* 9 (1): 1–8. doi:10.1167/9.1.32.

Holmqvist, K., M. Nystrom, R. Andersson, R. Dewhurst, H. Jarodzka, and J. van de Weijer. 2011. *Eye Tracking: A Comprehensive Guide to Methods and Measures*. New York: Oxford University Press.

Kaber, David B., Amy L. Alexander, Emily M. Stelzer, Sang-Hwan Kim, Karl Kaufmann, and Simon Hsiang. 2008. "Perceived Clutter in Advanced Cockpit Displays: Measurement and Modeling with Experienced Pilots." *Aviation, Space, and Environmental Medicine* 79 (11): 1007–1018. doi:10.3357/ASEM.2319.2008.

Kaber, David, Karl Kaufmann, Amy L. Alexander, Sang-Hwan Kim, James T. Naylor, Lawrence J. Prinzel, III, Carl Pankok Jr, and Guk-Ho Gil. 2013. "Testing and Validation of a Psychophysically Defined Metric of Display Clutter." *Journal of*

Aerospace Information Systems 10 (8): 359–368. doi:10.2514/1.I010048.
Kennedy, Robert S., Norman E. Lane, Kevin S. Berbaum, and Michael G. Lilienthal. 1993. "Simulator Sickness Questionnaire: An Enhanced Method for Quantifying Simulator Sickness." *The International Journal of Aviation Psychology* 3 (3): 203–220. doi:10.1207/s15327108ijap0303_3.
Kim, Sang-Hwan, Lawrence J. Prinzel, David B. Kaber, Amy L. Alexander, Emily M. Stelzer, Karl Kaufmann, and Theo Veil. 2011. "Multidimensional Measure of Display Clutter and Pilot Performance for Advanced Head-up Display." *Aviation, Space, and Environmental Medicine* 82 (11): 1013–1022. doi:10.3357/ASEM.3017.2011.
Lim, Ji Hyoun, Omer Tsimhoni, and Yili Liu. 2010. "Investigation of Driver Performance with Night Vision and Pedestrian Detection Systems—Part I: Empirical Study on Visual Clutter and Glance Behavior." *IEEE Transactions on Intelligent Transportation Systems* 11 (3): 670–677. doi:10.1109/TITS.2010.2049843.
Lohrenz, Maura C., and Melissa R. Beck. 2010. "Evidence of Clutter Avoidance in Complex Scenes." Paper presented at the Proceedings of the Human Factors and Ergonomics Society Annual Meeting, San Francisco, CA, USA.
Lohrenz, Maura C., J. Gregory Trafton, Melissa R. Beck, and Marlin L. Gendron. 2009. "A Model of Clutter for Complex, Multivariate Geospatial Displays." *Human Factors: The Journal of the Human Factors and Ergonomics Society* 51 (1): 90–101. doi:10.1177/0018720809333518.
Marr, David, and Ellen Hildreth. 1980. "Theory of Edge Detection." *Proceedings of the Royal Society of London B: Biological Sciences* 207 (1167): 187–217.
McCrobie, Daniel. 2000. "Visual Display Screen Clutter Metrics." Paper presented at the Proceedings of the Human Factors and Ergonomics Society Annual Meeting, San Diego, CA, USA.
Metz, Barbara, Nadja Schömig, and Hans-Peter Krüger. 2011. "Attention during Visual Secondary Tasks in Driving: Adaptation to the Demands of the Driving Task." *Transportation Research Part F: Traffic Psychology and Behaviour* 14 (5): 369–380. doi:10.1016/j.trf.2011.04.004.
Moacdieh, Nadine M., and Nadine B. Sarter. 2012. "Eye Tracking Metrics: A Toolbox for Assessing the Effects of Clutter on Attention Allocation." Paper presented at the Proceedings of the Human Factors and Ergonomics Society Annual Meeting, Boston, MA, USA.
Moacdieh, Nadine, and Nadine Sarter. 2015. "Display Clutter a Review of Definitions and Measurement Techniques." *Human Factors: The Journal of the Human Factors and Ergonomics Society* 57 (1): 61–100. doi:10.1177/0018720814541145.
National Center for Statistics and Analysis. 2016. "2015 Motor Vehicle Crashes: Overview." Edited by National Highway Traffic Safety Administration. Washington, DC: National Highway Traffic Safety Administraion.
National Highway Traffic Safety Administration. 2017. "Driver Electronic Device Use in 2016. Washington, DC: United States Department of Transportation. https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/812426
Neider, Mark B., and Gregory J. Zelinsky. 2011. "Cutting through the Clutter: Searching for Targets in Evolving Complex Scenes." *Journal of Vision* 11 (14): 1–16. doi:10.1167/11.14.7.
O'Hara, Stephen, and Nathan Dwyer. 2007. "An Agent-based Approach to Decluttering the Interfaces of Multi-UAV Command and Control Systems." Paper presented at the Defense and Security Symposium, Orlando, FL, USA.
Pankok, Carl John. 2015. "A Knowledge and Data-driven Integrated Measure of Display Clutter with Attention Allocation Factors." Doctoral diss., North Carolina State University.
Rosenholtz, Ruth, Yuanzhen Li, Jonathan Mansfield, and Zhenlan Jin. 2005. "Feature Congestion: A Measure of Display Clutter." Paper presented at the Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Portland, OR, USA.
Rosenholtz, Ruth, Yuanzhen Li, and Lisa Nakano. 2007. "Measuring Visual Clutter." *Journal of Vision* 7 (2): 1–22. doi:10.1167/7.2.17.
St. John, Mark, Harvey S. Smallman, Daniel I. Manes, Bela A. Feher, and Jeffrey G. Morrison.2005. "Heuristic Automation for Decluttering Tactical Displays." *Human Factors: The Journal of the Human Factors and Ergonomics Society* 47 (3): 509–525. doi:10.1177/154193120404800330.
Transportation Research Board. 2010. "Highway Capacity Manual." Washington, D.C.: Transportation Research Board.
van Orden, Karl F., Wendy Limbert, Scott Makeig, and Tzyy-Ping Jung. 2001. "Eye Activity Correlates of Workload during a Visuospatial Memory Task." *Human Factors: The Journal of the Human Factors and Ergonomics Society* 43 (1): 111–121. doi:10.1518/001872001775992570.
Wang, Ying, Bruce Mehler, Bryan Reimer, Vincent Lammers, Lisa A. D'Ambrosio, and Joseph F. Coughlin. 2010. "The Validity of Driving Simulation for Assessing Differences between in-Vehicle Informational Interfaces: A Comparison with Field Testing." *Ergonomics* 53 (3): 404–420. doi:10.1080/00140130903464358.
Wickens, Christopher D., Paul Kroft, and Michelle Yeh. 2000. "Data Base Overlay in Electronic Map Design: Testing a Computational Model." Paper presented at the Proceedings of the Human Factors and Ergonomics Society Annual Meeting, San Diego, CA, USA.
Wierwille, Walter W. 1993. "Visual and Manual Demands of In-Car Controls and Displays." In *Automotive Ergonomics*, edited by B. Peacock and W. Karwowski, 299–320. London: Taylor & Francis.
Xu, Dejiang, and Zelin Shi. 2013. "DSIM: A DisSIMilarity-based Image Clutter Metric for Targeting Performance." *IEEE Transactions on Image Processing* 22 (10): 4108–4122. doi:10.1109/TIP.2013.2270112.
Yu, Chen-Ping, Dimitris Samaras, and Gregory J. Zelinsky. 2014. "Modeling Visual Clutter Perception Using Proto-object Segmentation." *Journal of Vision* 14 (7): 1–16. doi:10.1167/14.7.4.

Taylor & Francis
Taylor & Francis Group

Ergonomics

ISSN: 0014-0139 (Print) 1366-5847 (Online) Journal homepage: http://www.tandfonline.com/loi/terg20

An integrated measure of display clutter based on feature content, user knowledge and attention allocation factors

Carl Pankok Jr. & David B. Kaber

To cite this article: Carl Pankok Jr. & David B. Kaber (2018) An integrated measure of display clutter based on feature content, user knowledge and attention allocation factors, Ergonomics, 61:5, 682-696, DOI: 10.1080/00140139.2017.1401669

To link to this article: https://doi.org/10.1080/00140139.2017.1401669

Accepted author version posted online: 03 Nov 2017.
Published online: 14 Nov 2017.