Perceptions and Affective Responses to Alternative Risk-Based Airport Security

Kenneth D. Nguyen[(✉)] and Richard S. John

Department of Psychology, University of Southern California,
Los Angeles, CA, USA
{hoangdun, richardj}@usc.edu

Abstract. The current research examines how American air travelers perceive various risk-based airport security screening policies that vary in terms of selection procedure and agency. Respondents were randomly assigned to one of six experimental conditions that differ in terms of procedures for selecting passengers for enhanced screening. Respondents were presented with a conventional security option and a risk-based option, and were asked to rate these policies in terms equity, safety, and convenience. They also indicated anticipated feelings if selected for an enhanced screening in the risk-based procedure. Results suggest that the conventional approach was perceived as safer and more equitable but less convenient. Importantly, while different passenger selection procedures for enhanced screening led to distinct perceptions and feelings, respondents were indifferent between an equivalent selection procedures conducted by humans versus by computer.

Keywords: Airport security policy · Passenger screening · Risk based security · Randomized security · Behavioral detection · Profiling · Human–computer agency

1 Introduction

Conventional wisdom suggests that security should be omnipresent. That is, an effective airport security measure should be pervasive and exhaustive, such that most, if not all travelers and their luggage are screened and checked thoroughly. Yet, security resources are limited. The U.S airports service more than 895.5 million of passengers in 2015 [1], and it is challenging to screen each of these travelers thoroughly. *Risk-based security* emerges as a promising alternative [2]. In this approach, scare security resources are efficiently assigned to areas in which the level of security threat is high and to passengers who are associated with a high threat level. Selective security screening is an example of this approach. For instance, the current TSA Pre identifies and requires a subgroup of passengers to undergo a more vigorous screening at security checkpoints.

Although risk-based security promises to enhance security efficiency, it presumes that individual travelers understand and support this approach. Yet, public perceptions are not always aligned with policy makers' visions. For example, one of the concerns with risk-based policy is that the public may view it as less safe than a conventional security policy. Indeed, a risk-based screening that probabilistically selects air travelers

for security checks is perceived as being riskier than a traditional security that searches everyone [3]. Such misperceptions can lead to unfortunate consequences. For example, travelers may opt for riskier transportation modes when they do not feel that the risk-based approach adequately protects them from security threats [4, 5]. Given the importance of understanding how the general public perceives and reacts to risk-based security, we investigate how a sample of Americans perceive various versions of risk-based security screening procedures. In particular, we explore how American air travelers perceive different risk-based screening policies that vary in the method of selection of passengers for enhanced screening and differ in terms of the intelligent agent, i.e. human versus computer, that decides who will be selected.

2   The Effects of Selection Procedures on Perceptions of Equity and Safety

Risk-based security is not without controversy. One key concern is that the TSA has relied on factors whose validity has not yet been established to classify high-risk air travelers [6]. Three examples of risk-based screening that have generated a great deal of public discussion are selection based on behavioral profiling, individual characteristics, and randomization. Although the explicit use of profiling is prohibited by law, demographic variables such as age and sex have been used in some screening policies. Indeed, the current policy allows passengers 75 years of age or older to keep their (light) jackets, belts, and shoes on when undergoing a security screening. Behavioral screening is another selection procedure that could be applied [6]. The rationale behind this technique is that potential terrorists who disguise themselves as regular air travelers may exhibit certain emotional responses such as anxiety, fear, and distress. Through extensive training, behavioral detection officers are believed to be able to detect micro changes in the facial expression of these potential terrorists, and apprehend these suspects before they can execute an attack. A third selection procedure is randomized security screening. A limited version of randomized screening has been applied for passengers who currently participate in the Trusted Travel Program. Recall that passengers who sign up for this program undergo an expedited screening. However, TSA officers still randomly select some of these travelers for additional screening.

Although the three screening policies result in the same inequitable outcome, i.e. only a selective group of travelers will undergo an enhanced screening, the three selection strategies are distinct in the procedures employed to select travelers for an enhanced screening. One interesting question is whether travelers perceive these selection strategies as equally effective in identifying high-risk passengers. A second question is whether travelers perceive these selection strategies as equally fair to all passengers. The latter question was indirectly addressed in a recent study in which air travelers were asked to make trade-offs for equity when the levels of equity were manipulated [7]. Respondents in this study were presented with two airline options: The first employs a conventional search-all policy whereas the second adopts a risk-based screening policy. The latter one is less fair than the former one in the sense that only some passengers will incur the cost of additional security searchers but it is also less costly, more convenient, and safer than the conventional security. Importantly,

the selection strategy used in the risk-based screening was manipulated. Results indicated that respondents were willing to pay more, wait longer in a security line, and give up some degree of safety and convenience to avoid a risk-based screening that selects travelers based on their individual characteristics such as sex, age, and nationalities compared to an enhanced security risk-based screening procedure that randomly selects travelers.

Even though these results imply that respondents perceived the selective screening based on individual characteristics as unfair, this implication is limited because perception of equity was not measured. The current research remedies this limitation by exploring how American air travelers perceive the level of equity in alternative risk-based screening policies that vary in their selection strategies. A possible result is that air travelers perceive selection for enhanced screening based on individual characteristics as worse (in terms of equity) than randomized screening, a finding consistent with the previously discussed study. Yet, people may still perceive biased patterns in a perfectly random process [8, 9], which suggests that individuals may believe and perceive a random selection process as equally or even more biased than a selection process based on individual characteristics. Likewise, it is unclear how air travelers perceive the level of risk or safety across different risk-based screening options. A possible result is that travelers may view the individual profiling as more effective in identifying terrorists than a randomized screening. This hypothesis, indeed, is the exactly the argument that has been used to justify a temporary travel ban against travelers from specific Muslim-majority countries [10]. Yet, an alternative and equally plausible pattern is that randomization is perceived as a superior strategy. One rationale for this hypothesis is that the randomized search process will deter adversaries from attacking the system because additional uncertainties make it difficult for attackers to observe weakness in the system and identify a preferred attack paths [11].

Given the conflicting nature of these hypotheses and the limited research on the topic, rather than testing a priori hypothesis, the first experiment aimed to explore how people judge the levels of equity and safety in different risk-based screening procedures that differ in their selection strategy for enhanced screening. The first experiment also explores how a sample of American travelers evaluates the convenience level for different selection procedures for enhanced screening.

3   The Effect of Security Automation on Perceptions of Safety and Equity

Because intentionality plays a key role in forming fairness attribution [12, 13], we expect that perception of equity is also a function of perceived intention. Because intentionality is judged differently depending on the type of intelligent agents conducting the action, it is expected that perception of equity also depends on the type of intelligent agent conducting the security checks. Although traditional security screening is implemented by humans, recent prospective security screenings are built around the concept man-machine system. Indeed, the TSA houses its own human factor division to study, examine, and develop integrated security prototypes that facilitates collaboration between human security personnel and computerized systems [14].

Table 1. Descriptions of the experimental contexts

Agent	Screening procedures	Coastal airline	Pacific Airline
Computer	Individual profiling	…some passengers will be selected for a more thorough screening based on their characteristics…The selection process is implemented **by a computer**…	…all passengers and their carry-on baggage are screened thoroughly.
	Behavioral selection	…some passengers will be selected for a more thorough screening based on behaviors that indicate malicious intent…… The behavioral selection is implemented **by a computer**…	…all passengers and their carry-on baggage are screened thoroughly.
	Randomization	…some passengers will be selected for a more thorough screening based on randomization…The randomized selection process is implemented **by a computer**…	…all passengers and their carry-on baggage are screened thoroughly.
Human	Individual profiling	… (Same as above)…The selection process is implemented by **a security officer**…	…all passengers and their carry-on baggage are screened thoroughly.
	Behavioral selection	… (Same as above) … The behavioral selection is implemented by **a security officer**…	…all passengers and their carry-on baggage are screened thoroughly.
	Randomization	… (Same as above)…The randomized selection process is implemented by **a security officer**…	…all passengers and their carry-on baggage are screened thoroughly.

Given the advances in autonomous security, the critical question is how air travelers perceive autonomous or semi-autonomous security screening systems. In particular, how do perceptions of safety and equity vary when the automation feature is combined with various selection strategies? Thus, the second aim of this research is to further explore the role of computer automation in airport security screening. The distinction between humans and computers is an important one. While humans are more prone to error and bias and tend to apply predefined rules in a haphazard way, computers have the superior capabilities to apply rules in a consistent manner. This means that a computerized procedure may be perceived as being more consistent than human [15]. Because consistent rule application forms the basis of fairness, a computerized security procedure may be perceived as being more equitable and less biased than one based one based on human judgment.

The computer versus human agent effect on traveler perceptions of safety and risk is less clear, however. On the one hand, an automated computerized system would be expected to have superior performance over humans because of its ability to apply the same set of presumably valid screening criteria consistently over time, without being prone to bias and without subject to fatigue. Thus, air travelers may perceive a computerized security system as being more effective and safer. On the other hand, an equally plausible hypothesis is that travelers view a security procedure implemented by security personnel as being safer than an autonomous computerized security system. This could be due to the lay belief that humans are uniquely capable of detecting micro facial and behavioral expressions, e.g. fear, anxiety, and stress, which a terrorist may express when carrying out his mission. Because these human behaviors are subtle, a computerized security program may easily overlook relevant cues and allow an attacker to pass through the security checkpoint. Furthermore, while anautonomous system is pre-programmed, human detectors may be perceived as more flexible and able to detect new and emerging risk factors that have not been programmed in the computer system. This is particularly important because security threats are constantly evolving over time. Terrorist adversaries can observe a screening infrastructures put into place by a defender, and adaptively modify an existing attack plan, or devices new attack plan [11]. Despite these conflicting hypotheses, there has been little research to address this issue.

In summary, this study was designed to explore the effects of different risk-based screening procedures and the influence of security automation on perceptions of equity and safety (and convenience). In addition, the first study also explored the role of affective response. This is because affective response, in addition to subjective perception of safety and fairness, are potentially powerful determinants of behavior. Furthermore, learning how travelers react to different alternative risk-based security screening strategies provides valuable information regarding public preferences and support for alternative selection procedures.

4 Method

4.1 Procedure

Six-hundred respondents were recruited from Amazon Mechanical Turk, and they were each paid 0.25 cents for their participation. This study is a 2 Intelligent Agent (human vs. computer)-by-3 Screening Procedure (behavior vs. profile vs. randomization) between-subject factorial design. Respondents were randomly assigned to one of the six experimental conditions. Respondents were told that they were taking part in a study to help researchers understand how travelers perceive different screening procedures at airports. They were also told to imagine that they were searching for an airplane ticket to a vacation spot within the United States, and they have narrowed down their search to two options: Pacific and Coastal Airlines. The two airlines are identical in most respects except for their security screening policies. Pacific Airline employs a conventional screening process under which all passengers are screened thoroughly whereas Coastal Airline utilizes a risk-based screening process under which

all passengers simply walk through a screening device but some passengers will be selected for additional screening. Respondents were asked to contrast the screening procedures of the two airlines in terms of safety, (security) risk, bias, consistency, convenience, and equity. Note that the comparison of two options is similar to actual decisions in which people compare and contrast multiple travel options. The final sample included 527 respondents[1]; 47% of were male, the median age was 33 years, and 79% self identified as Caucasian. Sixty-seven percent of the respondents had flown at least once in the past 12 months.

4.2   Experimental Manipulation

The key manipulation is the description of the Coastal Airline's screening policies. Respondents in each condition received a different version of the screening policy, in which the type of intelligent agents and the type of screening procedures were varied. Table 1 is the summary of the exact wording for each version. Respondents in all conditions were asked to compare the two airlines on the following attributes: equity, safety, risk, convenience, consistency, bias, and risk. Respondents rated the relative standing of the two airlines on each of these six attributes on a bipolar scale, where seven means the attribute is extremely high in the conventional screening policy (Pacific Airline) and one means the attribute is extremely high in the risk-based screening policy (Coastal Airline). The order of the attributes was randomized to avoid order effects. Once ratings were completed, respondents were asked to choose which airline they prefer to fly.

4.3   Emotion Elicitation

To elicit emotional responses, respondents were told to imagine how they would react when they are singled out for additional, enhanced security screening when flying on the airline that has a risk-based screening policy. Respondents were asked to provide three to five words describing their feelings, and were asked to complete a brief version of the Positive Affect and Negative Affect Scale [16].

4.4   Other Dependent Measures

In addition to the attribute and PANAS ratings, respondents completed several demographic questions, including sex, age, and ethnicity. They were also asked about their awareness of several actual security programs initiated by the TSA, and their perceptions of the current screening policy in practice at U.S. airports.

---

[1] Seventy-three responses were removed because respondents failed an attention check question or did not indicate their Mechanical Turk IDs—an indication that the respondents did not take the survey seriously because without the IDs, we could not assign their credits or assign penalties for careless responses.

5 Results

5.1 Analytical Approach

Attribute ratings were recoded into a bipolar scale from −3 to 3, symmetric around 0. Negative ratings indicate that the risk-based screening is stronger in the attribute under study than the conventional screening whereas positive ratings indicate otherwise. A value of zero indicates that the two screening policies achieve the same level of the attribute. Examinations of reliability coefficients, Chronbach's Alpha, suggest that ratings of several attributes could be combined. Specifically, ratings for safety and risk (reverse coded) attributes were combined to create a safety index ($\alpha = 0.87$). Ratings for equity, bias (reverse coded), and consistency were combined to form an equity index ($\alpha = 0.71$). Convenience ratings remained as a separate attribute scale.

In addition, respondents' qualitative responses during the emotion elicitation were also coded. Two specific emotions, anger and shame were the two most commonly expressed, and were coded according to whether they were present or absent for each respondent. A code of 1 was given when a respondent expressed at least an anger-laden word (e.g. frustrated, irritated, and annoyed) and 0 otherwise. Similarly, a code of 1 was given when a respondent expressed at least a shame-laden word (e.g. ashamed, judged, and embarrassed).

A series of ANOVAs were conducted to examine the effects of the 3×2 experimental manipulation of the risk-based screening procedures on the attribute scores and affective responses. Binary logistic regressions were conducted to predict the expression of specific emotions, as well as to predict their choices between the risk-based vs. the conventional policy.

Table 2 presents the Pearson correlations among the dependent variables as well as summary statistics (on the diagonal). It is particularly noteworthy that equity is significantly correlated with safety, Spearman's $r = 0.49$, $p < 0.001$, and safety is significantly correlated with convenience, r = −0.22, $p < 0.001$.

5.2 The Experimental Effects on Security Attributes and Affective Responses

Series of ANOVAs were conducted with the experimental groups as the independent variables and equity, safety, and convenience, negative and positive affect, and behavioral intention as dependent variables. The tests revealed that there were main effects of Screening Procedures on perception of equity, negative affect, and positive affect, $F(2, 521) = 8.329$, $p < 0.001$; $F(2, 521) = 11.256$, $p < 0.001$; $F(2, 521) = 3.855$, $p = 0.022$, respectively. Surprisingly, there was no significant effect of agency (computer vs. human) on screening attributes and affective responses.

Follow-up t-tests with Bonferroni's correction for family-wise error rate (FWER) were conducted to explore the nature of the effects of screening procedures on equity, negative affect, and positive affect. Figure 1 illustrates these significant effects. Results indicate that the risk-based screening that applies to random selection was perceived as being more equitable than the risk-based screening that selects travelers based on their

individual characteristics, $t(348) = 3.978$, $p < 0.001$. Similarly, the randomized screening was perceived as being fairer than the risk-based behavioral screening, t $(348) = 2.711$, $p = 0.007$.

Results also indicated that being randomly selected for an enhanced screening led respondents to feel less negative than being selected based on individual characteristic, $t(333) = 3.404$, $p < 0.001$, or being selected based on behavioral and emotional expressions, $t(333) = 3.480$, $p < 0.001$. Unexpectedly, respondents felt more positive when being selected based on their individual characteristics than when being selected for enhanced screening based on their behavioral and emotional expressions, t $(347) = -2.826$, $p = 0.004$.

Table 2. Correlations among dependent variables

	Equity	Safety	Convenience	Positive	Negative	Anger	Shame	Decision
Equity	1.545 [1.249]							
Safety	0.47*	0.637 [0.760]						
Convenience	-0.09**	-0.17*	-1.347 [1.823]					
Positive affect	-0.06	0.01	0.08	2.169 [0.981]				
Negative affect	0.12*	0.06	0.16*	-0.28*	2.118 [1.159]			
Anger (1 = Presence)	0.14*	-0.02	-0.04	-0.13*	0.12*	67% [0.471]		
Shame (1 = presence)	0.11*	0.11*	0	-0.04	0.1**	0.06	0.12 [0.327]	
Decision (1 = convention)							63% [0.48]	

Note: * p < .05; means and standard deviations are provided in the diagonal

5.3   Predicting Shame, Anger, and Decision

Binary Logistic Regression (BLR) was used to predict the emotions of shame (0 = absence or 1 = presence) and anger (0 = absence or 1 = presence) and the choice between airlines using risk-based screening (coded 1) versus conventional screening (coded 0). The followings were included as predictors in BLR models: the experimental group (random selection is the reference group), race (White vs. non-White), sex (males coded 1), and age. In the anger BLR model, only age significantly predicted the presence of anger, OR (odds ratio) = 0.97, 95% CI [0.96–0.99], $p < 0.001$. A one-year increase in age was significantly associated with 0.973 decreases in the odds of reporting feelings of anger. In other words, older respondents were less likely to experience anger than were younger respondents when being selected for an additional security search. For the BLR model predicting shame responses, the contrast between randomization versus profiling was significant, OR = 2.88, 95% CI [1.18–7.80], $p < .01$. Hence, the odds for

respondents in the individual profiling group reporting feelings of shame were 2.88 times higher than the odds for respondents in the randomization group reporting the same emotion. Note that the 95% confidence was wide, which was partially due to the low number of respondents experiencing shame (12%).

In the airline choioce BLR model, both sex and the contrast between Whites vs. non-Whites significantly predicted respondents' choice of the screening policy. Compared to male respondents, the odds for female respondents choosing the risk-based screening were 0.49 times smaller with a 95% CI [0.34–0.71], $p < 0.001$. In addition, the odds for non-White respondents to choose the risk-based screening were 0.49 times smaller than the odds for White respondents to choose the same screening with a 95% CI [0.30–0.79], $p = 0.004$.

Fig. 1. Subjective ratings of attributes and affect. A higher mean rating indicates a higher level of an attribute. Negative ratings indicate the risk- based screening is stronger in an attribute compared to the conventional screening and positive screening indicates otherwise

6   Conclusion

This study investigated how people perceive various versions of risk-based airport security screening. Interestingly, while respondents perceive the three distinctive selection procedures in risk-based screening as being equally effective in enhancing airport security, they viewed the methods of selecting passengers for additional security checks based on individual profiling and behavioral screening as being less fair than a randomized selection procedure. Respondents also indicated that they would react more negatively and would be more likely to feel ashamed if they were selected for an additional search based on their personal characteristics such as sex, age, and nationality compared to a randomized selection. Most respondents indicated a preference for the conventional screening procedure over a risk-based procedure.

One of the surprising results in the first experiment is the absence of an agency effect. Although our results suggest that travelers perceive no safety or equity differences between computerized and human-based screening selection, this result may be due to the fact that power may have been limited. In particular, the experimental stimuli were somewhat subtle: The only difference between computer and human group manipulation was a single word, i.e. "computer" versus "human". Had the experimental manipulation of agency been more potent, we would have had a better chance of detecting an agency effect on perceptions of equity and safety. We are currently experiments with more potent manipulations of agency.

Public support and cooperation for passenger screening is essential for keeping air travel safe. Although the TSA has proposed and piloted a number of security initiatives, little research has been conducted to examine public support for these policies and procedures. The present research bridges this gap by showing that the traveling public does not hold a positive view of risk-based security compared to the conventional security approach. However, there may be an opportunity to enhance perceived fairness of airport security fairer by expanding on the use of randomized selection procedures. Future research that explores how various aspects of airport security measures contribute to travelers' affect, risk perceptions, and consumer choice behaviors will contribute to making air traveling safer, fairer, and more cost-effective.

References

1. Bureau of Transportation Statistics https://www.rita.dot.gov/bts/press_releases/bts018_16
2. Morral, A.R., Price, C.C., Ortiz, D.S., Wilson, B., LaTourrette, T., Mobley, B.W., Willis, H.H.: Modeling terrorism risk to the air transportation system: An independent assessment of TSA's risk management analysis tool and associated methods. Technical report, RAND Corporation (2012)
3. Scurich, N., John, R.S.: Perceptions of randomized security schedules. Risk Anal. **34**(4), 765–770 (2014)
4. Gigerenzer, G.: Dread risk, September 11, and fatal traffic accidents. Psychol. Sci. **15**(4), 286–287 (2004)
5. Gigerenzer, G.: Out of the frying pan into the fire: behavioral reactions to terrorist attacks. Risk Anal. **26**(2), 347–35 (2006)
6. U.S. Government Accountability Office. Aviation security: TSA should limit future funding for behavior detection activities: Report to congressional requesters. Technical report, United States Government Accountability Office (2013)
7. Nguyen, D.K., Rosoff, H., John, R.S.: Valuing Equal Protection in Aviation Security Screening. Risk Anal. (2017). doi:10.1111/risa.12814
8. Lopes, L.: Doing the impossible: a note on induction and the experience of randomness. J. Exp. Psychol. Learn. Mem. Cogn. **8**(6), 626–636 (1982)
9. Lopes, L.L., Oden, G.C.: Distinguishing between random and nonrandom events. J. Exp. Psychol. Learn. Mem. Cogn. **13**(3), 392–400 (1987)
10. National Public Radio. http://www.npr.org/sections/thetwo-way/2017/02/09/514384549/appeals-court-rejects-bid-to-reinstate-trumps-travel-ban

11. Jain, M., Tsai, J., Pita, J., Kiekintveld, C., Rathi, S., Tambe, M., Ordóñez, F.: Software assistants for randomized patrol planning for the LAX airport police and the federal air marshal service. Interfaces **40**(4), 267–290 (2010)
12. Falk, A., Fehr, E., Fischbacher, U.: Testing theories of fairness—Intentions matter. Games Econ. Behav. **62**(1), 287–303 (2008)
13. Nelson, W.R.: Equity or intention: It is the thought that counts. J. Econ. Behav. Organ. **48**(4), 423–430 (2002)
14. Greene, F., Kudrick, B., Muse, K.: Human factors engineering at the Transportation Security Administration. In Proceedings of the Human Factors and Ergonomics Society Annual Meeting, Chicago (2014)
15. Madhavan, P., Wiegmann, D.A.: Effects of information source, pedigree, and reliability on operator interaction with decision support systems. Hum. Factors **49**(5), 773–785 (2007)
16. Thompson, E.R.: Development and validation of an internationally reliable short-form of the positive and negative affect schedule PANASJ. Cross Cult. Psychol. **38**(2), 227–242 (2007)