Experimental Designs for Testing Metal Detectors at a Large Sports Stadium

Christie Nelson
CCICADA
Rutgers University
Piscataway, New Jersey

Paul Kantor
CCICADA
Rutgers University
Piscataway New Jersey

Dennis Egan
CCICADA
Rutgers University
Piscataway, New Jersey

Alisa Matlin
CCICADA
Rutgers University
Piscataway, New Jersey

Brian Nakamura
CCICADA
Rutgers University
Piscataway, New Jersey

Brian Ricks
CCICADA
Rutgers University
Piscataway, New Jersey

Fred Roberts
CCICADA
Rutgers University
Piscataway, New Jersey

Michael Tobia
Independent Researcher,
Dept. of Homeland Security
Washington, D.C.

Ryan Whytlaw
CCICADA
Rutgers University
Piscataway, New Jersey

Michael Young
Consultant
Brecksville, Ohio

Abstract— **When utilizing metal detectors at a large venue such as a sports stadium, there are the competing objectives of accuracy of the patron screening and the speed of throughput. This research, carried out in collaboration with the security staff at MetLife Stadium in New Jersey as well as other stadiums, analyzed two patron screening methods: handheld metal detectors ("wands") and walk-through metal detectors ("walk-throughs"). An initial experimental design was created to understand the effectiveness of wanding. This design was used with MetLife Stadium security during three training sessions. The data collected was used to understand (a) if the prohibited item was found and (b) in how many attempts. Various prohibited as well as allowable metal items were hidden at random in various locations on the body of individuals, who were then scored based on the importance weight of the item (guns were given more weight than keys for example). Trainees were then assigned a performance score based on speed and accuracy and were tested until they reached a minimum required score. Building on this initial experiment, a second more formal experiment was created to help MetLife Stadium staff understand how walk-throughs would perform outdoors at different security settings. This experiment focused less on training security staff and more on understanding the performance of walk-throughs in real situations (as opposed to idealized lab situations). This experiment was created to understand the walk-through performance at each setting in the outdoor environment; e.g., does a walk-through catch each of the pre-specified prohibited items, and is this consistent across machines on the same setting? Because of the number of factors to be considered (type of item, location, orientation, walk-through setting, etc.), designing the experiment required a sophisticated approach called Combinatorial Experimental Design. The experiment was part of two DHS-supported projects on best practices for stadium security.**

Keywords: walk-through metal detectors; experimental design; Combinatorial Experimental Design; stadiums; large venues; effectiveness; training; handheld metal detectors; patron screening;

I. INTRODUCTION

Security officials have several objectives as they screen patrons at large venues such as sports stadiums. The accuracy of the patron screening both competes with the patron or fan experience and can directly impact the speed of throughput. The research to be presented in this paper was carried out at MetLife Stadium in New Jersey, including field observations and utilizing MetLife Stadiums' facility for experiments, along with observations at other large stadiums. Patron screening methods based on two technologies were examined: handheld metal detectors ("wands") and walk-through metal detectors ("walk-throughs").

Initially, an experimental design was created to understand the effectiveness of the wanding process. This design specifically targeted the training of wanding. This design was used during three training sessions of MetLife Stadium security employees. Data was collected to document accuracy, including if the prohibited or allowable metal item was found, and if so, in how many attempts.

Building on this initial experiment with wands, a second more formal experiment was created for walk-throughs. It was designed to help MetLife Stadium and other large sports venues understand how walk-throughs would perform in a natural outdoor environment at different security level settings. This experiment focused less on training security staff and more on understanding the performance of walk-throughs in real situations (as opposed to idealized lab situations).

The outline of this paper is as follows: first, patron screening at large venues is discussed. Next, background on handheld metal detectors ("wands") and walk-through metal detectors ("walk-throughs") is presented. Then, the details of the wanding training experiments are presented. Experimental designs for walk-throughs along with initial results are presented. The paper concludes with suggestions for additional research.

II. Background

A. Patron Screening at Large Venues

There are many different types of patron screening techniques at large venues such as stadiums. The most frequently used patron screening techniques are visual inspection, patdown, wanding, and walk-throughs. Visual inspection simply involves looking at patrons in crowds for prohibited objects or suspicious behavior. Patdowns are performed by security personnel using their hands to pat down a person to feel for objects such as knives or guns. Wanding is using a handheld metal detector held a few inches from a patron's body to search for metal objects. Walk-through metal detector screening requires patrons to walk through a machine that detects certain levels of metal. For the purposes of this paper, we will be focusing solely on wanding and walk-throughs.

When screening patrons in various operational environments, higher security/sensitivity settings on the device (wand or walk-through), enable it to detect smaller metal objects. (Later, we will distinguish between security settings and sensitivity settings.) Some use cases of metal detectors, both wands and walk-throughs, are in indoor environments such as courthouses or airports where screening accuracy is more important than the patron experience. However, in a large sports venue or stadium scenario, screening time is extremely important. Not only do stadiums want patrons to be safe, but they value the fan experience as well, which includes time taken waiting in line to go through security screening. Therefore, there is an important balance between the security level/sensitivity of the screening and the time taken to screen each person. Another issue is nuisance alarms. Nuisance alarms are caused by metal objects that are not prohibited (e.g. cell phones). A high nuisance alarm rate can cause a significant slow-down in the screening process. This makes determining an appropriate security/sensitivity level on the walk-through machines especially important when dealing with large crowds all trying to get into a stadium at once.

B. Hand-Held Metal Detectors

Hand-held metal detectors are small, portable, and battery operated mechanisms to detect metal. These devices may or may not have adjustable security/sensitivity settings, depending on the make and model. Sensitivity is often determined simply by how far away from the patron's body the device is held. "By moving the wand of a hand-held metal detector around and close to a scannee's body, the operator can fairly accurately locate sources of metal that may be on, or even in, a person's body. When a suspect area is located, the hand-held device will generally give off an annoying squeal" [7]. A common use of hand-held metal detectors is secondary screening at airports, when a person has set off an alarm and security staff want to find the source of the alarm. While wanding is an often used screening method, it does have some downsides in sports venue applications. Specifically, by our initial observations, the time taken to properly screen patrons with wands is quite a bit faster than with walk-through metal detectors. Choosing a wanding approach means security has to deal with additional issues, including the inconsistencies between physical hardware (wands); potential inconsistencies between wanders (people); not as effective as walk-throughs; more invasive to guests than other methods such as walk-throughs; and there is more training required for wanding.

C. Walk-Through Metal Detectors

A walk-through metal detector is "a free-standing screening device having an electromagnetic field within its portal structure (aperture) for detecting metallic objects, including some nuclear shielding materials, carried by persons walking through the aperture" [2]. Walk-throughs tend to be less invasive than wands for patron screening, and quite a bit faster to screen patrons than with wants, which would appeal to those emphasizing the fan experience for stadium venues.

Walk-throughs also have the ability to change the amounts of metal they detect, in comparison with wands, many of which do not have this capability. As such, walk-throughs can be used in large venue situations, targeting whatever amounts of metal the venue prefers to target (e.g., based on a risk assessment). Obviously, the smaller the quantity desired to be detected, the slower the screening would take on average. So there is also a balancing act to determine an appropriate level of security/sensitivity that is sufficient without being too slow, which is what the second part of the paper will discuss.

III. Wanding and Effectiveness

The first part of this research focused on hand-held metal detectors (wands). This initial research was performed at MetLife Stadium and specifically focused on training of security staff. Traditional training practices include instructional demonstration videos on proper wand use, written exams to test comprehension, and practical examinations to verify staff competence with the equipment. We designed the practical testing procedure experimentally to develop a reliable assessment instrument to evaluate user capability with the hand-held metal detectors. This was implemented at three separate staff training sessions of MetLife Stadium security staff. Evaluation data measured (a) whether the prohibited items were located by security staff trainees and (b) the number of attempts a trainee undertook before successfully locating the item. Various prohibited and allowable metal items were hidden at random in various locations on the body of individuals who were screened by the trainees. Screeners were given a designated time limit within which to screen a group of 'practice patrons'. Trainee performance was scored based on several factors: 1) the importance weight of the item (guns were weighted higher than keys for example), 2) whether or not the item was located by the trainee, and 3) the number of attempts until success. Trainees were assigned a performance score based on these factors which reflects the speed and accuracy of trainee screening performance. Those who were unsuccessful were re-tested until they reached a minimum required score.

The practical training procedure includes the following: first, trainees are given ample time to practice wanding; next

978-1-4799-1737-2/15/$31.00 ©2015 IEEE

they proceed to the testing area to demonstrate screening ability during a timed session. Performance scores are calculated at the end of the allotted time. Trainees instantly receive a pass/fail score and are re-tested until they achieve a minimum score. For the practice, trainees are encouraged to practice the wanding technique on one another. The training area is separate from the testing area to ensure that all item locations are unknown to the trainees. Item locations are rotated at random for each trainee so that trainees are unable to disclose the location to their co-workers. The testing environment is non-competitive. It is valuable to note that because the staff work as a team, they are able to assist one another in gaining competence during the training session. Trainees are tested by screening a set number of patrons, with and without target items, within a limited time window, monitored by stopwatch. The goal for the screener is to find as many of the prohibited items as possible within the allotted time. Scoring is evaluated as follows: trainees are given a score based on the weighted value of the number of items found within the designated time; if all practice patrons are screened before the designated time, the total screening time is recorded. See Figure 1.

Trainer: Date

	CANDIDATE		First Exam	
	Last Name	First Name	Number Found	Time
1				
2				
3				
4				
5				
6				
7				

Retest		Second retest	
Number Found	Time	Number Found	Time

Figure 1. Wanding Trainee Scoring Table

The development of such a practical training exercise to evaluate staff effectiveness with hand-held metal detectors is the first to be reported in the field and is novel research. The hand-held metal detectors are widely used at stadiums and stadium staff are normally provided with an instructional video demonstrating proper tool use, along with a comprehensive written exam, but they often do not include a practical examination. The practical exam helped to increase staff's familiarity with the hand-held metal detector. Further research on the practical training plan could demonstrate the approach's validity and reliability by testing at additional sites and within multiple training sessions. After validation, the training plan will be available for wider dissemination to venues that wish to implement a similar standard of practical training. The approach and design of this training provided the inspiration for further experimental procedures concerning the use of walk-through metal detectors, which we describe in more detail in the next section.

Using the concept of Orthogonal Latin Squares of order 7, using the CRC handbook, a screening schedule was created, shown in Figure 2. Depending on class size, an appropriate number of testing groups is formed. We show only one of them here. Note that each of the teams will need a specific number of pieces of contraband or test objects. People who do not pass will see the red team with *different* contraband when they are retested.

	Person Number						
Round	**1**	**2**	**3**	**4**	**5**	**6**	**7**
1	1 A	2 S	3 B	4 N	5 A B	6 S A	7 BPS
2	2 B	3 N	4 A B	5 S A	6 BPS	7 A	1 S
3	3 A B	4 S A	5 BPS	6 A	7 S	1 B	2 N
4	4 BPS	5 A	6 S	7 B	1 N	2 A B	3 S A
5	5 S	6 B	7 N	1 A B	2 S A	3 BPS	4 A
6	6 N	7 A B	1 S A	2 BPS	3 A	4 S	5 B
7	7 S A	1 BPS	2 A	3 S	4 B	5 N	6 A B

Figure 2. Notional Screening Schedule for the Red Team

Possible locations for contraband are not reported here for reasons of operational security. The seven combinations of contraband locations are also not reported here for reasons of operational security. They are represented in Figure 2, but we do not reveal the coding scheme here. Experience with this system shows that it is possible to achieve operational testing of trainees in a very compressed time scale, suitable for use in a training session that lasts no more than several hours. Details of the operational aspects and specific numerical values may be obtained from the authors, by suitable security agencies.

IV. Walk-Through Metal Detectors and Effectiveness

A. *Initial Approach*

Building on this initial work with wands, an experiment was created to help MetLife Stadium and other large venue sports stadiums understand how walk-throughs would perform in a natural outdoor environment when set at different security levels. This is novel research, as even though much work has been performed at controlled indoor settings, such as labs, airports, courthouses, schools, etc., outdoor settings with environmental issues along with issues of large numbers of patrons having to be screened in short amounts of time have not been researched in this capacity. This experiment focused on understanding the performance of walk-throughs in real outdoor and stadium situations (as opposed to idealized lab situations). For example, does a walk-through catch each of the pre-specified prohibited items, and is this consistent across different machines at the same venue on the same setting? Is this consistent across venues? Do machines perform differently in certain types of weather? Because of the number of factors to be considered (type of item, location, orientation, walk-through setting, etc.), designing the experiment is extremely challenging. This approach required a sophisticated experimental design approach called Combinatorial Experimental Design to reduce the number of necessary test cases.

First, we observed a roll-out of walk-throughs installed at one gate at an international soccer event at MetLife Stadium to understand how walk-through metal detectors are used, and the challenges that the venues face when screening large numbers of patrons. We began to explore the question of how to determine the appropriate security level or levels that the walk-through metal detectors should be set on. In particular, we focused first on the National Institute of Law Enforcement and Criminal Justice (NILECJ) security levels 1 through 5. Note that in addition to the security levels, walkthroughs also can have different sensitivity levels. Indeed, for some of the walk-throughs we studied, there were 100 different sensitivity levels for each security setting. Experiments were designed to understand the difference between sensitivity levels. We return to this below.

We created an experimental design to help understand the practical differences between these security levels in a real field setting. The initial factors considered were three types of test objects (knife, gun, keys) at eight different locations on a person's body (hat; shirt sleeve – left or right; hand – left or right; behind belt; inside sock – left or right) - oriented one of three ways, with the person walking fast (1m/s) or slow (0.2 m/s), for a total of 144 (3 X 8 X 3 X 2) possible test conditions Note that the locations used in this experimental design differ from those used in the wanding approach. Thus, a complete factorial design that can evaluate all two-way interactions among the independent variables would require a large number (144) of test conditions, each of which must be run multiple times on multiple machines, and would be very costly and time consuming to run. Instead, we used a very efficient design approach called Combinatorial Experimental Design [11] or CED, which can assess the main effects and all pairwise interactions of the experimental factors.

Test Number	Speed	Object Location	Object Orientation	Test Object
1	Fast (1m/s)	Hat	Orientation 1	Gun
2	Slow (0.2m/s)	Hat	Orientation 2	Knife
3	Fast (1m/s)	Hat	Orientation 3	Keys
4	Slow (0.2m/s)	Shirt Sleeve – Left	Orientation 1	Knife
5	Fast (1m/s)	Shirt Sleeve – Left	Orientation 2	Keys
6	Slow (0.2m/s)	Shirt Sleeve – Left	Orientation 3	Gun
7	Slow (0.2m/s)	Shirt Sleeve – Right	Orientation 1	Keys
8	Fast (1m/s)	Shirt Sleeve – Right	Orientation 2	Gun
9	Fast (1m/s)	Shirt Sleeve – Right	Orientation 3	Knife
10	Slow (0.2m/s)	Hand – Left	Orientation 1	Keys
11	Fast (1m/s)	Hand – Left	Orientation 2	Gun
12	Fast (1m/s)	Hand – Left	Orientation 3	Knife
13	Slow (0.2m/s)	Hand – Right	Orientation 1	Keys
14	Fast (1m/s)	Hand – Right	Orientation 2	Gun
15	Fast (1m/s)	Hand – Right	Orientation 3	Knife
16	Slow (0.2m/s)	Belt	Orientation 1	Keys
17	Fast (1m/s)	Belt	Orientation 2	Gun
18	Slow (0.2m/s)	Belt	Orientation 3	Knife
19	Slow (0.2m/s)	Sock – Left	Orientation 1	Keys
20	Fast (1m/s)	Sock – Left	Orientation 2	Gun
21	Fast (1m/s)	Sock – Left	Orientation 3	Knife
22	Slow (0.2m/s)	Sock – Right	Orientation 1	Keys
23	Fast (1m/s)	Sock – Right	Orientation 2	Gun
24	Fast (1m/s)	Sock - Right	Orientation 3	Knife
25	no threat object			

Figure 3. Walk-Through Test Configurations

Applying CED led to 25 "configurations" that we needed to examine, which is only 17% of the original 144 configurations that were required (see Figure 3). Each of the 25 configurations were to be tested repeatedly. In the literature [7], in order for a setting to "pass the test", the following would be required: in a given configuration, 19 or more out of 20 trials must signal an alert if a threat object is present. If there is no threat object, 19 or more out of 20 trials must not signal an alert. For each configuration, if the setting is doing poorly (2 mistakes) we can stop testing and go on to the next configuration. In addition, a test could stop after 19 trials if there were no misses. One could ask if the 19 out of 20 requirement is too stringent for stadiums, but that is up to the user or to those who develop standards for stadium security, and could be the subject of future research as well.

B. *Revised Approach*

Prior to doing any formal tests, pilot experiments were performed to help determine what the more formal experiments should be. The various pilot experiments focused on height, horizontal location, and security setting for different test objects. Heights examined were ground level, knee height, middle of body height, shoulder height, and above head height. Horizontal location was that a test object was held close to the metal detector pole, and then moved away from the pole toward the center between the two poles further and further. The test objects used were standard NILECJ security level 3 (yellow) and level 4 (orange) test objects, see Figure 4. However, "test objects" could be "any metallic object used to evaluate the detection capability of a detector" [3]. We initially looked at different locations horizontally, and then later focused on holding the test objects in the center of the body horizontally between the two walk-through poles at random orientations. Security levels were focused on NILECJ levels 1 through 3, varying the security sensitivity within each level. Only one or two trials were performed for each experiment, as this was a pilot to help determine what the more formal experiments should be.

In addition to height, we also wanted to understand the effect of changed sensitivity levels. There are standard security settings on the walk-throughs, including NILECJ levels 1-5. We started with default NILECJ levels 1 through 3. Each of these default settings can have built in sensitivities (1-100). The tests were carried out for each of the security levels. We then increased the sensitivity higher than the standard setting for each security level and also looked at decreasing the sensitivity from the preset default for each NILECJ level. When testing, it is important to first make sure that the person performing the tests is a "clean tester," "a person who does not carry any objects which would significantly alter the signal produced when the person carries a test object" [3].

We found, for instance, in some sample trials that when increasing the sensitivity on level 2, that it was more often able to catch the yellow level 3 test object than at the standard level 2 sensitivity setting. We also observed from the initial pilot tests that the test objects became more difficult to detect as they moved away from the legs of the walk-through (toward the middle of the machine). In addition, from the initial pilot tests, it appeared that very high and very low heights were more difficult to detect than the mid-range heights. These hypotheses need further testing to confirm but were the basis of the revised approach.

978-1-4799-1737-2/15/$31.00 ©2015 IEEE

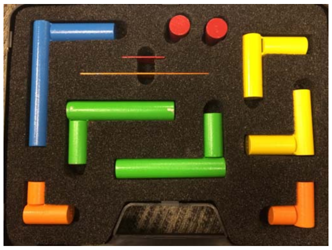

Figure 4. NILECJ-STD-0601 1-5 Test Objects

The pilot experiments revealed that the metal detectors may have certain areas that are more difficult to detect than others. So instead of testing every object at every height, and trying to cover many orientations of each object for each security setting, we could create fewer more targeted tests, which would be more feasible to perform in an operational environment. These would test our hypothesized areas which we suspect may be more difficult to detect; these areas are illustrated in Figure 5. We think this is the case, as in addition to our pilot test findings, detectors have the most overlap at mid-height and less sensitivity at very short or very tall heights; in addition, by their very nature, detectors are more sensitive close to their sides, so the center of the machine is less sensitive than the left or right positions horizontally; in addition, machines are calibrated in the field to not pick up metals in the ground, so the lower height may also be less likely to detect small metal objects. So rather than including all heights, we created tests to determine which were the most difficult heights to detect, and could then narrow down to a few specific heights to then test object orientation.

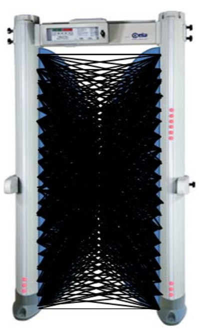

Figure 5. Illustrative example of walk-through metal detector areas of detection levels

A new two-phase test plan was created. The first phase involved testing NILECJ security levels 2 and 3. The experimental factors included two test objects: NILECJ orange 4140 test object, and NILECJ yellow 303 test object (see Figure 4). We are about to start on the new testing plan, phase 2. In this plan, the test object is held at the center of the body, with the tester trying to hold it close to the center of the machine (i.e. the center between the left and right posts of the machine). The object is then held at varying heights, starting at 0" from the ground, and then at 4" intervals until 7' (for 22 total heights). The orientations for these tests are either left or right facing, oriented so that the test object presents a paper-edge-like (thin as possible) side to the metal detector. (The object is perfectly in line with the two walk-through posts and not at an angle.) This is thought to be the most difficult orientation to detect, so we will be starting with these orientations. 20 trials will be performed for each test (or fewer: 19 if every test "passed"; or stop after 2 failed tests for a "failure"). For each test, the orientation is randomly either left or right. The experiments are shown in Figure 6.

Test	Test Item	Height
1	OrangeTestItem4140	0in
2	OrangeTestItem4140	4in
3	YellowTestItem303	8in
4	OrangeTestItem4140	1ft
5	YellowTestItem303	1ft4in
6	OrangeTestItem4140	1ft8in
7	YellowTestItem303	2ft
8	OrangeTestItem4140	2ft4in
9	YellowTestItem303	2ft8in
10	OrangeTestItem4140	3ft
11	YellowTestItem303	3ft4in
12	OrangeTestItem4140	3ft8in
13	YellowTestItem303	4ft
14	OrangeTestItem4140	4ft4in
15	YellowTestItem303	4ft8in
16	OrangeTestItem4140	5ft
17	YellowTestItem303	5ft4in
18	OrangeTestItem4140	5ft8in
19	YellowTestItem303	6ft
20	OrangeTestItem4140	6ft4in
21	YellowTestItem303	6ft8in
22	OrangeTestItem4140	7ft
23	None	

Figure 6. Planned Phase 1 Walk-Through Tests

The second phase of this experimental design is to determine the heights where there are the most failures. From there, one will limit this to a handful of specific heights, and test different orientations of the same two test objects (Yellow 303 and Orange 4140). These orientations will be predefined and require more field testing to determine. But it is thought that they will be similar to phase one, showing the thinnest edge of the object between the two left and right poles of the walk-throughs.

These tests will be carried out across walk-throughs at real venues in the locations of use. These locations include outdoor locations in all weather conditions, and indoor locations with difficult conditions such as interference from underground vibrations, etc. These walk-throughs can be battery powered or plug in, and will be set up as they are actually used at various venue locations. We have performed testing at MetLife Stadium and are in progress of performing testing at other locations.

This work demonstrates several unique approaches to examining patron inspection. One novel aspect is that we have made use of Combinatorial Experimental Design, a sophisticated experimental design tool that enables a small number of tests, to explore the levels of walk-through screening. Because the factors affecting successful screening are so vast, it would be a significant challenge for any venue to test for every possible relevant factor. Our use of advanced experimental design techniques has created a reliable sampling

978-1-4799-1737-2/15/$31.00 ©2015 IEEE

and testing procedure to address this issue. The objective of our research is to produce results that can be applied at any sport stadium or arena. Due to observed variations in screening based on factors such as weather, crowd and event type, and walk-through performance at different security and sensitivity settings, we recommend that each sport venue undertake such tests periodically in order to make sure their detectors are working properly and to test their calibrations. Without efficient experimental designs, this could not be practically feasible.

V. CONCLUDING THOUGHTS AND FUTURE WORK

When utilizing metal detectors at a large venue such as a sports stadium, there are the competing objectives of accuracy of the patron screening and the speed of throughput. This research, analyzed handheld metal detectors ("wands") and walk-through metal detectors ("walk-throughs"). Initially, a novel experimental design was created to understand the effectiveness of wanding. This design was successfully used at MetLife Stadium security during three training sessions. Data was collected to understand (a) if the prohibited item was found and (b) in how many "tries." Trainees were then assigned a performance score based on speed and accuracy, and were tested until they reached a minimum required score. Building on this initial experiment, a second experiment was created to help MetLife Stadium and other large venues understand how walk-throughs would perform in a natural outdoor environment at different settings. This experiment focused on understanding the performance of walk-throughs in real situations (as opposed to idealized lab situations). This experiment was created to understand the walk-through performance at each setting in the outdoor environment; e.g., does a walk-through catch each of the pre-specified prohibited items, and is performance consistent across machines on the same setting, across different venue locations, and how does performance change under more subtle changes of sensitivity level. We have illustrated the framework for these novel experiments.

The development of a practical training exercise to evaluate staff effectiveness with hand-held metal detectors is the first to be reported in the field. Further research on the practical training plan would demonstrate the validity and reliability by testing at additional sites and within multiple training sessions. In particular, a study could be set up to look at the amount of training time versus effectiveness of wanding in a real setting. The same could also be set up for walk-throughs to determine if training time affects screening effectiveness, and if there is a "sweet spot" in terms of length of training to acceptable effectiveness. After validation, the training plan will be available for wider dissemination to venues that wish to implement a similar standard of practical training.

Efficient testing protocols for analyzing performance of walkthroughs in real field settings are central to their effective roll-out in a large number of stadiums nationwide, as planned. To achieve effective usage, users at stadiums need to be able to run tests to compare performance in their own settings and to run tests to calibrate their machines routinely. Experimental tests of the kind we have discussed and will continue to develop are critically important in this rapidly growing area for use of walk-throughs.

Venues also must balance the security goals with economic and fan experience objectives. This means that the kinds of analyses reported here must be made as efficient as possible. This leads to a number of challenging research questions. What collection of threat objects and object locations is appropriate for a given venue and event? How frequently should training on the user of screening equipment be conducted? What is an acceptable level of performance for screeners for a given venue and event type? In testing the walk-throughs, is it technically feasible to combine several threat objects in a single walk-through, or must they be tested for separately at various positions? How many tests should be done of each device prior to a game, to achieve reasonable certainty that the device is working correctly?

Additional future work is to implement the walk-through experiments across venues in various locations, with different brands of walk-through machines, to understand the effectiveness of these machines in actual field tests. An important research objective is to build new metrics to measure effectiveness in this unique large outdoor venue environment.

ACKNOWLEDGMENTS

The authors thank the Department of Homeland Security for support under grants DHS-2009-ST-061-CCI002-05-CTR.OF EXCELLENCE and DHS-2009-ST-061-CCI002-06-CTR.OF EXCELLENCE to Rutgers University. In addition, we would like to thank MetLife Stadium for allowing us to use their facility to conduct experiments and field observations.

Any findings and opinions presented in this report are those of the authors, and do not necessarily represent the positions of any sponsors, venues, organizations or manufacturers.

REFERENCES

[1] Ricks, Nakamura, Almaz, DeMarco, Hui, Kantor, Matlin, Nelson, Powell, Roberts, Thompson, and Tobia. "Modeling the Impact of Patron Screening at an NFL Stadium." In Proc. of ISERC. 2014.

[2] ASTM. "Standard Guide for Installation of Walk-Through Metal Detectors." 2012.

[3] ASTM. "Standard Practice for Evaluation of Metallic Weapons Detectors for Controlled Access Search and Screening." 2010.

[4] NIJ. "Walk-Through Metal Detectors for Use in Concealed Weapon and Contraband Detection; NIJ Standard-0601.01." 2000.

[5] NIJ. "Walk-Through Metal Detectors for Use in Concealed Weapon and Contraband Detection; NIJ Standard-0601.02." 2003.

[6] Tobia, Michael A. "Public Safety Crowd Management." 2012 National and Educational Training Division, PREV-017.

[7] Green, Mary W. "The Appropriate and Effective Use of Security Technologies in U.S. Schools." National Institute of Justice. 1999.

[8] NILECJ. "NILECJ Standard for Walk-Through Metal Detectors for Use in Weapons Detection." 1974.

[9] CEIA. "Metal Detectors Utilities NILECJ-STD-0601.00 Security Levels 1, 2, & 3 Walk-through Metal Detectors."

[10] National Center for Spectator Sports Safety and Security. "Laboratory Assessment Report: CEIA Walk-through Metal Detector." 2010.

[11] Cohen, D., M., Dalal, S., Fredman, M. and Patton, G. C. The AETG system: An approach to testing based on combinatorial design. IEEE Transactions on Software Engineering, 23(7):437-444. 1997.

[12] Automated Efficient Test Generation System (AETG). User Manual. 2014. http://aetgweb.argreenhouse.com/doc

978-1-4799-1737-2/15/$31.00 ©2015 IEEE