Predicting Airport Screening Officers' Visual Search Competency With a Rapid Assessment

Stephen R. Mitroff, The George Washington University, Washington, DC,
Justin M. Ericson, and **Benjamin Sharpe**, Kedlin Company, Bellevue, Washington

Objective: The study's objective was to assess a new personnel selection and assessment tool for aviation security screeners. A mobile app was modified to create a tool, and the question was whether it could predict professional screeners' on-job performance.

Background: A variety of professions (airport security, radiology, the military, etc.) rely on visual search performance—being able to detect targets. Given the importance of such professions, it is necessary to maximize performance, and one means to do so is to select individuals who excel at visual search. A critical question is whether it is possible to predict search competency within a professional search environment.

Method: Professional searchers from the USA Transportation Security Administration (TSA) completed a rapid assessment on a tablet-based X-ray simulator (XRAY Screener, derived from the mobile technology app Airport Scanner; Kedlin Company). The assessment contained 72 trials that were simulated X-ray images of bags. Participants searched for prohibited items and tapped on them with their finger.

Results: Performance on the assessment significantly related to on-job performance measures for the TSA officers such that those who were better XRAY Screener performers were both more accurate and faster at the actual airport checkpoint.

Conclusion: XRAY Screener successfully predicted on-job performance for professional aviation security officers. While questions remain about the underlying cognitive mechanisms, this quick assessment was found to significantly predict on-job success for a task that relies on visual search performance.

Application: It may be possible to quickly assess an individual's visual search competency, which could help organizations select new hires and assess their current workforce.

Keywords: visual search, performance prediction, aviation security, personnel selection, personnel assessment

INTRODUCTION

Aviation security organizations are faced with the daunting task of ensuring the safety of the traveling public and performing efficiently so that they can process enormous numbers of passengers (and their belongings) without causing travel delays. Few industries have two primary goals that directly contradict one another—but that is exactly the case for aviation security where performance is required to be both accurate and fast. This unenviable position requires creative solutions that can optimize operations so that the public can travel safely without burdensome wait times. One solution is to adopt a risk-based screening strategy where screening efforts are focused on areas of uncertainty and minimized when risks are low. The U.S. Transportation Security Administration's (TSA) precheck program is an example of such a strategy; passengers diminish their risk by providing information prior to travel, and thus TSA is able to save resources by subjecting these individuals to less intensive screening. For aviation security organizations to fully adopt such risk-based strategies, it is vital to explore all possible tools that can improve screener performance as well as screener selection, assessment, and training procedures. Moreover, it is important that such tools be cost-effective in their use and implementation for them to have a realistic chance of impacting operations.

The current project explored the viability of a new software-based program to assess screeners' effectiveness and efficiency at the checkpoint. The new program, XRAY Screener, is a rapid assessment tool that is administered via a handheld tablet. XRAY Screener was built from the preexisting iOS mobile gaming platform Airport Scanner (Kedlin Company). Airport Scanner is a publicly available mobile game where players serve as aviation security screeners and look for prohibited items in simulated X-ray bags. Airport Scanner has provided a

Address correspondence to Stephen R. Mitroff, The George Washington University, 2125 G St NW, Washington, DC 20052, USA; e-mail: mitroff@gwu.edu.

HUMAN FACTORS
Vol. XX, No. X, Month XXXX, pp. 1–11
DOI: 10.1177/0018720817743886

valuable tool for academic research into the nature of visual search; as of November 2017, Airport Scanner has provided researchers with over 3.3 billion trials of data from over 13 million users. This has led to a number of publications (Biggs, Adamo, Dowd, & Mitroff, 2015; Biggs, Adamo, & Mitroff, 2014; Ericson, Kravitz, & Mitroff, 2017; Mitroff & Biggs, 2014; Mitroff et al., 2015; Mitroff & Sharpe, 2017) and represents the largest visual search data set currently being utilized for academic research. Through a TSA contract to Kedlin Company, the Airport Scanner app was modified to create the XRAY Screener program employed in the current project.

The primary goal of this project was to determine if the XRAY Screener program could reliably predict airport security screeners' accuracy and speed at processing passenger carry-on bags at airport checkpoints. TSA officers completed an assessment component of XRAY Screener, and their performance was compared to their accuracy at finding prohibited items in real bags at the checkpoint and their speed of evaluating passengers' bags. In short, this project examined if this rapid, tablet-based program could meaningfully predict whether professional screeners are relatively good or bad at a particular aspect of their actual job.

The essence of the XRAY Screener program is a focus on "visual search"—the ability to detect targets among distractors. Visual search is a core cognitive ability that underlies a large number of daily activities (e.g., finding a friend in a crowd, locating an e-mail in your inbox, identifying a ripe apple at the grocery store). Many occupations rely on accurate search performance (e.g., radiology, lifeguarding, aviation security) such that this particular skill lies at the heart of whether or not individuals are able to properly perform their job. Visual search has been well studied in the cognitive psychology literature (for reviews, see Eckstein, 2011; Nakayama & Martini, 2011), and interestingly, early academic pursuits of human visual search were largely driven by real-world questions about how some individuals might perform better than others (see Clark, Cain, Adamo, & Mitroff, 2012). For example, some of the first research into the nature of search performance was focused on how radar operators located enemy ships (Koopman, 1956a, 1956b). Aviation security screeners complete many tasks (e.g., passenger pat-downs, ID checks, physical bag searches), but arguably their most important task is visually searching X-ray images at the security checkpoint to determine if a prohibited item is present. As such, XRAY Screener was built to primarily focus on visual search—with the intention of evaluating whether a particular individual has a relatively high or low visual search competency.

MATERIALS AND METHODS

This research project was conducted by Kedlin Company through a contract from the TSA. An earlier phase of the project was conducted at four U.S. airports, and those data served as a pilot for the data presented here (see Appendix). The current data were collected in October 2016 from TSA Officers at an airport in the U.S. Southeast. This research complied with the American Psychological Association Code of Ethics and was approved by the Institutional Review Board at The George Washington University. Informed consent was obtained from each participant. Officers were provided the opportunity to participate or to not participate, and this choice was made privately without their supervisors' knowledge. All data were confidential, and no results were shared with the TSA that linked individuals or target groups to their performance.

Participants

A total of 289 TSA officers created an account in the XRAY Screener program, and of those, data were successfully obtained from 278. The remaining 11 individuals either chose not to participate or withdrew prior to completing the assessment, as was their right as a participant in this research project. The 278 participants (female = 117) had an average age of 42.48 years (SD = 11.80 years) and were employed by the TSA for an average of 7.56 years (SD = 5.11). Participants had their data excluded from further analyses for four reasons. First, data were removed from 4 individuals who completed the assessment but not on their first attempt (e.g., they started, stopped, and then restarted again later). Second, an additional 6 participants completed the assessment, but the data were not correctly saved (they may have quit the program too early, or a connection issue occurred while data were being transferred). Third, 17 individuals who completed the assessment did not have TSA-relevant data available for analyses. Finally, data were removed from 3 participants who had an average tap time of less than 0.5 seconds. These 3 participants were removed from analyses as their performance suggested they were purposely not trying (their accuracy was also much worse than the population average).

XRAY Screener

XRAY Screener is a software-based tool that was implemented on Apple iPad tablets for this project. Users created a unique login, and during the creation of their account, they were asked questions related to their TSA employment. XRAY Screener presented users with simulated bag images (Figure 1), and users tapped with their finger on any prohibited items they detected. Trials were generated in real time and consisted of simulated drawings of possible prohibited and allowed items. None of the stimuli were real X-ray images (i.e., none of the images were sensitive or classified). Approximately 50% of the bags had a single threat item present, and the remaining bags had no prohibited items. While screeners presumably experience a lower prevalence rate at the checkpoint (i.e., half of all passengers' bags do not contain a threat), this approximate 50% prevalence rate was chosen for two reasons. First, the overarching goal of XRAY Screener is to measure general visual search competency—not necessarily visual search ability tied to specific stimuli or specific operational parameters. Second, XRAY Screener is meant to be an engaging and brief assessment tool—50% prevalence provides a nice balance of target present and target absent trials that can keep the user engaged and provide sufficient data for analyses. Each bag contained 5 to 15 allowed items. The program measured participants' accuracy, response time, and a variety of other metrics of performance.

Users first completed two tutorial levels that introduced the touch mechanics of the app and included a small set of possible prohibited items. The tutorials consisted of 12 and 24 trials, respectively, and included pop-up messages that provided tips and instructions for the users (e.g., informing the user that they could swipe bags off the screen to go faster). Both tutorials had to be successfully completed before a user could move to the assessment level. The assessment appeared immediately after the tutorial levels and consisted of 72 trials. Users were encouraged to be equally fast and accurate given that security personnel are tasked with being both efficient and effective. Users were given performance feedback throughout based on the Airport Scanner platform. There were 20 unique prohibited items (e.g., pistol, bullets, large water bottle, doctor's scalpel) that could appear in the assessment, and they appeared with various rates of frequency (ranging from 0.5% to 5%). The prohibited items also varied in relative salience such that some were easier to detect than others. There were 100 unique allowed items (e.g., pants, batteries, cell phone charger, brush) that could appear. Allowed items were randomly drawn, with replacement, for each trial.

The primary dependent variable for XRAY Screener performance was inverse efficiency—response time (target present tap time) divided by hit rate. Inverse efficiency is a useful variable as it combines speed and accuracy into a single measure (Townsend & Ashby, 1978, 1983). A large value represents worse performance, signifying the individual was slow and/or inaccurate, while a smaller value indicates greater accuracy and/or faster search speed.

Inverse efficiency was specifically chosen as the dependent variable of interest given that aviation security operations place both speed and accuracy demands on the screeners. Top performers in an aviation security setting should be both fast and accurate—having the ability to process more passengers while reducing the chances of a threat passing through the checkpoint. This simultaneous combination of speed and accuracy can be assessed via inverse efficiency. However, since it is also potentially interesting from a mechanistic viewpoint, response time and accuracy were assessed separately as well.

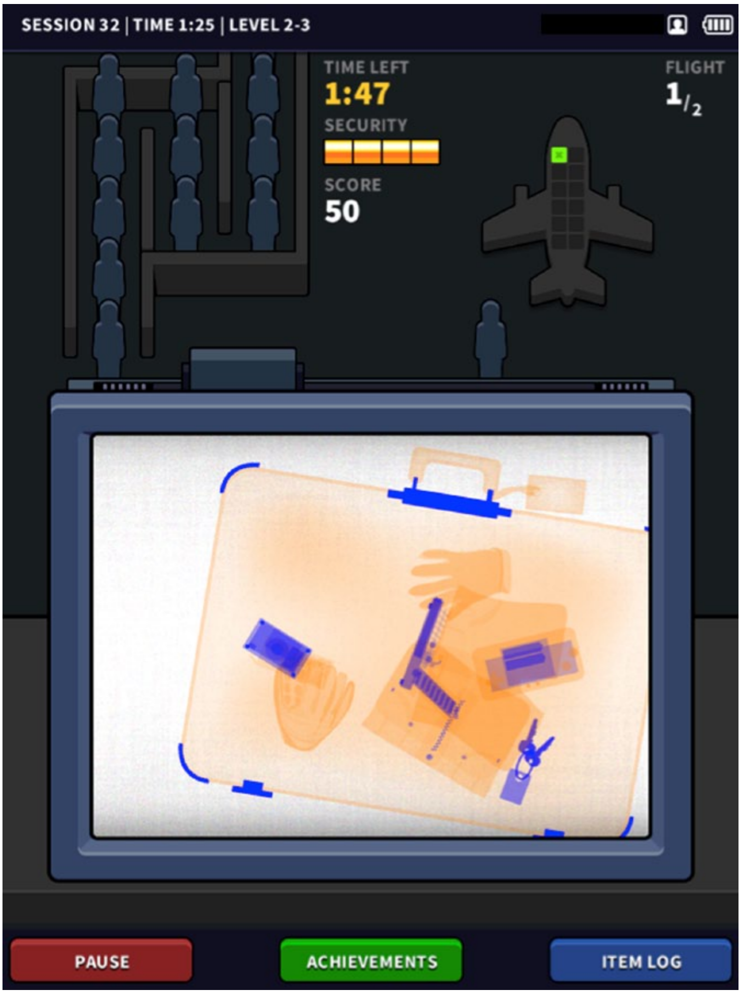

Figure 1. Sample display from XRAY Screener program.

Measures of TSA Officers' Performance

Three primary measures of TSA Officers' performance were evaluated for this project: covert testing miss rate, threat image projection (TIP) miss rate, and throughput rate.

Covert testing performance. Covert testing is the process of attempting to pass prohibited items through the security checkpoint to assess the workforce's effectiveness. The TSA provided the research team with covert testing data that were collected from October 2015 to October 2016 at the testing airport. These data are from threat awareness countermeasure training (TACT) activities conducted at the airport; the data represent local testing done by the airport leadership team where individuals unknown to the screeners attempt to pass prohibited items through the checkpoint. Tests were limited to "in property" tests (only tests that involved a prohibited item placed in a carry-on bag or parcel that went through the X-ray machine) and tests wherein the X-ray operator (the officer at the X-ray machine) was not aware that they were being assessed. Participants' data were only included in the analyses if they had at least five covert test events. While five events is arguably a low minimum threshold for inclusion, it highlights the difficulties of covert testing—it is time-consuming and difficult to properly administer these tests. It is not easy to scale covert tests to a large workforce in a way that would garner large numbers of trials for each individual. These issues emphasize the need for a more flexible and brief assessment tool. The primary measure of interest was miss rate—the percentage of covert tests not detected.

TIP performance. TIP is an aviation security tool that projects images of prohibited items into X-ray images of passengers' bags at the checkpoint (e.g., Hofer & Schwaninger, 2005; Schwaninger, 2006). Screeners are provided with immediate feedback for detected and missed TIP images, and monthly reports are generated for airport security leadership. For the current analyses, data were averaged across monthly TIP reports from October 2013 to October 2016. Participants had to have a minimum of 100 TIP exposures to be included in the analyses. The primary measure of interest was miss rate—the percentage of TIP images missed in bags in which a TIP was present.

Throughput (speed) performance. The primary dependent variable for throughput was the average time it took an officer to clear a bag (in seconds). Data were limited to bags from standard lanes (i.e., no precheck lanes) and collapsed over peak and off-peak times (i.e., busy and slow times). The throughput time represents the period between photodiode detection points within the X-ray machine—with one photodiode at the entrance and one at the exit of the X-ray chamber. Specific data points were removed that had values over 45 seconds. Only data from participants who contributed at least 1,000 trials were considered.

RESULTS

The primary finding from the current project was that XRAY Screener performance significantly and meaningfully predicted both effectiveness (e.g., covert testing performance) and efficiency (i.e., throughput rates at the checkpoint). Correlational data are presented first and then data separated into top and bottom performer groups. Note that the raw data are sensitive in nature so they are not presented here; relative performance metrics and statistical tests are presented, but raw data are intentionally excluded.

Correlation Analyses

The inverse efficiency scores from the XRAY Screener assessment significantly correlated with covert testing miss rate, TIP miss rate, and throughput (Table 1).

As noted, inverse efficiency was chosen as the primary dependent variable from XRAY Screener given a simultaneous focus on both speed and accuracy. However, it is also important to consider how speed and accuracy alone relate to the on-job performance measures. Speed was operationally defined as response time on target absent trials (how long it takes an individual to decide to terminate a search). Target-absent response time significantly related to covert testing miss rate, $r(181) = 0.27, p < .001$, and TIP miss rate, $r(227) = 0.14, p = .038$, with slower responders making more mistakes. Target-absent response time did not significantly relate to throughput, $r(147) = -0.09, p = .295$. Accuracy was assessed as *d*-prime, which incorporates target present hit rate and target absent false alarm rate. XRAY Screener *d*-prime was marginally related to covert testing miss rate, $r(181) = -0.13$, p = .088; highly related to TIP miss rate, $r(227) = -0.26, p < .001$; and not significantly related to throughput, $r(147) = 0.06$, $p = .443$. Of the XRAY Screener measures analyzed, inverse efficiency, which combines speed and accuracy, was the strongest predictor of both covert testing and throughput. This may suggest that successful aviation security XRAY interpretation requires a particular blend of speed and accuracy competency.

Quartile Analyses

Another way to examine if XRAY Screener performance related to on-job performance measures was to sort participants into those who performed well on the assessment and those who performed poorly. The prediction was that the high XRAY Screener performers would be significantly better than the poor performers on TIP, covert testing, and throughput. Participants were divided into four quartiles based on their performance in the XRAY Screener assessment in terms of inverse efficiency. Those who were in the bottom 25% (bottom quartile) in performance were those with the worst inverse efficiency scores, and those in the top 25% (top quartile) had the best scores. Inverse efficiency is target-present response time divided by hit rate so larger values represent worse performance. As seen in Figure 2, the top quartile

TABLE 1: Correlation Table for the Four Key Variables of Interest: XRAY Screener Assessment Inverse Efficiency, Covert Testing Miss Rate, TIP Miss Rate, and Throughput (Time/Bag)

		Covert Testing Miss Rate	TIP Miss Rate	Throughput (Time/Bag)
XRAY Screener inverse efficiency (higher values = worse performance)	Correlation (*r*)	0.32**	0.27**	0.28**
	Significance test (*t*)	4.48	4.29	4.35
	Significance (*p*)	<.001	<.001	<.001
	Count	183	229	225
Covert testing miss rate	Correlation (*r*)		0.31**	0.10
	Significance test (*t*)		5.89	1.73
	Significance (*p*)		<.001	.085
	Count		322	315
TIP miss rate	Correlation (*r*)			0.11*
	Significance test (*t*)			2.44
	Significance (*p*)			.015
	Count			469

Note. The significance values represent the *p* value for the correlation. *p* values lower than .05 are considered significant. The counts represent the number of participants contributing to each correlation. TIP = threat image projection.
$^{*}p < .05$. $^{**}p < .01$.

XRAY Screener performers were significantly better than the bottom quartile performers in terms of covert testing, $t(86) = 3.22$, $p = .002$; TIP, $t(111) = 4.19$, $p < .001$; and throughput, $t(109) = 3.37$, $p = .001$. Specifically, compared to the bottom quartile XRAY Screener performers, the top quartile performers were 32% more accurate at covert testing, 32% more accurate at TIP, and 12% quicker to process bags. Critically, these analyses demonstrate that the top XRAY Screener performers were both more effective and more efficient screeners in their daily tasks—they detect more covert threats and clear bags more quickly while on the job.

DISCUSSION

The primary goal of this project was to explore if a new rapid assessment tool that was designed to focus on visual search competency could reliably predict aviation security screeners' performance at the actual checkpoint. This was successfully accomplished as TSA officers who were top performers in XRAY Screener were both more accurate and faster performers when evaluating real bags at the checkpoint. Aviation security officers have many tasks, but arguably the X-ray operator role is the most important; the detection of prohibited items in the X-ray images is a critical step to ensure the safety of the traveling public. It is therefore vital to make sure that the individuals who are executing this task are the best suited to do so. The current project suggests that it might be possible to use a visual search assessment program to help inform the process of identifying superior performers (both in terms of accuracy and speed).

When comparing how well XRAY Screener and TIP predicted performance, it is noteworthy that XRAY Screener and TIP were roughly equal predictors of covert testing performance (the ability to detect real threat items at the checkpoint) and XRAY Screener was stronger than TIP as a predictor of throughput (speed). TIP was designed as a vigilance tool to allow screeners to experience threat items at a regular rate (e.g., Schwaninger, 2006). This was meant to provide systematic exposure given that (thankfully) the rate of true threats at the checkpoint is remarkably low. However, regardless of its original intent, various aviation security organizations have used TIP as a workforce assessment tool. TIP may have some validity as an assessment (however this is not always clear, see

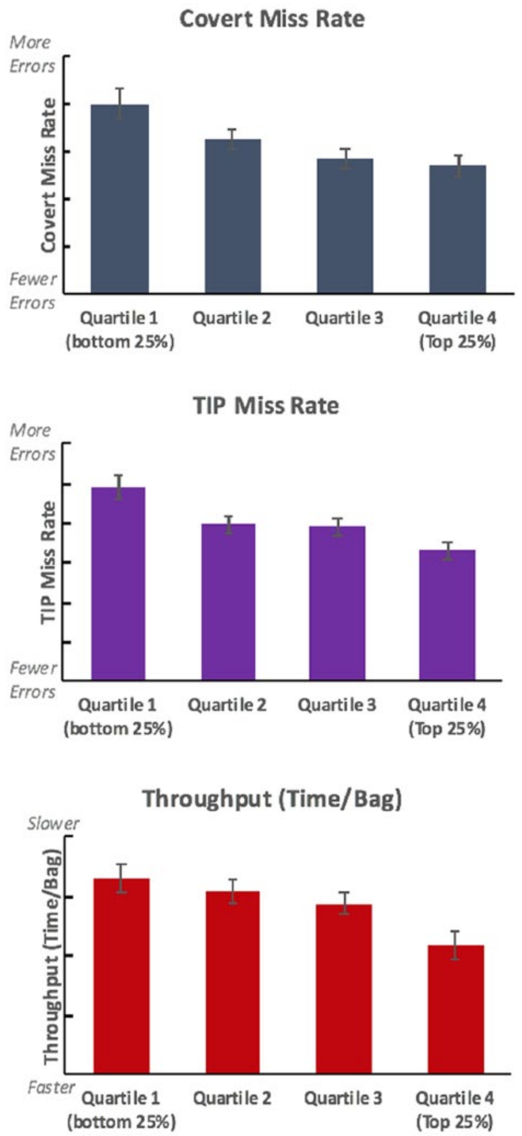

Figure 2. Quartile breakdowns of performance for covert testing miss rate, threat image projection (TIP) miss rate, and throughput (time/bag in seconds). Quartiles based on inverse efficiency performance in the XRAY Screener assessment. Axes values were removed from the graphs to protect sensitive data. Error bars represent standard error.

Appendix Table A1), yet it takes much longer to become potentially meaningful compared to XRAY Screener. TIP data are collected at the checkpoint, and it takes months post-training to gather sufficient data to assess how performance relates to covert testing accuracy. In contrast, the XRAY Screener data reported here were gathered from a single 15-minute assessment session. Moreover, this 15-minute assessment can be administered pretraining and pre-hire since the images are not sensitive or classified. The current findings suggest that a tool designed specifically as an assessment measure and focused on basic visual search competency may provide the industry with a needed resource.

An important aspect of XRAY Screener is that it does not use sensitive or classified images. There are several reasons for using simulated images. First, these images can be shown to anyone—even new applicants who have not yet been vetted for security purposes. Second, simulated images allow for greater control over the stimuli. For example, the 20 prohibited items were systematically varied in their relative frequency (their rate of appearing as a search target) and their relative salience (how easy they are to find). These factors were not further analyzed here as more data will be needed to tease apart such nuance, but this level of experimental control is best achieved with simulated bags that can be constructed in real time. Finally, this design allows for greater flexibility than using real bag images—it is possible to make simple adjustments to alter the assessment for a variety of future questions.

An open question is whether the use of simulated images might weaken the generalizable nature of XRAY Screener—can it really inform aviation security screening without using actual X-ray images? The simplest answer to this concern is that it did in fact predict actual X-ray screening performance. More generally, the logic of the current project was that the core cognitive ability of visual search is critical for an aviation security screener, and thus basic visual search competency should predict on-job success. That XRAY Screener uses images similar to those at the actual checkpoint is irrelevant—the critical point is that it taps into visual search competency in such a way that it can predict who is well suited to conduct aviation security screening.

Another open question is whether XRAY Screener assesses visual search performance per se or whether it is a proxy for a more generalized measure of engagement or ability. That is, are high performers in XRAY Screener specifically better at visual search, or are they individuals who are willing to try harder, are more conscientious, and/or have overall heightened cognitive abilities? It was recently shown that TSA officers who report higher levels of conscientiousness are better at visual search (Biggs, Clark, & Mitroff, 2017), so it is a viable interpretation that XRAY Screener could be a proxy for motivation and effort. However, another recent study demonstrated that early performance in Airport Scanner (the public version of XRAY Screener) is predictive of later success—those who start out as stronger searchers stay stronger searchers (Ericson et al., 2017). These data come from individuals who willingly downloaded and played the Airport Scanner game (and are presumably motivated to do well) and suggest that the app can distinguish individuals based on their visual search performance. Regardless, more work will be needed to definitively state whether XRAY Screener is a pure measure of visual search performance or whether it may also tap into more generalized engagement and/or ability.

PRACTICAL IMPLICATIONS

For the aviation security industry, the current project has three primary potential implications. First, XRAY Screener can potentially provide the ability to effectively assess current workforces. Organizations can identify which individuals are best suited for the X-ray operator role. Several organizations are considering "specialization"—where employees focus on a subset of tasks instead of rotating through all tasks. XRAY Screener could help with determining assignments for the specialization process. Second, XRAY Screener can potentially provide a tool for assessing new applicants to see if they have the necessary visual search skills to succeed in this important position. Given that the assessment does not involve sensitive or classified images, it can be used at any point in the hiring process to vet individuals for the particular skill of visual search. This can potentially save an organization both time and money as they can assess an individual at the beginning stages of the hiring process before they have invested in hiring and training the individual. Finally, by hiring individuals who are better suited for the job and better understanding the strengths and weaknesses of the current workforce, an organization can reduce its risk.

This project highlights that there are individual differences in visual search that can be effectively measured. Prior work has clearly shown that performance can improve with experience (e.g., Stafford & Dewar, 2014), but it is important to consider that some individuals might start out better suited than others. For example, a prior study using data from the Airport Scanner game data found that individuals who start as stronger visual search performers are likely to remain the stronger performers (Ericson et al., 2017). As such, optimal performance can potentially be achieved by hiring the right individuals from the start (those who are predisposed to excel at search) and then training those individuals to make them as strong as they possibly can become.

CONCLUSIONS

The current project had a specific and narrow focus—demonstrating that a rapid assessment tool could successfully predict professional screeners' on-job performance. The success of this project opens many potential future directions. For example, a natural extension of this project is to explore ways in which XRAY Screener can be used as a training tool. The current implications are for personnel selection and assessment, but the program framework can be used to also administer prescribed experiences that could strengthen visual search performance. Likewise, the program can be used to inform additional industries beyond aviation security. In sum, visual search is a critical skill for an aviation security screener, and the XRAY Screener program leverages this fact to identify those individuals who may be better suited than others for this highly critical occupation.

APPENDIX

PILOT DATA

The data presented in the current report were from a research project conducted at an airport in the U.S. Southeast. Prior to this study, a pilot version of the project was run at four other U.S. airports. This work was funded through the same source (TSA Contract No. HSTS04-15-C-CT7031). This Appendix reports the relevant data, which demonstrate the same results as the primary study.

Methods

TSA officers at four U.S. airports completed the XRAY Screener assessment. Covert testing and threat image projection (TIP) data were provided for a subset of the participants, and XRAY Screener performance was compared to these on-job performance metrics. Throughput (speed) data were not available. The primary dependent variable for XRAY Screener was *d*-prime, which provides a measure of the likelihood of finding a prohibited item on target-present trials and for correctly withholding a response for bags that did not contain a prohibited item.

Results

XRAY Screener performance significantly related to both covert testing and TIP performance, such that those who performed well in XRAY Screener were more likely to perform well in these two measures of on-job performance (Table A1).

TABLE A1: Correlation Table for the Three Key Variables of Interest: XRAY Screener Assessment *d*-Prime, Covert Testing Miss Rate, and TIP Miss Rate

		Covert Testing Miss Rate	TIP Miss Rate
XRAY Screener *d*-prime (higher values = better performance)	Correlation (*r*)	–0.38**	–0.23**
	Significance test (*t*)	2.74	3.68
	Significance (*p*)	.009	<.001
	Count	48	241
Covert testing miss rate	Correlation (*r*)		0.04
	Significance test (*t*)		.029
	Significance (*p*)		.774
	Count		61

Note. The significance values represent the *p* value for the correlation. *p* values lower than .05 are considered significant. The counts represent the number of participants contributing to each correlation. TIP = threat image projection.
*$p < .05$. **$p < .01$.

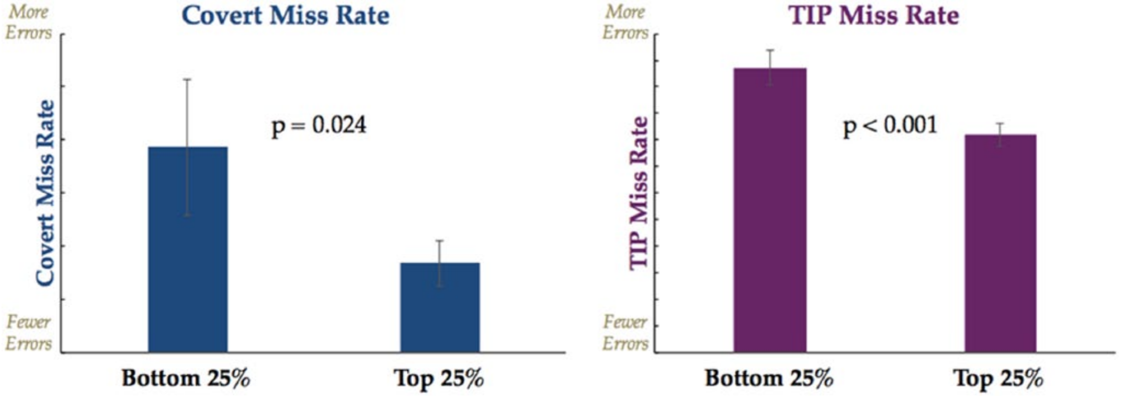

Figure A1. Top versus bottom quartile breakdowns of performance for covert testing miss rate and threat image projection (TIP) miss rate. Quartiles based on *d*-prime performance in the XRAY Screener assessment. Axes values were removed from the graphs to protect sensitive data. Error bars represent standard error.

Participants were divided into quartiles based on their XRAY Screener performance, and the top performers were significantly more accurate at both covert testing and TIP (Figure A1).

ACKNOWLEDGMENTS

The research team would like to thank Darryl Smith, Bonnie Kudrick, Roxanne Doyle, and the TSA Office of Requirements and Capabilities for help with conducting this research project and Thomas Liljetoft and Alex Hartley for their roles in creating the XRAY Screener program. We would especially like to thank the TSA officers and local airport management teams that helped with the project and participated in the research. We also thank the GW Visual Cognition lab, Dwight Kravitz, and others in the GW vision community for helpful feedback. This project was funded by TSA Contract No. HSTS04-15-C-CT7031, and this manuscript has been approved for public release. Two of the authors (Mitroff, Sharpe) are co-founders of a company that sells the XRAY Screener product. The research reported here was done under a research and development project funded by the TSA, and the outcome led to the creation of this new company.

KEY POINTS

- Aviation security screeners must search passengers' bags both accurately and quickly.
- Employing the best visual searchers can help optimize security operations.
- The current study examined a rapid tool to identify top and bottom visual searchers.
- The rapid tablet-based assessment reliably predicted screeners' on-job performance.
- This tool can potentially be used to assess new applicants and current workforces.

REFERENCES

Biggs, A. T., Adamo, S. H., Dowd, E. W., & Mitroff, S. R. (2015). Examining perceptual and conceptual set biases in multiple-target visual search. *Attention, Perception, & Psychophysics*, *77*, 844–855.

Biggs, A. T., Adamo, S. H., & Mitroff, S. R. (2014). Rare, but obviously there: Effects of target frequency and salience on visual search accuracy. *Acta Psychologica*, *152*, 158–165.

Biggs, A. T., Clark, K., & Mitroff, S. R. (2017). Who should be searching? Differences in personality can affect visual search accuracy. *Personality and Individual Differences*, *116*, 353–358.

Clark, K., Cain, M. S., Adamo, S. H., & Mitroff, S. R. (2012). Examining influences on applied visual search performance. In M. D. Dodd & J. H. Flowers (Eds.), *The influence of attention, learning, and motivation on visual search* (Vol. 59, pp. 147–181). New York, NY: Springer.

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

Ericson, J. M., Kravitz, D. J., & Mitroff, S. R. (2017). Visual search: You are who you are (+ a learning curve). *Perception*. Advance online publication. doi:10.1177/0301006617721091

Hofer, F., & Schwaninger, A. (2005). Using threat image projection data for assessing individual screener performance. *WIT Transactions on the Built Environment*, *82*, 417–426.

Koopman, B. O. (1956a). The theory of search: I. Kinematic bases. *Operations Research*, *4*(3), 324–346.

Koopman, B. O. (1956b). The theory of search: II. Target detection. *Operations Research*, *4*(5), 503–531.

Mitroff, S. R., & Biggs, A. T. (2014). The ultra-rare-item effect: Visual search for exceedingly rare items is highly susceptible to error. *Psychological Science*, *25*, 284–289.

Mitroff, S. R., Biggs, A. T., Adamo, S. H., Dowd, E. W., Winkle, J., & Clark, K. (2015). What can 1 billion trials tell us about visual search? *Journal of Experimental Psychology: Human Perception & Performance*, *41*(1), 1–5.

Mitroff, S. R., & Sharpe, B. (2017). Using big data to solve real problems through academic and industry partnerships. *Current Opinion in Behavioral Sciences*, *18C*, 91–96.

Nakayama, K., & Martini, P. (2011). Situating visual search. *Vision Research*, *51*(13), 1526–1537.

Schwaninger, A. (2006, December). Threat image projection: Enhancing performance? *Aviation Security International*, 36–41.

Stafford, T., & Dewar, M. (2014). Tracing the trajectory of skill learning with a very large sample of online game players. *Psychological Science*, *25*, 511–518.

Townsend, J. T., & Ashby, F. G. (1978). Methods of modeling capacity in simple processing systems. In J. Castellan & F. Restle (Eds.), *Cognitive theory* (Vol. 3, pp. 200–239). Hillsdale, NJ: Erlbaum.

Townsend, J. T., & Ashby, F. G. (1983). *Stochastic modeling of elementary psychological processes*. Cambridge, UK: Cambridge University Press.

Stephen R. Mitroff earned his PhD in cognitive psychology from Harvard University in 2002. He is currently an associate professor in the Department of Psychology at The George Washington University and is the Chief Science Officer for Kedlin Screening International.

Justin M. Ericson earned his PhD in cognitive psychology from Louisiana State University in 2014. He is currently a User Research Specialist at Microsoft.

Benjamin Sharpe earned his BS in electrical engineering from the University of Washington in 2003. He is the Founder of Kedlin Company and the CEO of Kedlin Screening International.

Date received: July 14, 2017
Date accepted: October 29, 2017