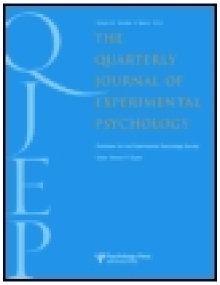
This article was downloaded by: [Duke University Libraries]
On: 12 December 2013, At: 12:20
Publisher: Routledge
Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



The Quarterly Journal of Experimental Psychology
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/pqje20

Different predictors of multiple-target search accuracy between nonprofessional and professional visual searchers

Adam T. Biggs[a] & Stephen R. Mitroff[a]

[a] Department of Psychology & Neuroscience, Center for Cognitive Neuroscience, Duke University, Durham, NC, USA

Accepted author version posted online: 22 Nov 2013.Published online: 09 Dec 2013.

To cite this article: Adam T. Biggs & Stephen R. Mitroff , The Quarterly Journal of Experimental Psychology (2013): Different predictors of multiple-target search accuracy between nonprofessional and professional visual searchers, The Quarterly Journal of Experimental Psychology, DOI: 10.1080/17470218.2013.859715

To link to this article: http://dx.doi.org/10.1080/17470218.2013.859715

PLEASE SCROLL DOWN FOR ARTICLE

Taylor & Francis makes every effort to ensure the accuracy of all the information (the “Content”) contained in the publications on our platform. However, Taylor & Francis, our agents, and our licensors make no representations or warranties whatsoever as to the accuracy, completeness, or suitability for any purpose of the Content. Any opinions and views expressed in this publication are the opinions and views of the authors, and are not the views of or endorsed by Taylor & Francis. The accuracy of the Content should not be relied upon and should be independently verified with primary sources of information. Taylor and Francis shall not be liable for any losses, actions, claims, proceedings, demands, costs, expenses, damages, and other liabilities whatsoever or howsoever caused arising directly or indirectly in connection with, in relation to or arising out of the use of the Content.

This article may be used for research, teaching, and private study purposes. Any substantial or systematic reproduction, redistribution, reselling, loan, sub-licensing, systematic supply, or distribution in any form to anyone is expressly forbidden. Terms & Conditions of access and use can be found at http://www.tandfonline.com/page/terms-and-conditions

THE QUARTERLY JOURNAL OF EXPERIMENTAL PSYCHOLOGY, 2013
http://dx.doi.org/10.1080/17470218.2013.859715

Routledge
Taylor & Francis Group

Different predictors of multiple-target search accuracy between nonprofessional and professional visual searchers

Adam T. Biggs and Stephen R. Mitroff

Department of Psychology & Neuroscience, Center for Cognitive Neuroscience, Duke University, Durham, NC, USA

Downloaded by [Duke University Libraries] at 12:20 12 December 2013

Visual search, locating target items among distractors, underlies daily activities ranging from critical tasks (e.g., looking for dangerous objects during security screening) to commonplace ones (e.g., finding your friends in a crowded bar). Both professional and nonprofessional individuals conduct visual searches, and the present investigation is aimed at understanding how they perform similarly and differently. We administered a multiple-target visual search task to both professional (airport security officers) and nonprofessional participants (members of the Duke University community) to determine how search abilities differ between these populations and what factors might predict accuracy. There were minimal overall accuracy differences, although the professionals were generally slower to respond. However, the factors that predicted accuracy varied drastically between groups; variability in search consistency—how similarly an individual searched from trial to trial in terms of speed—best explained accuracy for professional searchers (more consistent professionals were more accurate), whereas search speed—how long an individual took to complete a search when no targets were present—best explained accuracy for nonprofessional searchers (slower nonprofessionals were more accurate). These findings suggest that professional searchers may utilize different search strategies from those of nonprofessionals, and that search consistency, in particular, may provide a valuable tool for enhancing professional search accuracy.

Keywords: Visual search; Individual differences; Expertise; Multiple targets.

Visual search, the act of looking for targets amongst distractors, is a common cognitive task performed countless times each day. Individuals conduct visual searches in a wide variety of settings, including mundane tasks such as finding car keys and highly important tasks such as radiologists screening for cancerous tumours. Although accurate search is useful in almost any context, accuracy is essential for many professional searchers as missed targets can have life-threatening consequences. As

---

Correspondence should be addressed to Adam T. Biggs, Center for Cognitive Neuroscience, B203 LSRC, Box 90999, Duke University, Durham, NC 27708, USA. E-mail: adam.biggs@duke.edu

We thank Michael Zunk, Hung Nguyen, Bob Kingan, Kellis Turner, Ken Macdonald, Richard Menard, and Everett Vaughn for making it possible to conduct research at Raleigh-Durham International airport (RDU). We especially thank Dave Evans and Stuart Campbell for their extraordinary efforts to support the research at RDU. We also thank Stephen Adamo and Emma Wu Dowd for helpful comments, and Elise Darling and undergraduate research assistants for assistance with data collection at Duke University.

This work was partially supported by the Army Research Office [number 54528LS] and partially through a subcontract with the Institute for Homeland Security Solutions, a research consortium sponsored by the Resilient Systems Division in the Department of Homeland Security (DHS). This material is based upon work supported by the DHS [contract number HSHQDC-08-C-00100]. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the official policy or position of DHS or of the U.S. Government. The study is approved for public release.

© 2013 The Experimental Psychology Society

Downloaded by [Duke University Libraries] at 12:20 12 December 2013

such, it is important to understand the difficulties involved in professional searches and how various factors might predict variability in search accuracy. The present study was designed to assess how professional searchers might differ from nonprofessional searchers when faced with a particular challenge common to complex visual searches—the presence of multiple targets.

Laboratory-based visual search tasks often have only one target present, and a given trial ends as soon as that target is found (e.g., Chun & Wolfe, 1996; Eckstein, 2011; Nakayama & Martini, 2011). However, real-world searches often can have more than one target present at a time; for example, radiographs could have a fracture *and* a tumour, or X-rays of carry-on luggage could have a water bottle *and* a gun. In some scenarios, finding a particular type of target is sufficient to end the search (e.g., you only need to find one bomb in a carry-on bag to know that something is awry) whereas other target items are more innocuous (e.g., an oversized toothpaste tube in a carry-on bag), and a searcher will probably continue looking for additional targets after detection. It has been demonstrated that finding one target during a search can negatively affect the ability to find other targets in the same search; this particular error—missing a second target after finding a first—is known as "satisfaction of search" (Smith, 1967; Tuddenham, 1962).

Satisfaction of search is a pervasive source of error, thought to account for one fifth to one third of misses under certain conditions in radiology (Berbaum, Franken, Caldwell, & Schartz, 2010; Krupinski, 2010) and a potential majority of errors in diagnostic medicine (Kuhn, 2002; Voytovich, Rippey, & Suffredini, 1985). Several theories have been proposed regarding the underlying cause of these errors (Berbaum, 2012). For example, the original explanation—and the source of the name—explained these errors through premature search termination due to the searcher becoming "satisfied" with the underlying meaning of the scan after finding the first target (e.g., Smith, 1967; Tuddenham, 1962). Other theories have explained satisfaction of search errors through a "resource depletion" account wherein newly created memory representations of found targets take away available cognitive resources for subsequent searches (e.g., Cain & Mitroff, 2013), or a "perceptual set" account wherein searchers miss additional targets that are not perceptually similar to the found target (Berbaum et al., 2010).

The original "satisfaction" explanation for satisfaction of search errors cannot account for the entirety of the phenomenon as participants generally do continue searching after finding a first target (Berbaum, Dorfman, Franken, & Caldwell, 2000; Berbaum et al., 1991; Berbaum et al., 2010; Fleck et al., 2010). As several alternative explanations probably contribute to multiple-target search misses, the satisfaction of search name is an unfortunate misnomer. As such, we have rebranded the base phenomenon as the "subsequent search misses" (SSM) effect (Adamo, Cain, & Mitroff, 2013). SSM errors can occur for any number of mechanistic reasons, above and beyond the original "satisfaction" proposal (e.g., Adamo et al., 2013; Cain & Mitroff, 2013), so it is important to move away from a misleading label.

The goal of the current study is to better understand the mechanisms of multiple-target visual search. From an academic point of view, this can provide valuable insight into the nature of visual search and attention. Single-target visual search has provided an invaluable tool for understanding human cognition, and multiple-target search (and the accompanying errors) can provide additional evidence that could not be collected in a single-target paradigm. From a practical point of view, reducing SSM errors can literally save lives. Recent efforts have increased the research-related focus on SSM both in academic radiology (e.g., Berbaum, 2012) and in cognitive psychology (e.g., Fleck et al., 2010), but much remains unknown. To better understand multiple-target search, we compared professional and nonprofessional searchers on the same multiple-target search task to determine whether similar or different factors underlie search accuracy across these two groups. This study builds upon two recent projects from our lab (Biggs, Cain, Clark, Darling, & Mitroff,

Downloaded by [Duke University Libraries] at 12:20 12 December 2013

2013; Clark, Samei, Baker, & Mitroff, 2011) that allow for a carefully directed look at how and why targets are missed in multiple-target visual searches.

Predicting accuracy in a single-target visual search task

Often visual search studies focus on accuracy or response time—the tasks are designed such that displays are briefly presented, and accuracy is the primary dependent measure, or targets are easily discerned once fixated, and the time it takes to find the targets is the primary dependent measure. We recently employed a single-target visual search task, which required significant time to find targets while simultaneously producing considerable variability in detection accuracy (Biggs et al., 2013). As such, we were able to demonstrate that variability in accuracy could be predicted largely by two simple factors—participants' search speed, and the consistency in their search speed.

Our previous work defined *search speed* as how long an individual took, on average, to complete a search when no targets were present. This measure provides an assessment of search diligence—how long a searcher is willing to spend on a trial without finding a target. We defined *consistency* as the variability in how long searchers spent looking on trials in which they did not find a target (Biggs et al., 2013). That is, this measure assesses how consistent a searcher was, from trial to trial, in terms of the time they were willing to dedicate to the search process. By focusing on trials in which searchers terminated the search of their own volition without finding a target (i.e., not due to timeouts or successful detection), this timing-based consistency measure provides insight into search strategy (Biggs et al., 2013).

These two basic response time factors—speed and consistency—were able to account for a significant amount of the participants' accuracy performance, and these factors were obtained through easily collected response timing measures. Both proved to be excellent predictors of performance—in some cases, cumulatively accounting for more than 60% of the variability in accuracy across individuals. Interestingly though, they were differently predictive for professional versus nonprofessional searchers. Performance was compared between working Transportation Security Administration (TSA) officers and members of the Duke University community who did not have any specialized training in visual search. Search speed was the primary predictor of search accuracy for nonprofessional searchers and for the least experienced (i.e., early-career) professional searchers, whereas consistency was the primary predictor for the most experienced professional searchers (Biggs et al., 2013). Given that differences arose between professional and nonprofessional searchers for a simple single-target search task, there is potential promise in extending this analysis to the much richer environment of multiple-target visual search. In particular, consistency may be especially important for multiple-target searches because it could help ensure a thorough search following a found target.

Multiple-target search and expertise

We have recently begun an initial comparison of performance between professional and nonprofessional searchers on a multiple-target search task, comparing trained radiologists to nontrained searchers (Clark et al., 2011). This initial step has proven to be a fruitful endeavour as we have found significant differences, including that professionals took significantly longer to complete the search, and professionals were less influenced by visual salience (i.e., how difficult a target was to find) than were nonprofessionals. While the data from this earlier study are intriguing, there were too few participants (9 radiologists) to meaningfully assess potentially compelling differences in predictors of search accuracy. As such, the goal of the present study is to leverage the strengths of both the Biggs et al. (2013) and the Clark et al. (2011) studies by utilizing the multiple-target search paradigm of Clark et al. along with the individual differences and accuracy prediction methods of Biggs et al. (2013). Our primary interests are whether different factors could explain search

Downloaded by [Duke University Libraries] at 12:20 12 December 2013

accuracy between professional and nonprofessional searchers, and whether consistency can benefit accuracy for multiple-target search as well as single-target search.

EXPERIMENTAL STUDY

Method

Professional participants

Seventy-six members of the Transportation Security Administration participated as a part of their regular workday.[1] Four participants were excluded from analysis (see Data Filtering in the Results section below). The 72 participants remaining in the analysis reported their ages via age range options (18–25 years: $N = 2$; 26–34 years: $N = 18$; 35–49 years: $N = 27$; 50–65 years: $N = 25$).

Nonprofessional participants

One hundred and eight members of the Duke University community participated for $10 or partial fulfilment of a class requirement. Five participants were excluded from analysis (see Data Filtering in the Results section). The 103 participants remaining in the analysis ranged in age from 18–22 years ($M = 19.54$ years, $SD = 1.15$).

Stimuli

The stimuli were based on prior studies (e.g., Cain, Dunsmoor, LaBar, & Mitroff, 2011; Clark et al., 2011; Fleck, Samei, & Mitroff, 2010). Each display included 25 total items arranged upon an invisible 8 × 7 grid, with each item randomly offset 0–10 pixels from perfect grid alignment (Figure 1). Targets were perfect "T" shapes and appeared in one of two salience levels (high salience: 57–65% black; low salience: 22–45% black). Distractors were non-T shapes drawn from the same salience ranges. Each item was composed of two rectangles (approximate width of 0.3° of visual angle and length of 1.0° with participants seated approximately 57 cm from the screen without head restraint) oriented perpendicularly and slightly separated; each item was 1.3° × 1.3° at its widest point. Each item appeared in one of four possible rotations, and all were on a background of grey "clouds" (4–37% black).

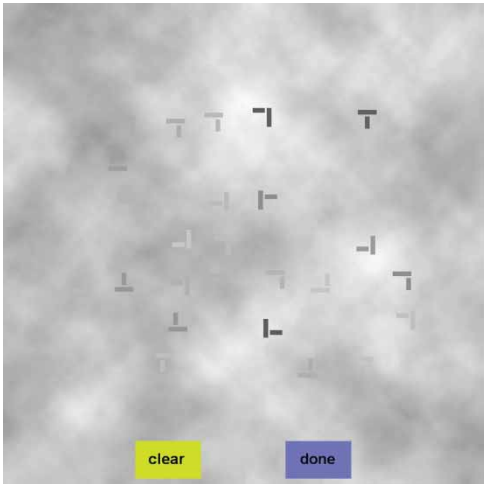


Figure 1. *Sample display. Targets were perfect "T" shapes, and two are present here (one high-salience target in the upper right and one low-salience target in the lower left). To view this figure in colour, please visit the online issue of the Journal.*

Apparatus

Professional visual searchers recruited from the TSA were tested in a private room at Raleigh-Durham International Airport (RDU) on Dell Vostro 260 computers and 23.6-inch widescreen LCD monitors. Nonprofessional visual searchers from the Duke community were tested at the Duke Visual Cognition Laboratory on the Duke University campus with Dell Inspiron computers with 20-inch CRT monitors. The computer displays were adjusted so that both the RDU lab and Duke testing stations presented the same physical display sizes to participants. All testing stations used Matlab software (The MathWorks, Natick, MA) and the Psychophysics Toolbox Version 3.0.8 (Brainard, 1997; Kleiner, Brainard, & Pelli,

---

[1] Because TSA officers participated as part of their regular workday, they were given the option of allowing their data to be used only for TSA purposes or for both TSA and research purposes. There were an additional 9 professional participants who were tested, but did not provide consent for their data to be used for research. As such, data from these individuals are not included in the analyses.

Downloaded by [Duke University Libraries] at 12:20 12 December 2013

2007; Pelli, 1997) for experimental presentation and data collection.

Procedure

Each participant completed a practice block of 25 trials and 255 experimental trials. The experimental trials were divided as follows and presented in a unique random order for each participant: 125 with a single, high-salience target; 40 with a single, low-salience target; 40 with both a low-salience and a high-salience target (dual-target trials); and the remaining 50 of trials with no targets. The trial numbers are divided accordingly based upon previous work (e.g., Fleck et al., 2010) that has shown this distribution to effectively elicit SSM effects. Participants were instructed to make a mouse click on each of the targets they found and then click a button marked "Done" when they had completed their search. A click within a 35-pixel radius of the centre of an item was considered a click on that item. The "Done" button appeared once an item had been clicked or after 3s had elapsed without a click. This delay was introduced to prevent the target-absent response from becoming prepotent and executed habitually (Fleck & Mitroff, 2007). If the "Done" button was not clicked within 15 s, the trial was terminated, and a message appeared encouraging the participant to search faster. Feedback on misses and false alarms was given during the practice block, but no feedback was given during the experimental trials. Displays also included a "Clear" button at the bottom of the screen, which allowed participants to undo any previous mouse clicks made by accident during that particular trial. Any trials with a "Clear" button click were excluded from the analyses.

Planned analyses

The visual search task employed here was relatively difficult, which resulted in variability for both accuracy and response timing. This variability allowed us to use response time variables as predictors of accuracy in regression analyses (e.g., Biggs et al., 2013). For the regression analyses reported below, we sought to explain the variability in accuracy through two variables: search speed and search consistency. Search speed is the average time spent before voluntarily ending a trial when no target was present and therefore represents how long participants took to complete their search without successfully finding a target.

Search consistency represents the variability in response time for trials where participants voluntarily ended the trial without finding the maximum number of targets. For the present experiment, the consistency measure was calculated from trials where participants made fewer than two mouse clicks (see Biggs et al., 2013, for a consistency calculation for a single-target visual search tasks that involved multiple set sizes). Given a maximum of two targets, this approach omits trials where the participant ended the trial after finding two targets, or if the participant ended the trial after finding a target and making a false alarm (i.e., believing that he or she had located the maximum number of targets). Consistency was calculated by taking the standard deviation of response times and dividing by the average of response times for trials where the participant voluntarily terminated search. The standard deviation represents the variability in the decision time to volitionally end a trial without finding the maximum number of targets and dividing by the average response time accounts for the difference in potential variance due to longer response times.

For dual-target trials, we can assess how finding a first target affects finding a second target (i.e., the SSM effect). Previous work from our lab that used a similar paradigm to the one employed here assessed SSM errors by comparing accuracy for low-salience targets on single-target trials versus accuracy for low-salience targets on dual-target trials when the high-salience target was also found (e.g., Fleck et al., 2010), or, more conservatively, when the high-salience target was found first (e.g., Clark, Cain, Adcock, & Mitroff, 2013). That is, low-salience target accuracy is compared between two situations—when it was the only target present versus when there was another found target in the same search array. Single-target accuracy provides an appropriate baseline for comparison as it represents the likelihood of a searcher identifying that particular target. The second component is more specific than general accuracy on dual-target trials because

Downloaded by [Duke University Libraries] at 12:20 12 December 2013

it must incorporate and specifically address whether accuracy changes following a target being found. Significant SSM errors occur when the difference between these two accuracy rates is significantly greater than zero.

Prior SSM investigations (Clark et al., 2011; Fleck et al., 2010) omitted any dual-target trials from SSM analyses if the first target found was low salience because there were too few trials to allow for a viable analysis. However, in the current study, professionals and nonprofessionals differed on how frequently the high-salience target was found before the low-salience target (see Results). This difference is potentially problematic for the standard SSM approach as significantly more trials would enter the calculation for nonprofessional searchers than professional searchers. To address this discrepancy, we modified the SSM calculation to also include those trials where the low-salience target trial was found first. There were two calculations taken. First, we compared accuracy for low-salience targets on single-target trials to that for dual-target trials when the high-salience target was found first. Second, we compared accuracy for high-salience targets on single-target trials to that for dual-target trials when the low-salience target was found first. Then, each calculation (i.e., SSM for both high-salience and low-salience targets) was weighted according to the number of times a high-salience or low-salience target was found first. The result is a weighted SSM calculation that compares second target accuracy on dual-target trials to the corresponding baseline as established by single-target trial accuracy.

Results

Data filtering

All trials including either a timeout or a "Clear" button click were excluded from data analyses. The timeout trials prevented participants from voluntarily terminating search, whereas the "Clear" button press trials indicated that one or more of the mouse clicks was accidental (e.g., a motor error). Participants' data were excluded from further analyses if the excluded trial total (timeouts plus "Clear" button presses) was more than 3 standard deviations from their group mean. Data from one professional searcher and two nonprofessional searchers were excluded based upon contributing too few experimental trials.

Participants' data were also excluded if their accuracy on single-target trials was more than 3 standard deviations beyond the group mean, or if their dual-target accuracy was below 5% correct, suggesting that the participant was not performing a multiple-target search. Data from two professional searchers and one nonprofessional searcher were excluded due to single-target accuracy more than 3 standard deviations beyond the group mean. Data from one professional searcher and two nonprofessional searchers were excluded due to dual-target accuracy below 5% correct. Seventy-two professional and 103 nonprofessional searchers remained for data analyses.

Response timing metrics

See Table 1 for full results of the response time metrics. Professionals had significantly more trials reaching timeout ($M = 11.00$, $SE = 1.50$), than nonprofessionals ($M = 6.55$, $SE = 0.65$), $t(173) = 3.02$, $p < .01$, $d = .44$. There was no difference between groups on the "Clear" button press trials (professionals: $M = 4.68$/experiment, $SE = 0.65$; nonprofessionals: $M = 4.33$/experiment, $SE = 0.40$); $t(173) = 0.48$, $p = .63$. These two trial types —those including timeouts or "Clear" button presses—were omitted from further analyses.

Professionals took longer to search on average ($M = 9.72$s, $SE = 0.17$) than nonprofessionals ($M = 8.21$s, $SE = 0.16$), $t(173) = 6.33$, $p < .001$, $d = .98$. Given that the present search task included multiple possible targets, the slower overall rates could come from a longer search time before finding the first target, more time spent after finding a target, or both. Professionals took longer to find a first target ($M = 4.14$ s, $SE = 0.10$) than nonprofessionals ($M = 3.39$ s, $SE = 0.08$), $t(173) = 6.10$, $p < .001$, $d = .93$; additionally, professionals took longer to terminate search after finding a first target without finding a second target ($M = 5.50$ s, $SE = 0.15$) than did nonprofessionals ($M = 4.98$ s, $SE = 0.14$),

Downloaded by [Duke University Libraries] at 12:20 12 December 2013

Table 1. *Response timing metrics for professional and nonprofessional searchers*

Measure	*Group* *Professionals*	*Nonprofessionals*	*Comparison*
A. Timeouts	11.00 (1.50)	6.55 (0.65)	$t(173) = 3.02, p < .01^{*}$
B. Mean response time (s)	9.72 (0.17)	8.21 (0.16)	$t(173) = 6.33, p < .001^{**}$
C. 1st target response time (s)	4.14 (0.10)	3.39 (0.08)	$t(173) = 6.10, p < .001^{**}$
D. 1st-to-done response time (s)	5.50 (0.15)	4.98 (0.14)	$t(173) = 2.40, p = .02^{*}$
E. Consistency	0.20 (0.01)	0.24 (0.01)	$t(173) = 3.56, p < .001^{**}$
F. % of high salience found 1st	82.74 (1.30)	86.09 (1.00)	$t(173) = 2.07, p = .04^{*}$

Note: A: Average number of trials where the participant took more than provided time limit to respond; B: Average response time across all trials during the experiment; C: Average response time to correctly locate the first target during search; D: Average response time after finding a first target to voluntarily terminate search without finding a second target; E: How consistently an observer terminated search from trial-to-trial (lower values signify more consistent performance; see Method). F: Percentage of dual-target trials where the first target found was high salience versus low salience.

*indicates $p < .05$.

**indicates $p < .001$.

$t(173) = 2.40$, $p = .02$, $d = .38$. Therefore, professionals took more time than nonprofessionals to locate a target and more time after locating a target before voluntarily deciding to end their search.

We also observed differences in response timing that are suggestive of search strategy differences between the professionals and nonprofessionals. Professionals found the high-salience target first on a lower percentage of trials (82.74%, $SE = 1.30\%$) than did nonprofessionals ($M = 86.09\%$, $SE = 1.00\%$), $t(173) = 2.07$, $p = .04$, $d = .32$. Finally, we measured temporal consistency, or how similarly an observer spent searching in terms of time from trial to trial. Temporal consistency is calculated by taking the standard deviation of response time and dividing by the average response time (see the Planned Analyses of the Method section above for more details). Lower values indicate more consistent visual search, and professionals were more consistent ($M = 0.20$, $SE = 0.01$) than nonprofessionals ($M = 0.24$, $SE = 0.01$), $t(173) = 4.27$, $p < .001$, $d = .50$.

Accuracy metrics

See Table 2 for full accuracy results. Professionals were more likely to make a false alarm ($M = 9.63$/experiment, $SE = 1.18$) than nonprofessionals ($M = 5.32$/experiment, $SE = 0.89$), $t(173) = 2.96$, $p < .01$, $d = .45$. An accurate trial was defined as when the observer found all targets present (whether 0, 1, or 2) and made no false alarms. Professionals and nonprofessionals did not differ significantly on accuracy for single-target trials with a high-salience target (professionals: $M = 90.09\%$, $SE = 0.89$; nonprofessionals: $M = 90.32\%$, $SE = 0.65$), $t(173) = 0.43$, $p = .67$, or a low-salience target (professionals: $M = 56.95\%$, $SE = 1.84$; nonprofessionals: $M = 53.26\%$, $SE = 1.69$), $t(173) = 1.55$, $p = .12$, nor on accuracy for dual-target trials (professionals: $M = 46.84\%$, $SE = 1.94$; nonprofessionals: $M = 46.32\%$, $SE = 1.78$), $t(173) = 0.20$, $p = .84$.

Dual-target trials can also be analysed based upon the number of targets found rather than whether both targets were found; this approach provides a more nuanced analysis because it incorporates the difference between an error trial of finding zero targets versus an error trial of finding only one target. Similar to the above comparisons, the groups did not differ on the average number of found targets on dual-target trials (professionals: $M = 1.44$ targets/trial, $SE = 0.02$; nonprofessionals: $M = 1.43$ targets/trial, $SE = 0.02$), $t(173) = 0.22$, $p = .82$.

Dual-target trials also allow for an SSM assessment (see Planned Analyses section) as SSM errors represent a difference score between single-target

Table 2. *Accuracy metrics for professional and nonprofessional searchers*

Measure	*Professionals*	*Nonprofessionals*	*Comparison*
A. False alarms	9.63 (1.18)	5.32 (0.89)	$t(173) = 2.96, p < .01^{**}$
B. Single-target, high-salience (%)	90.09 (0.89)	90.32 (0.65)	$t(173) = 0.22, p = .83$
C. Single-target, low-salience (%)	56.95 (1.84)	53.26 (1.69)	$t(173) = 1.45, p = .15$
D. Dual-target trials (%)	46.84 (1.94)	46.32 (1.78)	$t(173) = .20, p = .84$
E. Targets found on dual-target trials	1.44 (0.02)	1.43 (0.02)	$t(173) = 0.22, p = .82$
F. Subsequent search misses (SSM; %)	12.70 (1.46)	8.72 (1.15)	$t(173) = 2.17, p = .03^{*}$

Group spans the Professionals, Nonprofessionals and Comparison columns.

Note: A: Average number of clicks made on a nontarget during the experiment; B: Accuracy for trials with only a high-salience target present; C: Accuracy for trials with only a low-salience target present; D: Accuracy for trials with both a high-salience and a low-salience target present; E: Average number of targets found on trials when two targets were present; F: Subsequent search misses (see text for more detail).

*indicates $p < .05$.

**indicates $p < .001$.

and dual-target accuracy. While the professionals and nonprofessionals did not significantly differ in single-target and dual-target performance, the professionals had a near significantly higher accuracy rate for the single-target low-salience targets. This manifests into a significant difference score such that with the weighted SSM measure, professionals showed a larger SSM effect ($M =$ 12.70%, $SE = 1.46$%) than nonprofessionals ($M = 8.72$%, $SE = 1.15$%), $t(173) = 2.17$, $p = .03$, $d = .33$.

Age

One potential concern about the current study is the significant age difference between the professional and nonprofessional populations. While we have previously shown that age is not a driving factor when comparing TSA officers to members of the Duke community (Biggs et al., 2013), this is nevertheless worth considering further. We cannot directly examine age differences upon performance among professional searchers because there is a confounding variable in that older participants have more experience at the TSA than younger participants: age and years of experience at TSA, $\chi^2(12, N = 72) = 42.67$, $p < .001$. However, we have enough participants with 8 to 11 years of experience at the TSA[2] to compare two different age ranges within the same level of experience: 35 to 49 years ($N = 18$), and 50 to 65 years ($N = 17$). This approach equates the two groups on the amount of experience with the TSA and allows us to examine differences solely due to age. These two groups did not differ on any of the response timing metrics we examined (timeouts: $p = .54$; average response time: $p = .17$; time to find the first target: $p = .63$; time taken to end a trial after finding a first target: $p = .22$; consistency: $p = .28$; or percentage of dual-target trials where the high salience target was found first: $p = .58$), nor did they differ on most of the accuracy metrics (false alarms: $p = .56$; high-salience single-target accuracy: $p = .66$; low-salience single-target accuracy: $p = .11$; SSM: $p = .47$). The one exception concerned accuracy on dual-target trials, where younger participants were more likely to find both targets than older participants (35–49 years: 51.45%; 50–65 years: 40.16%), $t(33) = 2.05$, $p = .05$, $d = .70$. There appears to be a minimal influence of age on the current results.

Predicting accuracy

A stepwise linear regression was conducted separately for each group (professionals and

---

[2] Professional experience is constrained by the creation of the TSA on November 19, 2001.

Downloaded by [Duke University Libraries] at 12:20 12 December 2013

Downloaded by [Duke University Libraries] at 12:20 12 December 2013

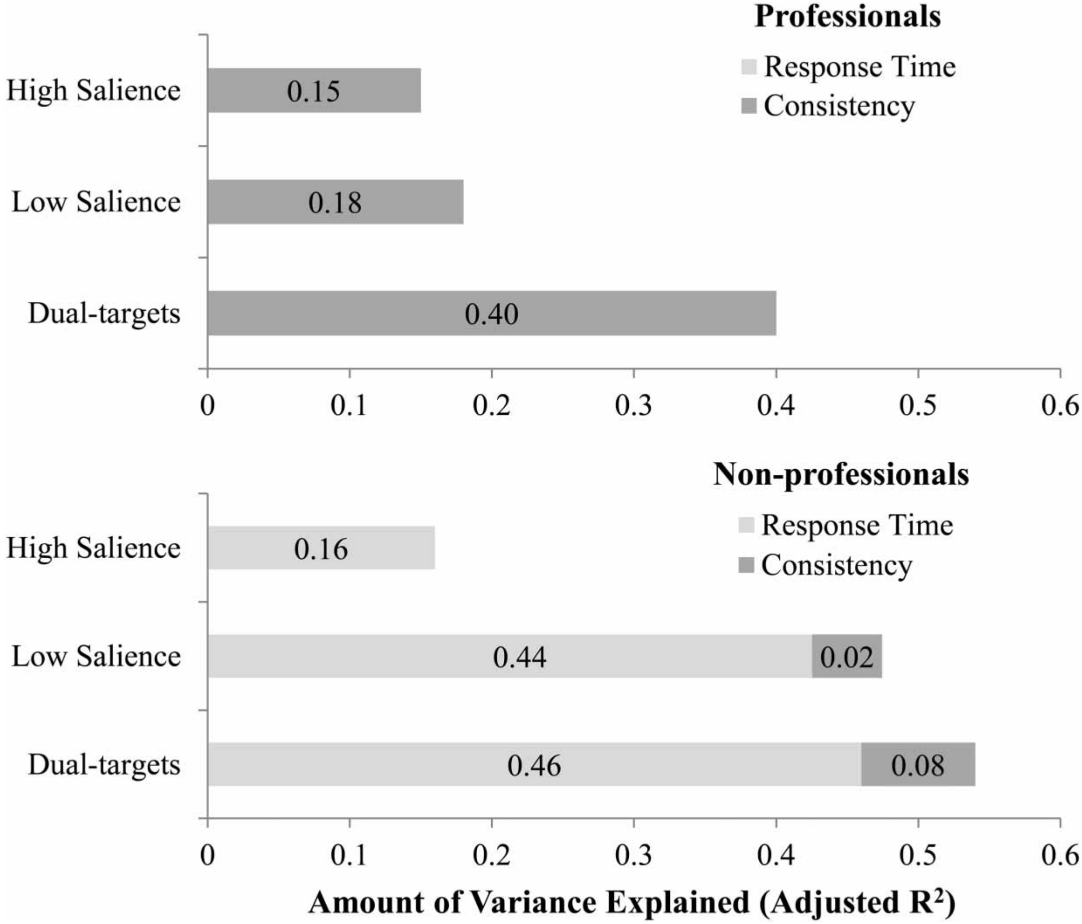


Figure 2. *Accuracy variance as explained by search speed (response time) and consistency for high-salience single-target trials, low-salience single-target trials, and dual-target trials.*

nonprofessionals) with two primary variables of interest: search speed and search consistency (see Planned Analyses for more details). Outliers were assessed based upon a Cook's *D* equal to or greater than 1 (Cook & Weisberg, 1982), but no data points were trimmed due to this criterion. Additionally, because search speed and consistency are calculated from similar response time metrics, we assessed collinearity diagnostics. No factor included in the following models had a variance inflation factor above 1.81 when two factors were present, which indicates that collinearity is not an issue for our regression models (O'Brien, 2007).

Professional visual searchers. As seen in Figure 2, consistency explained a significant amount of the accuracy variance for single-target trials, both for high-salience targets [adjusted (adj.) $R^2 = .15$, $F(1, 70) = 13.94$, $p < .001$; consistency: $\beta = -.407$, $t(71) = 3.73$, $p < .001$] and low-salience targets [adj. $R^2 = .18$, $F(1, 70) = 16.45$, $p < .001$; consistency: $\beta = -.436$, $t(71) = 4.06$, $p < .001$]. Consistency was also the only variable explaining accuracy on dual-target trials [adj. $R^2 = .40$, $F(1, 70) = 48.35$, $p < .001$; consistency: $\beta = -.639$, $t(71) = 6.95$, $p < .001$]. In all cases for the professionals, more consistent search resulted in increased accuracy, especially for multiple-target searches.

Downloaded by [Duke University Libraries] at 12:20 12 December 2013

Nonprofessional visual searchers.. As seen in Figure 2, response time was the sole predictor of accuracy for single-target trials with high-salience targets [adj. $R^2 = .16$, $F(1, 101) = 21.07$, $p < .001$; search speed: $\beta = .415$, $t(102) = 4.59$, $p < .001$]. For single-target trials with low-salience targets, both speed and consistency explained a significant amount of the accuracy variance for the nonprofessional visual searchers [adj. $R^2 = .46$, $F(2, 100) = 44.71$, $p < .001$]. Speed was the primary contributor to the model [$\Delta R^2 = .44$, $\beta = .518$, $t(102) = 5.30$, $p < .001$] with consistency as the secondary contributor [$\Delta R^2 = .02$, $\beta = -.223$, $t(102) = 2.28$, $p = .03$]. For dual-target trials, both speed and consistency explained a significant amount of the accuracy variance for the nonprofessional visual searchers (adj. $R^2 = .54$, $F(2, 100) = 59.76$, $p < .001$). Speed was the primary contributor to the model ($\Delta R^2 = .46$, $\beta = .432$, $t(102) = 4.76$, $p < .001$) with consistency as the secondary contributor ($\Delta R^2 = .08$, $\beta = -.376$, $t(102) = 4.14$, $p < .01$). In all cases, participants were more accurate as they took longer to complete their search, which suggests that nonprofessionals are exhibiting a classic speed–accuracy trade-off in performance. Consistency was only able to explain a small amount of variance above and beyond response time for single-target trials with a low-salience target and dual-target trials, and more consistent search (as represented by lower values, or a smaller standard deviation/response time) resulted in increased dual-target accuracy.

GENERAL DISCUSSION

Real-world search tasks contain a number of challenges that hinder accuracy, including the possible presence of multiple targets. For example, an individual trying to sneak contraband through a security screening checkpoint may have disassembled a weapon into multiple parts to make the assembled weapon harder to recognize, or a particular bag could contain both a water bottle *and* a gun. These possibilities are particularly troublesome for professional visual searchers where poor accuracy can have life-threatening consequences. As such, it is important to understand what factors can best predict accuracy. The present study compared whether professional and nonprofessional searchers differed in their performance on a search task that could contain multiple targets and whether the factors that best predicted accuracy differed between professional and nonprofessional searchers.

Professional searchers differed significantly from nonprofessional searchers on a number of response timing metrics. First, professional searchers took longer to complete their search than nonprofessional searchers, replicating the previous counterintuitive evidence that, rather than an expedited search process, professional visual searchers take longer to complete a visual search task unrelated to their domain-specific task (Biggs et al, 2013; Clark et al., 2011; Jackson, Clark, & Mitroff, 2013; Jackson, Mitroff, Clark, Proffit, Lee, & Nguyen, 2013). Our previous study also demonstrated that these slower overall search speeds did not vary based upon the years of experience conducting professional visual searches (Biggs et al., 2013). We were able to extend these findings in the current study by showing that professionals took longer both to find a first target and to voluntarily discontinue their search after having found a first target. This evidence suggests that the slower overall responses are probably due to a more diligent search process rather than any one component, such as slower motor responses, longer time to locate a single target, or longer search times after locating a first target. In addition, professional searchers were less likely than nonprofessionals to locate a high-salience target before a low-salience target on dual-target trials, which suggests that nonprofessionals are beginning with a more global search before trying to methodically find a low-salience target. This difference suggests that professional searchers may use different search strategies from those of nonprofessional searchers, even for searches that are not domain-specific.

Implications for multiple-target visual search

Search strategy may be especially important for searches that can contain more than one target. Multiple-target searches are vulnerable to a specific

Downloaded by [Duke University Libraries] at 12:20 12 December 2013

problem wherein a target is more likely to be missed if another target had already been found in the same array (Berbaum, 2012). This phenomenon (originally called "satisfaction of search", which we have relabelled "subsequent search misses"; SSM) can have dire implications for real-world searches. Interestingly, we observed a larger SSM effect for professional searchers than for the nonprofessionals. This was primarily driven by the fact that SSM errors represent a difference score between single-target and dual-target performance, and the professionals had a relatively higher single-target accuracy rate for the low-salience targets (thus making this difference score more robust). Professional searchers, particularly security officers, may show such a pattern of results if they are trained to act upon locating a single target. Fortunately, there are a number of ways to counteract such vulnerability. There could be multiple searchers with different assignments to better ensure that a second, different target is not missed (Menneer, Barrett, Phillips, Donnelly, & Cave, 2007), or the act of rerunning a specific search scenario such as rerunning a bag of airport luggage could limit subsequent search errors (Cain, Biggs, Darling, & Mitroff, 2013, submitted).

Predicting accuracy from response time measures

Although overall accuracy performance was similar between groups, the factors predicting performance differed drastically between professional and nonprofessional searchers. Search speed was the primary predictor of performance amongst nonprofessionals for single-target and multiple-target trials, whereas consistency was the primary predictor of performance amongst professionals for single-target and multiple-target trials. This notable difference between the groups raises an important question—why would consistency benefit professionals and not nonprofessionals? Our interpretation is that consistency is a learned skill, which can lead to marked benefits with sufficient practice. In most visual search tasks, searchers need to remember where they have and have not already looked (Beck, Peterson, & Vomela, 2006; Dickinson & Zelinsky, 2007; but see Horowitz & Wolfe, 1998). This burden may be particularly onerous in the complex searches often encountered by professional searchers, which in turn would require more cognitive resources to successfully complete the search. If the searcher consistently used the same search strategy, whatever it might be, there would be less need to remember where search has or has not already been conducted—fitting with evidence from a resource depletion theory of multiple-target search errors (e.g., Cain & Mitroff, 2013). Consistent search thus potentially frees up cognitive resources that would otherwise be spent maintaining where search has previously been conducted. The exact reason that consistency benefits search requires more study, but it does appear that professional searchers benefit from consistency more than nonprofessional searchers.

Benefits from utilizing a domain-general task

The present study compared performance between professional and nonprofessional searchers on a task that did not require domain-specific knowledge. Examining domain-specific stimuli, such as radiological scans or X-ray images of carry-on luggage, inherently requires knowledge and training to successfully complete the search, and directed training can clearly have an impact on within-domain performance. For example, previous efforts to improve accuracy in an X-ray image search akin to airport baggage screening have observed benefits by focusing on the influence of image-based effects (e.g., Schwaninger, 2003, 2005; Schwaninger, Hardmeier, & Hofer, 2004), or an observer's ability to recognize potential target items (McCarley, Kramer, Wickens, Vidoni, & Boot, 2004). Here we focused on search behaviours by minimizing any domain-specific knowledge through the use of a search task with simple targets. Whenever the searcher does not have a specific target to find (i.e., they are not provided with a single, well-defined target item), efficient search behaviours offer the most practical means to enhance search accuracy. Given that many professional search tasks involve underspecified targets, there could be substantial

Downloaded by [Duke University Libraries] at 12:20 12 December 2013

value in training the specific search behaviours used by professionals, especially since expertise performance can be enhanced through adaptation to task constraints (Ericsson & Lehmann, 1996). Additionally, we would not have been able to compare professional and nonprofessional searchers if we had used the domain-specific tasks for which our professional participants were trained; the nonprofessionals would not be adequately informed to perform a meaningful search of X-ray images of airport luggage.

Addressing age concerns

A possible concern with this study is that the results could potentially be influenced by age differences. Age is a known source of differences on cognitive tasks—for example, elderly participants typically show a strong speed–accuracy trade-off with slower, but more accurate responses (e.g., Rabbitt, 1979; Ratcliff, Thapar, & McKoon, 2007). However, some age differences may be specific to speed without an impact on cognitive processing—for example, when general slowing was accounted for in a comparison between younger and older adults, no age effects were found for the ability to utilize top-down control in a visual search task (Costello, Madden, Shepler, Mitroff, & Leber, 2010). Consistent with our previous investigation of TSA officers and members of the Duke community (Biggs et al., 2013), age differences cannot adequately explain the present results. First, there were very few differences on either response time or accuracy metrics between our groups matched on professional experience. Second, the significant differences were not observed in the direct accuracy comparisons between professional and nonprofessional searchers, but rather they were observed in what factors were most predictive of accuracy within a group. This result suggests a more complex influence than a simple speed–accuracy trade-off among the professional visual searchers, and one that is not simply explained by age.

Conclusions

Professional searches are conducted in real-world settings, and these searches are subject to a variety of factors that can significantly hinder accuracy; there could be multiple targets, targets that are difficult to find, and an unknown and variable set of potential targets. These problems exist within search tasks that demand the highest possible accuracy, which emphasizes the need to explore every possible means of improving performance. Consistency is one potential means of improving accuracy, particularly for professional searchers engaging in a multiple-target search. The current study found that the professional searchers who were more consistent in their search time were more accurate for both single- and dual-target trials, suggesting that consistency may be an especially important factor. Finding a target in a multiple-target visual search saps valuable cognitive resources for subsequent search performance (Cain & Mitroff, 2013), so it is vital to identify means to boost performance. The current results suggest that training consistency might provide just such a boost; if consistency benefits a searcher by reducing the cognitive burden necessary to maintain a search strategy, consistency training might offer a viable means to improve accuracy. Repeatedly executing a search strategy until the strategy can be executed with minimal effort might produce measurable benefits that can enhance search performance.

Original manuscript received 2 July 2013
Accepted revision received 7 October 2013
First published online 10 December 2013

REFERENCES

Adamo, S. H., Cain, M. S., & Mitroff, S. R. (2013). Self-induced attentional blink: A cause of errors in multiple-target search. *Psychological Science*. doi: 10.1177/0956797613497970

Beck, M. R., Peterson, M. S., & Vomela, M. (2006). Memory for where, but not what, is used during visual search. *Journal of Experimental Psychology: Human Perception and Performance*, *32*(2), 235–250.

Berbaum, K. S. (2012, February). Satisfaction of search experiments in advanced imaging. In *IS&T/SPIE electronic imaging* (pp. 82910 V-182910 V-10). International Society for Optics and Photonics.

Downloaded by [Duke University Libraries] at 12:20 12 December 2013

Berbaum, K. S., Dorfman, D. D., Franken, E. A. Jr, & Caldwell, R. T. (2000). Proper ROC analysis and joint ROC analysis of the satisfaction of search effect in chest radiography. *Academic Radiology*, *7*, 945–958.

Berbaum, K. S., Franken, E. A. Jr, Dorfman, D. D., Rooholamini, S. A., Coffman, C. E., Cornell, S. H., … & Smith, T. P. (1991). Time course of satisfaction of search. *Investigative Radiology*, *26*, 640–648.

Berbaum, K. S., Franklin, E. A. Jr, Caldwell, R. T., & Schartz, K. M. (2010). Satisfaction of search in traditional radiographic imaging. In E. Samei & E. Krupinski (Eds.), *The handbook of medical image perception and techniques* (pp. 107–138). Cambridge: University Press.

Biggs, A. T., Cain, M. S., Clark, K., Darling, E. F., Mitroff, S. R. (2013). Assessing visual search performance differences between transportation security administration officers and nonprofessional searchers. *Visual Cognition*, *21*, 330–352.

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, *10*, 433–436.

Cain, M. S., Biggs, A. T., Darling, E. F., & Mitroff, S. R. (2013). A little bit of history repeating: Splitting up multiple-targets visual searches decreases second-target miss errors. Submitted for publication.

Cain, M. S., Dunsmoor, J. E., LaBar, K. S., & Mitroff, S. R. (2011). Anticipatory anxiety hinders detection of a second target in dual target search. *Psychological Science*, *22*, 866–871.

Cain, M. S., & Mitroff, S. R. (2013). Memory for found targets interferes with subsequent performance in multiple-target visual search. *Journal of Experimental Psychology: Human Perception and Performance*, *39*(5), 1398–1408.

Chun, M. M., & Wolfe, J. M. (1996). Just say no: How are visual searches terminated when there is no target present? *Cognitive Psychology*, *30*, 39–78.

Clark, K., Cain, M. S., Adcock, R. A., & Mitroff, S. R. (2013). Context matters: The structure of task goals affects accuracy in multiple-target visual search. *Applied Ergonomics*. doi: 10.1016/j.apergo.2013.07.008

Clark, K., Samei, E., Baker, J., & Mitroff, S. R. (2011, November). Expertise in radiological screening and satisfaction of search. Poster presented at the annual Object Perception, Attention, and Memory meeting. Seattle, WA.

Cook, R. D., & Weisberg, S. (1982). *Residuals and influence in regression* (Vol. *5*). New York: Chapman and Hall.

Costello, M. C., Madden, D. J., Shepler, A. M., Mitroff, S. R., & Leber, A. B. (2010). Age-related preservation of top-down control over distraction in visual search. *Experimental Aging Research*, *36*, 249–272.

Dickinson, C. A., & Zelinsky, G. J. (2007). Memory for the search path: Evidence for a high-capacity representation of search history. *Vision Research*, *47*, 1745–1755.

Eckstein, M. P. (2011). Visual search: A retrospective. *Journal of Vision*, *11*, 1–36.

Ericsson, K. A., & Lehmann, A. C. (1996). Expert and exceptional performance: Evidence of maximal adaptation to task constraints. *Annual Review of Psychology*, *47*, 273–305.

Fleck, M. S., & Mitroff, S. R. (2007). Rare targets are rarely missed in correctable search. *Psychological Science*, *18*, 943–947.

Fleck, M. S., Samei, E., & Mitroff, S. R. (2010). Generalized "satisfaction of search": Adverse influences on dual target search accuracy. *Journal of Experimental Psychology: Applied*, *16*, 60–71.

Horowitz, T. S., & Wolfe, J. M. (1998). Visual search has no memory. *Nature*, *357*, 575–577.

Jackson, T. H., Clark, K., & Mitroff, S. R. (2013). Enhanced facial symmetry assessment in orthodontists. *Visual Cognition*. doi: 10.1080/13506285.2013.832450

Jackson, T. H., Mitroff, S. R., Clark, K., Proffit, W. R., Lee, J. Y., & Nguyen, T. T. (2013). Face symmetry assessment abilities: Clinical implications for diagnosing asymmetry. *American Journal of Orthodontics and Dentofacial Orthopedics*, *114*(5), 663–671.

Kleiner, M., Brainard, D. H., & Pelli, D. G. (2007). What's new in psychtoolbox-3? *Perception*, 36, ECVP Abstract Supplement.

Krupinski, E. A. (2010). Current perspectives in medical image perception. *Attention, Perception, & Psychophysics*, *72*, 1205–1217.

Kuhn, G. J. (2002). Diagnostic errors. *Academic Emergency Medicine*, *9*, 740–750.

McCarley, J. S., Kramer, A. F., Wickens, C. D., Vidoni, E. D., & Boot, W. R. (2004). Visual skills in airport security screening. *Psychological Science*, *15*, 302–306.

Menneer, T., Barrett, D. J. K., Phillips, L., Donnelly, N., & Cave, K. R. (2007). Costs in searching for two targets: Dividing search across target types could improve airport security screening. *Applied Cognitive Psychology*, *21*, 915–932.

Nakayama, K., & Martini, P. (2011). Situating visual search. *Vision Research*, *51*, 1526–1537.

O'Brien, R. M. (2007). A caution regarding rules of thumb for variance inflation factors. *Quality & Quantity*, *41*(5), 673–690.

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, *10*, 437–442.

Rabbitt, P. M. A. (1979). How old and young subjects monitor and control responses for accuracy and speed. *British Journal of Psychology*, *70*, 305–311.

Ratcliff, R., Thapar, A., & McKoon, G. (2007). Application of the diffusion model to two-choice tasks for adults 75-90 years old. *Psychology and Aging*, *22*, 56–66.

Schwaninger, A. (2003). Evaluation and selection of airport security screeners. *AIRPORT*, *2003*(2), 14–15.

Schwaninger, A. (2005). Increasing efficiency in airport security screening. *WIT Transactions on the Built Environment*, *82*, 405–416.

Schwaninger, A., Hardmeier, D., & Hofer, F. (2004). Measuring visual abilities and visual knowledge of aviation security screeners. IEEE ICCST Proceedings, 258–264.

Smith, M. J. (1967). *Error and variation in diagnostic radiology*. Springfield, IL: Thomas.

Tuddenham, W. J. (1962). Visual search, image organization, and reader error in roentgen diagnosis: Studies of psychophysiology of roentgen image perception. *Radiology*, *78*, 694–704.

Voytovich, A., Rippey, R., & Suffredini, A. (1985). Premature conclusions in diagnostic reasoning. *Journal of Medical Education*, *60*, 302–307.

Downloaded by [Duke University Libraries] at 12:20 12 December 2013