NEED: Networking Activities for Enhanced Evacuation Drills: roadmap for enhanced evacuation drills

Amos, Martyn; Bénichou, Noureddine; Gwynne, Steve; Kinateder, Max

For the publisher's version, please access the DOI link below./ Pour consulter la version de l'éditeur, utilisez le lien DOI ci-dessous.

https://doi.org/10.4224/40001242

NRC·CNRC CONSTRUCTION

NEED: Networking Activities for Enhanced Evacuation Drills

Roadmap for Enhanced Evacuation Drills

Author(s): Martyn Amos (Northumbria University, UK), Noureddine Bénichou (NRC, Canada), Steve Gwynne (Movement Strategies, UK), Max Kinateder (NRC, Canada)

Report No.: A1-015092.1
Report Date: 30 September 2019
Contract No.: A1-015092
Agreement Date: 12 April 2019

National Research Council Canada | Conseil national de recherches Canada

Canada

NEED: Networking Activities for Enhanced Evacuation Drills
Roadmap for Enhanced Evacuation Drills

Author

Noureddine Bénichou, Ph.D. and Max Kinateder, Ph.D.

Approved

Ahmed Kashef, Ph.D.
Director, Research and Development
Fire Safety
NRC Construction

Report No: A1-015092.1
Report Date: 30 September 2019
Contract No: A1-015092
Agreement date: 12 April 2019
Program: Fire Safety

31 pages

Copy no. 1 of 1

Table of Contents

Table of Contents 5
1 Introduction and Background 6
2 Methodology 8
3 Training 9
3.1 What are the current methods used? 9
3.1.1 Drills 10
3.1.2 Summary of negative aspects identified 10
3.1.3 Teaching-based 10
3.1.4 Summary of negative aspects identified 11
3.2 What are the evaluation criteria for these approaches, and how do we measure them (that is, how might we assign “scores” for each criterion)? 11
4 Assessment 13
4.1 What are the current methods used? 13
4.1.1 Summary of negative aspects identified (with specific reference to drills) 13
4.2 What are the evaluation criteria, and how do we measure them (that is, how might we assign “scores” for each criterion)? 14
4.2.1 Cost 14
4.2.2 Safety 14
4.2.3 Actual versus expected performance 14
4.2.4 Adherence to protocol 15
4.2.5 Data collection 15
4.2.6 Consistency, accuracy and predictability of results 15
4.2.7 Robustness, adaptability and scope of procedures 15
4.2.8 User satisfaction with procedure 15
5 Discussion 16
6 Alternative Approaches 17
7 Next-Generation Evacuation Drills 19
References 21

NEED: Networking Activities for Enhanced Evacuation Drills Evacuation Drill Roadmap Document

Martyn Amos (Northumbria University, UK), Noureddine Bénichou (NRC, Canada), Steve Gwynne (Movement Strategies, UK) and Max Kinateder (NRC, Canada).

1 Introduction and Background

An *evacuation drill* is a pre-planned simulation of an emergency evacuation for a specific scenario, with the aim of improving and evaluating the performance of occupants and staff involved. Although drills are informed by safety legislation and building codes, their merits are still not well-understood, and their impact on evacuation performance has not been well-characterized (see for instance, Peacock, Averill, & Kuligowski, 2010; Peacock, Hoskins & Kuligowski, 2012; Proulx & Reid, 2006; Shields & Boyce, 2000; Sheilds, Boyce & McConnell, 2009). However, they are still seen as a key component of safety planning and building certification. There are *significant issues* with evacuation drills, as currently executed (Gwynne *et al*., 2016; Gwynne *et al*., 2017):

- The *effectiveness* of the evacuation drill model[1] is not well understood.
- Drills carry both an *inherent risk* to participants (that may be changing given demographic and social changes), and a *significant cost* (in terms of temporary loss of building functionality).
- Sub-populations are often *excluded* from drills (e.g., those with medical issues, or mission-critical staff), which affect the potential for effective training and assessment.
- Drills are conducted inconsistently across organizations, building types and between jurisdictions.
- Drills are employed to train, to assess, or to do both.

However, the availability of new approaches and technologies such as augmented / virtual reality, computational simulation, smart sensors, building intelligence and video analytics means that we have an *unprecedented opportunity* to enhance the way that we plan, deliver and analyze the results of evacuation drills or complementary activities, and to improve evacuation performance (and the assessment of such) (Feng, *et al*., 2018; Lovreglio *et al*., 2017; Williams-Bell, *et al*., 2015). This allows us to disentangle the training benefits and assessment of drill

[1] A model is a simplified abstraction of a real world phenomenon. Evacuation drills are one attempt to model emergency evacuation; others include computation simulations (“evacuation models”), behavioral experiments, or questionnaires.

effectiveness, following a *case-control* approach comparable to research practices in evidence-based medicine - eventually moving towards *evidence-based evacuation drills* (Kinateder *et al.*, 2014; Kinateder & Warren, 2016).

Here, we explore new technologies, methodologies and perspectives to (a) enhance the training component of evacuation drills, (b) improve the analysis and interpretation of their results, and (c) reduce both short-term risk to participants and operational disruption. Working towards these objectives may bring a range of beneficial outcomes, such as potentially reducing costs, improving training effectiveness and allowing more fine grained assessment of occupant behavior.

The NEED (Networking Activities for Enhanced Evacuation Drills) project established a consortium of experts in evacuation dynamics, fire safety, immersive tech, data science, artificial intelligence, and computational simulation, to focus on the following questions:

- What are the main *training limitations* of current evacuation drills?
- What are the main *assessment limitations* of current evacuation drills?
- How do these limitations vary across the building occupancy types and populations that are subject to evacuation drills?
- What are the *alternative approaches* for drill training and assessment?
- What are the *major benefits* of the alternative approaches identified? How can they be quantified?
- What are the *challenges* (regulatory, public perception, technical, cost etc.) in adopting such alternatives?
- What are the *next steps* for ensuring these challenges are further addressed through collaboration between UK and Canadian partners?

By answering this set of questions, NEED addressed its *key objectives*:

- Identify significant *outstanding issues* in evacuation drills
- Highlight *possible approaches* to addressing these issues.
- Provide a *framework for sustainable UK-Canada collaboration* on these and other issues, and identify future opportunities.
- Develop *engagement with non-academic stakeholders* (e.g., small and medium-sized enterprises, policy makers), and help maximize the possibility of uptake and impact.
- Encourage participation of *early-career researchers*/students.

This roadmap document serves as the definitive output of the NEED project, capturing the first two key objectives, and providing a foundation for achieving the other three. Our intention is that this will inform “next generation” evacuation drills. Although we use the term "roadmap", we differentiate the current document from more traditional technology-driven roadmaps, as our focus includes procedural and legal issues.

The rest of the document is structured as follows:

- We first describe our methodology for obtaining a diverse set of input into the roadmap document, and explain how it was refined into a core set of fundamental issues concerning evacuation drills.
- We then consider each of these issues in turn (based on discussions within our workshops), compare and contrast our findings with existing research, describe possible alternatives to the *status quo* that have been identified, and then conclude with a proposed plan of work for "next-generation" evacuation drills.

We hope that this document will serve as both a summary of issues with the current state-of-the-art, and a foundational platform for the development of innovative new approaches to evacuation training and evaluation.

2 Methodology

The project was structured around a number of in-person workshops with all in all 30 experts from 20 different organizations, held as follows (see Appendix A for full participant lists):

- Kick-off meeting/workshop, Northumbria University, Newcastle upon Tyne, UK, December 6-7 2018.
- Workshop, NRC, Ottawa, Canada, January 7-8 2019.
- Workshop, NRC, Ottawa, Canada, March 19-20, 2019.
- Wrap-up workshop, Movement Strategies, London, UK, April 29-30 2019.

The purpose of the workshop sequence was to iteratively refine the roadmap document, and to prepare follow-on activities. At the first workshop, we agreed that a useful starting point for the roadmap document was offered by the questions outlined in Gwynne, *et al*., 2017:

- How do we exploit egress drills more effectively to *train* and *assess* performance?
- What *alternative models* are available to influence and assess evacuee performance?
- How might we use egress drills and alternative models together?

Section 7 of the paper by Gwynne, et al., 2017, gives an initial comparative analysis of the relative strengths and weaknesses of various egress models, and it was agreed that a core component of the roadmap should be to both augment and extend this table, with various cost / benefit analyses. We supply this table, for reference, in Appendix B. In order to provide foundational input into the roadmap, over the course of the *first two workshops* we identified the following (non-exhaustive) list of questions:

- Who (in terms of organizations) is *currently performing* “realistic” evacuation drills? On whose expertise might we draw?
- What are the current relevant *standards* for evacuation drills, and how did they evolve / emerge?
- What are the *current barriers* (for organizations) to carrying out effective drills?

- How do we engage with *participants* in drills (as opposed to just those responsible for organizing drills)? How often are building occupants consulted?
- How do we engage with *responders*?
- Do drills *actually work* in terms of providing existing training? If they demonstrably do work, then which features of the drills are effective?
- More generally, what *metrics* are needed to measure the effectiveness of drills?
- What information is needed for *regulators* to make informed decisions about drills? How would we go about obtaining this information?
- How should our findings be made "*adoptable*" (e.g., by ISO or other standards organizations)? How should we be informed by current regulatory structures / protocols?
- Who is the target audience for the roadmap? Regulators (NBC, ISO, ICC), safety, resilience, and emergency managers, …

From this, we established the *three main areas of focus* for the current roadmap document:

- *Training* aspects of evacuation drills; how are they used to *prepare* individuals?
- *Assessmen*t aspects of evacuation drills; how might we *evaluate the performance* of drills (and of individuals)?
- What are the *alternative* approaches to existing evacuation drills?

In order to explore these issues, during the *third workshop*, we asked participants to consider the following questions - in what follows, we summarise the responses received:

- For training and assessment, (1) what are the methods *currently used*, and what are their main *positive and negative aspects*? (2) What are the *evaluation criteria*? (3) How *important* (relatively) are the criteria?
- What *alternatives* to existing methods might we consider?

3 Training

3.1 What are the current methods used?

Existing methods are often a combination of approaches described below, and are applied at different levels (task-oriented, individual acts, procedural enactment) and to different groups. They are often scheduled at different frequencies, depending on assessed risk, occupancy type, and the role(s) of the population under consideration.

As a result of our survey process, we essentially partitioned existing methods into two categories: (1) "traditional" *drills*, and (2) *teaching-based* approaches. While both categories are inherently artificial, the first is more experiential in nature, while the second tends to emphasise theory and pedagogy. We also see an emerging third category, based on immersive / game-based learning, which attempts to bridge the gap between the other two.

All of these methods essentially explore "what if?" scenarios, in order to establish the individual capacity to adapt and to make decisions under different conditions. However, there is the potential for different scenarios to be overlooked / excluded, or to be conflated. The performance of specific actions is related to individual roles within an evacuation procedure, and may be sensitive to specific factors related to the role, organization, procedure or occupancy.

3.1.1 Drills

These are often seen as the “standard” method, with Fire Evacuation Officers (FEOs) assisting occupants out of a building. We see a combination of *announced* (warning provided ahead of time) and *unannounced* drills. Partial information is often provided, indicating that something will occur, but not precisely when. We may also see partial drills, based on *zonal evacuation*. Recording of drills provides opportunities to support subsequent debriefing and / or simulation studies, and may allow for comparison between individual actions and strategic objectives. Drills allow us to test whether or not processes are actionable. Drills also are used to provide knowledge of a building by forcing occupants to experience areas of the building that they were not previously aware of. We also see “FEO-only” drills, involving only key personnel.

3.1.2 Summary of negative aspects identified

- Full-scale drills pose serious *safety challenges* when conducted in cities – people can be evacuated into traffic, for example. Methods are often selected based on risk of injury, rather than on the training benefits provided.
- *Different occupancies* have different population / procedures / staff / risks, even within the same occupancy class. What are the physical / emotional consequences of including / excluding certain sub-populations (e.g., those with mobility impairments)? Implications range from a perception of being patronised to degrading / uncomfortable / dangerous experiences.
- When releasing partial information, it is difficult to ensure that information does not spread beyond a target population.
- *Loss of building functionality* during a drill can incur significant overheads (financial costs - up to £50K per hour in some instances - and organizational / logistical costs may accrue in terms of loss of functionality and downtime, and impact on business continuity).
- If drills last for longer than 20 minutes then participants may lose focus, and either mentally or physically disengage – somewhat undermining the purpose of the event. There are also considerations of *discomfort and impact on health* (physical and emotional) – this precludes consideration of a number of realistic scenarios.

3.1.3 Teaching-based

A complementary approach is based on more *pedagogical* methods, using PowerPoint presentations, videos, online training (e.g., 15 minute module, taken annually, with as many attempts as required for a pass). It may be the case that a relatively “low tech” playback of a video (showing consequences of decisions, etc.) is sufficient to inform / practice decision making and provide feedback to participants. It is also possible to generate relatively realistic time-induced pressure on decision making using table-top or simulation approaches. Such methods are generally seen as flexible and convenient, relatively low cost, and easy to use.

They also allow for easy individual recording of completion of mandatory training (by keeping a register or online log of who has completed it).

3.1.4 Summary of negative aspects identified

- Mostly *knowledge-based* – they do not provide the physical experience of an actual evacuation, and may therefore be seen by participants as less "realistic". It is therefore unclear how well the training transfers to real situations in the physical world.
- Teaching methods are often based on *solitary completion*, and do not require any *interaction* with others. This may be significant in a real world scenario, as inter-personal interactions will play a role.
- *Costs* involved (employee time, cost of creating teaching materials).
- Teaching-based methods are often viewed by participants as an unmotivating "*box ticking*" exercise, which may diminish the impact of training.
- Such methods often do not require meaningful engagement with the material, and it is possible for participants to "pass" without active learning. In other words, it can be unclear how much information trainees are retaining and applying to their own personal situation.

3.2 What are the evaluation criteria for these approaches, and how do we measure them (that is, how might we assign "scores" for each criterion)?

Although we have attempted to make these criteria as general as possible, it is clear that many of them will only apply to traditional evacuation drills, and not to teaching-based approaches. The training format will also influence the relevance and weight of the criteria; for instance, it is often thought that drills are most effective when they are unannounced, so convenience (criterion 7) should probably be disregarded in this case. However, with any other form of training, reducing inconvenience to attend will improve participation.

We present these criteria in *descending order of perceived significance*, according to the aggregated view of our workshop participants (noting that this is, of course, an approximate ranking).

We should note that this is only ever a broad approximate ranking, and that the specific ordering of criteria will differ according to the organization, building type, population(s), etc. For example, if we consider a nuclear power plant, then risk (to health and safety in case of emergencies), will be more highly-weighted than cost (we would hope); whereas a stock exchange building might carry relatively low risk but incur hugely significant financial costs if impacted by a drill. Also, not all criteria will be appropriate or relevant to all scenarios and environments.

1. Adherence to regulatory requirements.

Does the activity allow the organization to fulfil its legal and other regulatory obligations, and / or to be seen to do so?

2. Achievability of training objective(s).

Will the activity take place? Will participants engage with the training, or will they potentially "switch off" or, even worse, avoid participating altogether? Can occupants complete the activity within a target time frame? Are participants equipped to be clearly aware of situations, procedures, and could they make appropriate decisions and respond adequately? Could they be able to recall the appropriate objectives, procedures and actions? Could people identify and later remember what they are supposed to learn from the activity?

3. Scenario coverage / transferability / adaptability.

Are a number of different scenarios covered? Are general principles taught that are applicable to different scenarios, or is the activity specific to one building/site?

4. Cost/opportunity cost.

What financial costs are imposed on the organization, in terms of loss of productivity, business activity, etc? Are there any opportunity costs? What are the delivery costs? What will be the cost (e.g., fines, potential incident) of *not* performing the activity? How much time (in terms of person hours) does the activity require?

5. Risk (physical/emotional/dignity), safety, and ethical concerns.

Are we able to account for all participants? Are all participants able to travel to a safe place at the completion of the activity? Are participants (un)comfortable during the activity? Could the activity lead to injuries or distress?

6. Reach / inclusion of populations.

Are the target population(s) reached with the activity? Are any populations excluded? Are participants consulted on the development and delivery of the activity?

7. Convenience and accessibility.

How easy do participants find it to comply with the requirements of the activity? Are participants able to complete the activity at a time / place convenient to them? Is the activity easy to understand, and are there any physical / logistical barriers to completion? Are there any personal costs (in terms of inconvenience, dignity, etc. or impact on work performance) associated with the activity? This criterion is concerned more with the *format* of the exercise than its *content*.

8. Performance / reliability / availability of procedural infrastructure.

Do communication channels (e.g., alarms, signage, and announcements) and the information provided therein work effectively? Does computer training work correctly? Are there any technical issues? Are all required facilities available at the time they are needed?

9. Data collection and subsequent actions.

Is it possible to automatically gather data on participant responses, movement or behaviour, or is it manually recordable? Is feedback obtained from / given to participants? Is the activity correctly “set up” to facilitate assessment of its effectiveness and performance?

4 Assessment

4.1 What are the current methods used?

By far the most common method for assessing procedures is the *traditional evacuation drill*, and this often focuses on *final outcomes* (i.e., the final evacuation time - were occupants evacuated from the building safely and within a specified time limit?), without considering *the internal dynamics* of the exercise. However, analysis of exercises does sometimes consider qualitative aspects (such as “orderliness”), as well as quantitative aspects such as pedestrian flow and / or exit usage, speed of movement, etc.

Drills are useful in that they require participants to *physically enact* procedures. This has two main benefits: (1) they create *observable behaviour* (generating usable data), and (2) they are capable of generating *uncertainty / randomness* that may not be captured or encoded in simulations or non-embodied methods. They also have the potential for the greatest possible level of *ecological validity* (for a range of scenarios).

Other methods that are used (less commonly) include computational simulation (often in conjunction with a traditional drill), or micro-level assessment of specific components of evacuation (such as the evacuation of people with mobility impairments), as well as treatment of hypothetical situations using mechanisms such as table-top exercises, walk-throughs, laboratory experiments, and virtual / augmented reality experiments.

When a traditional drill is used, a variety of data capture methods may be employed to record information about its execution. These include CCTV / video capture, head counts, marshal feedback, regulatory body observers, surveys (e.g., questionnaires), post-event interviews and debriefing. The data collected varies according to the method employed.

4.1.1 Summary of negative aspects identified (with specific reference to drills)

- Existing methods for assessing evacuation procedures are very *outcome-driven*, in that they tend to treat the activity as a “black box”. If undue focus is placed on “bottom line” metrics, then this may lead to an implicit / explicit bias in the design of the exercise.
- There is always the issue that a drill may fulfil (sometimes by accident) a specific criterion without participants necessarily adhering to procedures. In essence, it would then not be clear whether the outcome (successful or otherwise) was achieved through

the use of the intended procedure, or simply coincidentally, given other non-procedural factors (Table 1).

Table 1 Evaluation matrix for evacuation procedures

		Criteria Met?	
		YES	NO
Procedure Followed Accurately?	YES	☑	☒
	NO	☒	☒

- Drills do not generally interrogate all procedures, operational environments or scenarios, because of the exclusion of certain populations or execution in favourable conditions, announced versus unannounced status, and so on, so data collection is inherently partial. We must also consider the extreme cost of examining all possible factors that might affect performance.
- Because of the nature of drills, it is very difficult to establish longitudinal reliability, as they are effectively "one off" events.

4.2 What are the evaluation criteria, and how do we measure them (that is, how might we assign "scores" for each criterion)?

For this set of criteria, we argue that it is more difficult to obtain a definitive *rank ordering* of importance, as the ranking is dependent on both the organization and the scenario in question, so we simply list them here. Most of these criteria will apply (non-exclusively) to drills, and some may be more or less applicable to other delivery mechanisms, such as simulations, table-top exercises, etc.

4.2.1 Cost

What is the cost to the organization of performing the drill? Are there any non-financial costs involved (e.g., inconvenience to customers, reputational cost)? What are the direct resources required to conduct the drill, and what is their financial cost?

4.2.2 Safety

Is the drill safe to perform? What is the increased risk to those performing in the drill in comparison with their situation prior to the drill being performed? How are the inherent risks balanced against the potential benefits?

4.2.3 Actual versus expected performance

How do we define and operationalise "performance?" Total Evacuation Time? Floor clearance time? It is critical to define in advance what is being measured, and what is "acceptable". It is also important to understand that indicator values are only useful in the context of specific scenarios, and are often only *relative* (i.e., better or worse than previous performance), which provides only limited insights. Often, there may be no specific benchmark available in specific scenarios, unless there exists an *a priori* figure or scale. There may also be more qualitative indicators to consider, such as “orderliness”.

4.2.4 Adherence to protocol

Are all signals observed, and all paths used (if appropriate)? Is there a demonstrable understanding of systems and appropriate actions? Are all egress-related paths and exits used and taken? Are there any “non-standard” behaviours?

4.2.5 Data collection

Has the drill been documented sufficiently, with appropriate levels of granularity, scope and volume of data capture, to capture both individual behaviour, communication and cooperation between individuals and / or roles, and overall system-level performance? More fundamentally, has the evacuation protocol been designed in such a way as to *admit* an assessment of whether or not it has been followed by an arbitrarily-selected single individual? We might phrase this in terms of the following question: do we have enough data to allow us to recreate the scenario in an agent-based model?

4.2.6 Consistency, accuracy and predictability of results

How predictable and accurate are the results of the drill? Do we trust the data sources (e.g., individuals recording pedestrian flows)? If we repeat the same drill a number of times under similar conditions, do we obtain broadly similar results? How variable are the data across roles, individuals, groups, and the population as a whole? How context-dependent is the data?

4.2.7 Robustness, adaptability and scope of procedures

How critical are various aspects of the procedure to overall performance? How much tolerance does the procedure have to individual variability? If key components of the procedure are rendered unavailable, how well does the system cope? (How) do people react to dynamic changes in the scenario (e.g., lost staircase)? How critical are the emergence of "natural leaders" to the outcome? More generally, how wide a range of scenarios may be captured by the procedures?

4.2.8 User satisfaction with procedure

We might measure qualitative aspects such as the sense of self-empowerment, comfort, or safety reported by individuals. These may reflect both the “delivery” side of the drill (in which case, we would consider the self-reported confidence of staff to operationalise the procedures), and the “receiving” end, reported by evacuees.

5 Discussion

In Appendix B we present an excerpt from a recent paper that partially motivated the current study, and which provides one possible framework for the assessment of egress drills and possible alternative models / methodologies.

In particular, we are interested in the "considerations" listed in Table 2, which may be interpreted as assessment criteria. For the purposes of clarity, we list them here:

- Financial / organizational
- Ethical
- Perceived credibility
- Scope
- Potential insights
- Third party scrutiny
- Statistical
- Pedagogical

One of the key goals of the NEED workshop series was to obtain a *complementary set of criteria*, without biasing the discussion by presenting, *a priori,* the original criteria. We now show, in the table below, the alignment between the *original* criteria and those *subsequently* obtained through our workshop discussions. More criteria than the ones identified here are, of course, conceivable. For example, we might consider the degree of acceptance and availability of appropriate "technology", and the ability/willingness to use or adopt that technology by both trainer and trainees.

Gwynne, *et al.* criteria	**Our training criteria**	**Our assessment criteria**
Financial / organizational	Cost / opportunity cost. Convenience and accessibility. Performance / reliability / availability of procedural infrastructure.	Cost.
Ethical	Risk (physical / emotional / dignity), safety, and ethical concerns.	Safety.
Perceived credibility		User satisfaction with procedure.
Scope	Scenario coverage / transferability / adaptability. Reach / inclusion of populations.	Robustness, adaptability and scope of procedures.
Potential insights	Data collection and subsequent actions.	Consistency, accuracy and predictability of results. Actual versus expected performance.
Third party scrutiny	Adherence to regulatory requirements.	Adherence to protocol.
Statistical	Data collection and subsequent actions.	Data collection.
Pedagogical	Achievability of training objective(s).	

From this, we conclude that we have generated a robust set of criteria against which both training and assessment methods may be assessed.

6 Alternative Approaches

In previous work, Gwynne, *et al*. (2017) identified a number of alternatives to traditional egress drills, including simulations, laboratory experiments, immersive (Virtual / Augmented Reality) technologies, table-top exercises, mental rehearsals, walkaround sessions, briefings and scripted exercises.

In the radar diagrams below, we present one possible framework for comparing traditional egress drills with alternative approaches. For succinctness, we adopt the original criteria, noting their alignment with the criteria on an arbitrary scale we subsequently identified. Each "spoke" of the spider diagram corresponds to a criterion, which allows us to easily visualise the relative merits of two or more methods. For each spoke, the "weights" attached to each method are derived from Table 2 of Gwynne, *et al*., 2017 and expanded to a 7-point Likert scale ranging from 0 (no alignment) to 3 (perfect alignment) in 0.5 increments. The reader should note that these scales are based on the consensus reached in the NEED workshops; however, limitations of subjective data still apply.

In Figure 1, we show a comparison between traditional egress drills and computer simulations (or evacuation models). The traditional drill scores highly on "credibility" and "pedagogical" aspects, but carries relatively little statistical weight, and is limited in scope. On the other hand, simulations may yield significant insights, and are flexible (and thus broad in scope), but suffer in terms of credibility.

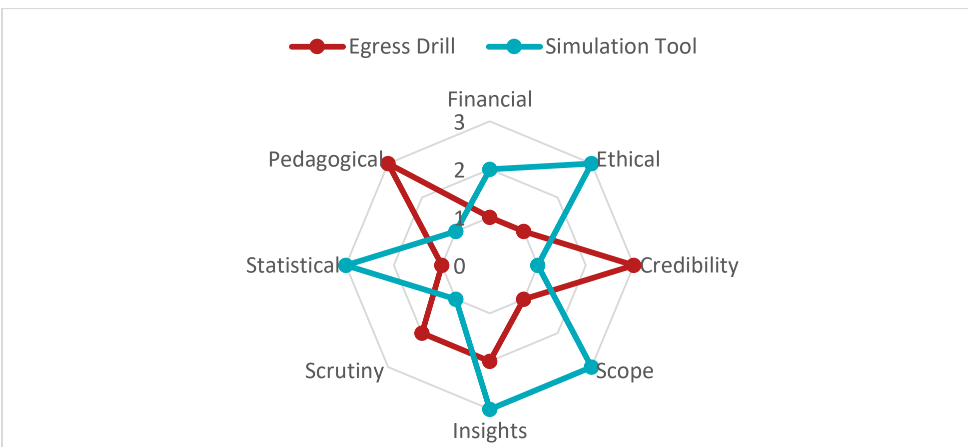

Figure 1 Radar diagram plotting alignment of egress drills and simulation tools with eight evaluation criteria. Data are presented on an arbitrary 7-point Likert scale ranging from 0 (no alignment with criterion) to 3 (perfect alignment with criterion).

We now present the remaining methods identified in the paper (Figure 2), in which egress drills are compared to (a) immersive technologies, (b) lab experiments, (c) briefings, (d) scripted exercises, (e) mental rehearsals, (f) walkarounds, and (g) table top exercises.

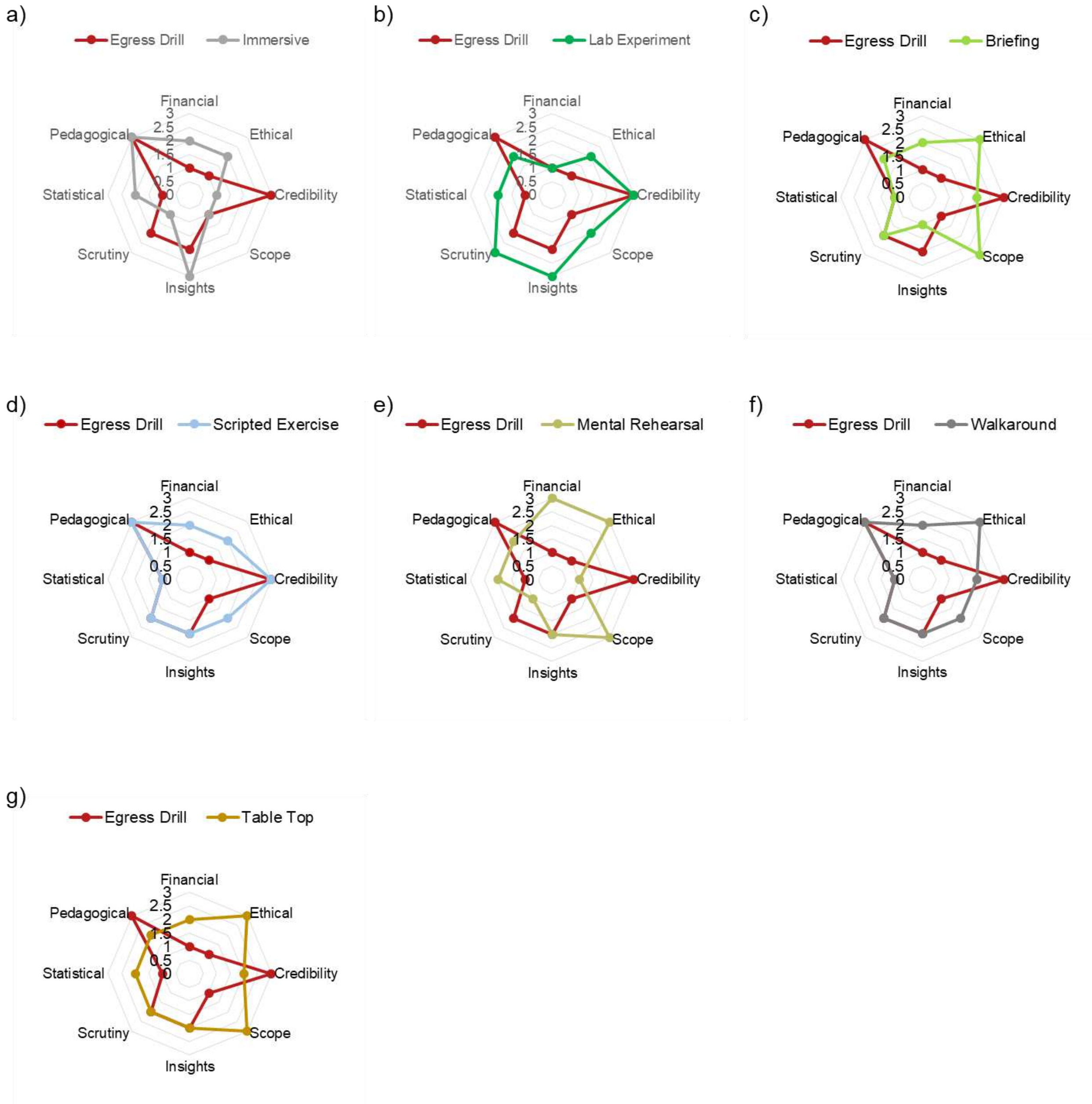

Figure 2 Radar diagrams plotting alignment with eight evaluation criteria of egress drills and a) immersive technologies (e.g., augmented or virtual reality tools), b) laboratory experiments (e.g., controlled observations of human behaviour in a laboratory environment), c) Briefing (e.g., meetings in which information about evacuation procedures are presented), d) scripted exercises (e.g., roleplaying type exercises in which participant assume certain roles and follow a predefined chain of events), e) mental rehearsal (e.g., imaginary exercise of what-if scenarios), f) walkarounds (e.g., inspection of egress routes by occupants), and g) table top exercises (e.g., theoretical rehearsal of evacuation scenarios conducted only by key personnel such as FEOs). Data are plotted as in Figure 1.

It is apparent that some of these approaches *dominate* the performance of egress drills; i.e., are at least as effective in all measures. However, specific approaches may not always be practical. Another approach is to suggest pairs that - *when combined* - then dominate egress drill performance (Figure 3).

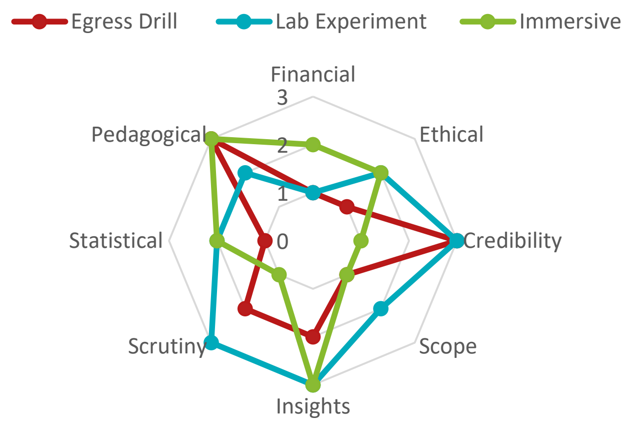

Figure 3 Radar diagram of egress drills, lab experiments and immersive technologies (data replotted from Figure 1 and Figure 2).

Here, we show how a combination of lab experiment and immersive experience might out-perform the traditional egress drill in terms of all criteria.

These comparisons suggest two immediate questions:

1. How might we best *quantify* the criteria that we have identified?
2. How might we obtain *evidence* for any scoring against quantifiable criteria?

In the next Section, we propose an outline programme of work that, we believe, will go some way towards addressing these questions.

7 Next-Generation Evacuation Drills

In this Section, we conclude the roadmap by summarising the main problems with traditional evacuation drills that we have identified, and propose one possible programme of work that might begin to address these.

Evacuation drills are pre-planned simulations of emergency evacuations for specific scenarios, with the aim of improving and assessing the performance of individuals involved. Although drills are informed by safety legislation and code, their *merits* are still not well-understood, and their impact on evacuation *performance* is not well-characterized (Gwynne *et al.*, 2016; Gwynne *et al.*, 2017). The benefits of such drills are unclear, given that:

1. Drills carry inherent *risks* to participants and significant *costs* (e.g., temporary loss of building functionality, added liability).
2. They are *inconsistently performed* and not fully exploited to meet their twin objectives of *training* and *performance assessment*. For example, sub-populations are often not included in drills such as mobility impaired occupants, or mission-critical staff, affecting the potential for realistic training / assessment.

3. The nature of potential emergency scenarios is constantly evolving (e.g., climate change extreme events, terrorism, active shooter incidents), as are population demographics.
4. Data is usually collected manually and often based on subjective assessments (e.g., manual timing; qualitative performance ratings)

Nonetheless, drills are still seen as key components of safety planning / building certification. Given this reliance, it is vital to understand whether drills, as currently performed, are fit for purpose, and if they need to be enhanced (or even replaced) in a cost-effective manner. *Emerging immersive / simulation technologies* offer potential alternatives to the existing drill model, and could potentially mitigate the challenges mentioned above. However, their effectiveness, credibility and validity need to be assessed against the existing approach. The findings from our NEED workshops have highlighted the *limitations* of the current approach and the *significant opportunities* offered by emerging technologies to improve the safety of building occupants, while reducing the negative side-effects of drills. We also identified a number of *criteria* against which different methodologies might be assessed (e.g., cost, realism, or credibility). We propose the following future work:

Goals

1. Establish an *evidence-based methodology* (case-control approach) for assessing evacuation drills and alternatives.
2. Harness *novel objective and automatable approaches to data capture and analytics* to better characterize performance (using smart sensors, artificial intelligence, computer vision, and machine learning).
3. Develop an *alternative to the current drill model*, based on emerging immersive / simulation technologies (e.g., virtual / augmented reality), and compare its performance to the *status quo*.
4. Develop *guidance for regulatory bodies* on the application and cost-benefits of each approach (e.g., relative performance gain, loss of individual / building time) for different scenarios.

Milestones

1. Develop a *test protocol* for evidence-based assessment of occupant evacuation performance during drills, to facilitate rigorous comparison of methodologies.
2. Establish *novel data capture and analytical methods* for in-person drills.
3. Develop a *technology-based training alternative* to existing drill model.
4. Develop a *simulation platform,* enabling scenario-basis for assessment, and generating more robust results through repetition.
5. Complete *controlled lab experiments,* testing training system.
6. Test the integrated system against the *status quo* in *real-world large-scale evacuations*, establishing the relative costs / benefits of our approach in terms of the criteria we have identified.
7. Develop *guidance,* in cooperation with regulatory systems (e.g., National Fire Code), on integration of evidence-based practices.

References

Feng, Z., González, V.A., Amor, R., Lovreglio, R. & Cabrera-Guerrero, G., 2018. Immersive virtual reality serious games for evacuation training and research: A systematic literature review. *Computers & Education*, 127, pp. 252-266. DOI: 10.1016/j.compedu.2018.09.002

Gwynne, S., Boyce, K., Kuligowski, E., Nilsson, D., Robbins, A., & Lovreglio, R., 2016. Pros and cons of egress drills. *Proc. 14th International Conference on Fire Science and Engineering (Interflam 2016)*, July 4-6, 2019, pp. 1657-1669.

Gwynne, S.M.V., Kuligowski, E.D., Boyce, K.E., Nilsson, D., Robbins, A.P., Lovreglio, R., Thomas, J.R. & Roy-Poirier, A., 2017. Enhancing egress drills: Preparation and assessment of evacuee performance. *Fire and Materials*. DOI: 10.1002/fam.2448.

Kinateder, M., Ronchi, E., Nilsson, D., Kobes, M., Müller, M., Pauli, P. & Mühlberger, A., 2014. Virtual reality for fire evacuation research. In *2014 Federated Conference on Computer Science and Information Systems*, pp. 313-321. DOI: 10.15439/2014F94.

Kinateder, M., & Warren, W. H. (2016). Social influence on evacuation behavior in real and virtual environments. *Frontiers in Robotics and AI,* 3, p. 43. DOI: 10.3389/frobt.2016.00043.

Lovreglio, R., Gonzalez, V., Amor, R., Spearpoint, M., Thomas, J., Trotter, M. & Sacks, R., 2017. The need for enhancing earthquake evacuee safety by using virtual reality serious games. In *Lean & Computing in Construction Congress*, pp. 1-9.

Peacock, R.D., Averill, J.D. & Kuligowski, E.D., 2010. Stairwell evacuation from buildings: what we know we don't know. In *Pedestrian and Evacuation Dynamics 2008*, pp. 55-66. Springer, Berlin, Heidelberg. DOI: 10.1007/978-3-642-04504-2 4.

Peacock, R.D., Hoskins, B.L. & Kuligowski, E.D., 2012. Overall and local movement speeds during fire drill evacuations in buildings up to 31 stories. *Safety Science*, 50(8), pp.1655-1664. DOI: 10.1016/j.ssci.2012.01.003.

Proulx, G. & Reid, I.M., 2006. Occupant behavior and evacuation during the Chicago Cook County Administration Building fire. *Journal of Fire Protection Engineering*, 16(4), pp. 283-309. DOI: 10.1177/1042391506065951.

Shields, T.J. & Boyce, K.E., 2000. A study of evacuation from large retail stores. *Fire Safety Journal*, 35(1), pp. 25-49. DOI: 10.1016/S0379-7112(00)00013-8.

Shields, T.J., Boyce, K.E. & McConnell, N., 2009. The behaviour and evacuation experiences of WTC 9/11 evacuees with self-designated mobility impairments. *Fire Safety Journal*, 44(6), pp. 881-893. DOI: 10.1016/j.firesaf.2009.04.004.

Williams-Bell, F.M., Kapralos, B., Hogue, A., Murphy, B.M. & Weckman, E.J., 2015. Using serious games and virtual simulation for training in the fire service: a review. *Fire Technology*, 51(3), pp.553-584. DOI: 10.1007/s10694-014-0398-1.

Appendix A - Workshop Participant Lists

Name	Organization	Country	Kickoff meeting	Workshop Ottawa	Workshop London
Virginia Alonso	AECOM, Madrid	Spain			Yes
Martyn Amos	Northumbria University/NEED	UK	Yes	Yes	Yes
Nicole Arbour	NRC	Canada		Yes	
Georgia Bateman	Imperial College London	UK			Yes
Noureddine Bénichou	NRC/NEED	Canada	Yes	Yes	Yes
Nicolai Bode	University of Bristol	UK			Yes
Karen Boyce	University of Ulster	UK	Yes	Yes	Yes
Natalia Cooper	NRC	Canada		Yes	
Arturo Cuesta	University of Cantabria	Spain		Yes	Yes
John Drury	University of Sussex	UK			Yes
Rita Fahy	NFPA	USA		Yes	
Steve Gwynne	Movement Strategies/NEED	UK	Yes	Yes	Yes
Helen Hinds	Newcastle City Council	UK	Yes		
Aoife Hunt	Movement Strategies	UK			Yes
François Jacquet	Ottawa Airport	Canada		Yes	
Max Kinateder	NRC/NEED	Canada	Yes	Yes	Yes
Erica Kuligowski (remote)	NIST	USA		Yes	
Ruggiero Lovreglio (remote)	Massey University	NZ		Yes	
Shayne Mintz	NFPA	Canada		Yes	Yes
Amanda Robbins	NRC	Canada		Yes	
Enrico Ronchi	Lund University	Sweden	Yes	Yes	
Brian Sheriff	Toronto Pearson Airport	Canada		Yes	
Alistair Shipman	Imperial College London	UK			Yes
Michael Spearpoint	OFR Consultants, Manchester	UK			Yes
Keith Todd	University College London	UK			Yes
Russ Thomas	NRC	Canada		Yes	
Lexi Thompson	NRC	Canada		Yes	
Pete Thompson	Autodesk	UK			Yes
Sean Tracey	Ottawa Firefighters	Canada		Yes	
Natalie van der Wal	University of Leeds	UK	Yes	Yes	Yes

Number of participants: total: 30; Kickoff meeting: 8; Workshop Ottawa: 20; Workshop London 17.

Number of organizations: 20.

Number of Countries: 5.

Appendix B - Excerpt of egress drill paper

Here we present an excerpt from Gwynne, S., et al., Enhancing egress drills: Preparation and assessment of evacuee performance. *Fire & Materials*, 2017. Note that internal references may not be consistent, as this is an out-of-context "snippet" of a larger paper.

A number of alternative egress models exist. These do not have the same capabilities or make the same assumptions. In the context of familiarising participants or assessing performance, there are several basic factors that can be used to differentiate the models and their potential contribution: *environment* (the physical space within which the event is conducted), *agency* (the degree of actor activity and influence during the event), *scenario* (the event narrative and path-dependency of the event), and *feedback* (the insights provided to the actors by the event manager [model user]). Each model type is described in more detail below.

The following models are included here as they might possibly assist with the assessment or familiarisation of evacuee performance, and may therefore reasonably complement egress drills in some form. A brief overview of the strengths and limitations of several of these models for egress training or assessment is presented in Table 2 [70]. This includes the following models:

1. ***Egress Simulation Tool*** - Computer simulation of an evacuation [40]. Initial scenario conditions are provided to the tool by the user that then simulates the progression of these conditions over time, depending on the response of a simulated evacuating population. With most tools, the user is able to passively monitor the conditions via a GUI as they evolve and the consequences of agent actions on the outcomes produced. Typically, data is generated on various aspects of the simulation reflecting the evolving conditions and the final outcome. As with the other computational approaches discussed, may have access to scenarios that would ethically impossible to conduct in reality; i.e., that place simulated evacuees in jeopardy.

2. ***Computational Virtual/Immersive Environment/Serious Gaming*** - Participants are 'located' within a virtual space allowing them to actively make decisions in response to the conditions faced, interact with it and perform actions (either through control of an avatar or through acting directly) [71-74]. The scope and nature of this interaction is typically limited by the nature and sophistication of the technology available. There is an attempt to represent the conditions that might be faced in the scenario of interest (e.g., in a real event). The realism of the experience is a factor in the credibility of the participant decisions taken and the measure of the task effectiveness. In recent years, the concept of serious gaming has been employed within such environments – to more effectively represent the decision-making process within such an environment. Serious games allow decisions to be made under credible, controlled and varied conditions, without exposing players to any physical risks, allowing more flexibility in scenario design [75-77]. Serious games have already been used for training purposes in the fire service [78]. Serious game players should be able to acquire some knowledge about emergency scenarios and improve their evacuation preparedness through the game experience,

and / or assess the implications of such behaviour (e.g., as part of an investigation, see [77, 79-83]). However, the impact of the realism of the environment and the long term effects of training using serious gaming has to be established [80].

3. ***Laboratory Experiment*** - Exposure of participants to a particular physical / psychological condition (ideally reflective of an evacuation) in a controlled artificial environment to assess performance of a specific task given the manipulation of this environmental condition [22]. This model represents a largely path-dependent (i.e., strictly controlled, pre-determined) decision-making process within an acknowledged artificial environment.

4. ***Table Top Exercise*** – Involves participants in a simulation of the decision-making process in response to the exercise scenario in order to test the effectiveness of the procedure tested and their part in it [70]. Should emergency responders be involved, then the table-top environment may be configured to reflect a command and control centre reflecting the physical setting under which these decisions might be made in a real incident; however, this would not be true of an evacuating population, were they involved in the exercise.

5. ***Mental Rehearsal*** – Individual attempts to visualize expected decisions, tasks and desired outcomes before the situation is experienced [84-86]. No attempt is made to create representative environmental conditions. The intention, here, is to enhance performance once the individual faces equivalent conditions. This approach has been used quite extensively in the context of training for physical activities, particularly in the area of sports coaching although it has been shown to be effective in other contexts. Prompting the individual to initiate a mental rehearsal of an envisioned event has been shown to improve both success and speed of execution. This approach, which creates minimal risk to the participant, can be undertaken at any time by the individual and can represent a very low level of investment in training other than the individual participant's time.

6. ***Hot/Cold Debriefing*** - – A review of the events, decisions and outcomes produced during the event, including those involved and those monitoring it. Ministry of Civil Defence Emergency Management, New Zealand (MCDEM) recommends that two debriefs be conducted after an exercise / drill [41]. A hot debrief allows participants to provide feedback on the recent event – to recall key events while they are fresh in the memory. The more formal cold debrief is held a month or so later allowing more reflection on the strengths and limitations of the event (after more detailed analysis of observations made) and then changes that need to be made to the procedure, tasks or the drill itself[41].

7. ***Information Session / Workshop / Seminar*** – Presentations / discussions held in order to engage the target population (individually, in groups or the entire population) in order to inform them of the scenarios that might occur, the conditions that might evolve [2,4], the procedures in place, their role within them, and / or desirable outcomes.

No attempt is made to recreate real-world actions or require the participants to simulate expected performance.

8. ***Walkaround –*** Participants (e.g., evacuees) are guided around to the key locations, routes, decision-points, and tasks by an experienced trainer/member of staff to familiarize them with the procedure and the activities that form it [2]. This walkaround will take place in the real-world environment, and will involve the target population experiencing these spaces and situations passively – without having to take decision.

9. ***Control Point Exercise*** – Staff who have a role in an emergency procedure are positioned in locations in which they would be expected to operate during an incident [4]. Aspects of communication between these locations, individual familiarity with each of the spaces, and the individual's ability to operate at these locations can be examined.

10. ***Controlled/Scripted Exercise*** – Participants are taken through an exercise following an entirely scripted narrative. Essentially, participants are acting out their roles in one controlled instance – in order to identify example outcomes and stated decisions made [4]. This is unlikely to be in a real-world setting given the required 'actor' interaction, but instead located in a single space where individual actors can more naturally communicate.

11. ***Free Play Exercise*** - In contrast to controlled or scripted exercises, actors are encouraged to role-play after being provided with a set of initial conditions [4]. This expands on the table top exercise to possibly include an evacuating population. Participants actively engage in an attempt to reach a stated goal given the initial conditions provided. This event is likely located in a single space where individual actors can communicate.

In reality, a combination of these approaches (either in partial or complete form) may be adopted to complement an egress drill, given they present a range of different strengths and weaknesses. It should be noted that the qualification of the impact / effectiveness of the models described in Table 2 is speculative and subjective at this stage. It would require additional research prior to these assessments being used in earnest. This material is presented to outline the differences that might exist and how these may be used to differentiate between the approaches available.

Table 2. Model limitations and capabilities.

	Considerations (derived from discussion presented in Section 6)							
Egress Model	Financial / Organiz.	Ethical	Methodological			3rd Party Scrutiny	Statistical	Pedagog.
			Perceived Credibility	Scope	Potential Insights			
Egress Drill	—	—	+	—	≈	≈	—	+
Simulation Tool	≈	+	—	+	+	—	+	—
VR / Immersive Expt. / Serious Game	≈	≈	—	-	+	—	≈	+
Lab Experiment	—	≈	+	≈	+	+	≈	≈
Table Top	≈	+	≈	+	≈	≈	≈	≈
Mental Rehearsal	+	+	—	+	≈	—	≈	≈
Information Session/ Walkaround	≈	+	≈	≈	≈	≈	—	+
Briefing	≈	+	≈	+	—	≈	—	≈
Control / Free / Scripted Exercise	≈	≈	+	≈	≈	≈	—	+

An indication of the other models that might be employed to assess/train the target population is shown in Table 2 (an extension on the work previously presented by [70]). The column headings employed in Table 2 reflect the key considerations identified in Section 6. Table 2 enables a comparison between the relative merits and limitations of these models and what they might contribute to population training and performance assessment. The limitations of an egress drill regarding the considerations outlined in Table 2 were described earlier in Section 6. This table extends that description to include the relative strengths and limitations of the 11 alternative models available. Cells with a '–' indicate a relative weakness in the model; a '+' symbol indicates a relative strength, and '≈' indicates that the model performs neutrally in relation to the other models available. This represents a crude and qualitative assessment of these models. However, these considerations are primarily intended to demonstrate that there are alternatives that have various capabilities and that these capabilities must be understood when assessing their relevance to the project at hand. For instance, a simulation tool causes little disruption to the organization with only a moderate possible financial cost (leading to the '≈' designation), it poses no risk or ethical concerns ('+'), and can repeatedly be employed to explore a range of scenarios ('+'); however, it is often perceived as lacking credibility given that it is a relatively new tool ('–'), and has limited educational capacity, in and of itself ('–'). Virtual reality (VR) environments can elicit similarly sceptical views given their recent development and fashionable status ('–'), although they also have enhanced training capabilities ('+'). Table top exercises provide a flexible approach, allowing a range of scenarios to be explored ('+'); however, they provide moderate insights into the conditions produced – focusing on the decision-making process rather than the specific (physical) impact of these decisions within the evacuated space ('≈'). Walkthroughs, briefings and exercises pose little ethical concerns regarding participants ('+'); however, they are unlikely to be sufficiently repeated to produce statistically credible insights ('–'). Finally, mental rehearsals may produce reasonable benefits to individual performance ('≈'); however, it is a relatively recent development again provoking scepticism not aided by the lack of external scrutiny available ('–').

It is apparent that these models have various strengths and weaknesses, and may also operate at different levels – informing or assessing the individual, group or population and relating to specific tasks or the entire procedure. As such, potential solutions may involve the use of several of these models in order to support the assessment and effective training of the target population [87] - to produce a more robust, informative and representative training and assessment program.

An example may help demonstrate how these approaches might be employed and the potential benefits. Let us assume that we are developing the emergency plan for an office building with 5000 occupants. The regulatory system with jurisdiction requires that this structure has quarterly full-building evacuation drills. A sub-set of the egress models outlined in Table 2 are combined to produce an alternative programme (labelled Approach A) to assess performance and enhance occupant / staff familiarity with the emergency procedure in place. An outline of such a program is shown in Table 3, formed from ten procedural steps. Obviously, this is one of many approaches that might be adopted.

Table 3 includes several elements for each of the ten suggested steps: the egress model employed; the frequency that this model is to be employed; the objective of this employment, in terms of the target population and the intended impact; the possible disruption that this step will cause to building operation; and the overall cost of these steps in terms of the number of person hours lost per year, calculated from the size of the population involved over the year, the time taken for the step to be completed, and the number of times it is performed per year. As noted earlier, this cost could be balanced by a more quantitative estimate of the benefit of each step, if we had a clearer understanding of the impact of these tools on staff / evacuee performance (e.g., the incremental increase in evacuee performance for each drill performed). In the absence of this knowledge, only the objective for each of the steps is stated, allowing a more qualitative assessment of what might potentially be achieved by each step.

Let us examine Step 1 in Table 3 as an example. The Information Session / Walkaround is to be conducted once per year for each occupant. This will not involve all occupants at the same time, but will be conducted individually or in small groups to reduce disruption, eventually addressing the entire population. Occupants walk through the routes that might be adopted during an emergency and then attend a seminar on the emergency procedure – the objectives, tasks, their role, resources, assumptions made, etc. As mentioned, given that the step will be applied to sub-sections of the population at any one time, the building will still be able to function as normal. Given that the entire population will be engaged over the course of the year and that the session is expected to last an hour, 5000 person.hours / year will be effectively 'lost'. The ongoing nature of this step (and others) also has the benefit of ensuring that emergency planning is seen to be a process rather than a one-off task [87]. Similar details on each of the 10 steps in the example training and assessment program is provided in Table 3.

This type of analysis for each of the steps in training and assessment program (e.g., all 10 steps in Table 3) allows us to determine the number of times the building functionality will be disrupted, the loss of time experienced by the population, the training and assessment tasks addressed and the target groups/level of these tasks (individual, procedural, etc.). From Table 3, the overall 'cost' of this approach is calculated as 35500 person.hours/year.

Table 3: Training and assessment program using multiple models.

Step	Egress Model	Frequency	Objective	Disruption	Cost (#people x #hours x #repeats)
[1]	Information Session / Walkaround	Each occupant does this every year.	Individual Level Training: Routes and systems available.	Individual/groups removed from routine activities. Building operational	5000x1x1=5000p.hr/yr No loss of building functionality.
[2]	Small-Scale Training (Individual scripted exercise)	Each occupant does this every year.	Individual Level Training: Individual training / familiarization of the procedure and their role within a given	Individual/groups removed from routine activities. Building operational.	5000x1x1=5000p.hr/yr No loss of building functionality.

			range of different scenarios.		
[3]	Small-Scale Test: in procedure; e.g. go to nearest exit, operate device, identify alarm, etc. (Individual drill)	Each occupant does this every year.	Individual Level Assessment: Individuals / groups asked to enact role and their ability to perform expected tasks is assessed.	Individual/groups removed from routine activities. Building operational.	5000x1x1=5000p. hr/yr No loss of building functionality.
[4]	TableTop Exercise	Each member of staff active in emergency plan does this twice per year.	Assessment – Individual / Procedure: Staff understanding and execution of procedure.	Staff removed from routine activities. Building operational.	50x2x2= 200p.hr/yr No loss of building functionality.
[5]	Seminar / Review	Each member of staff active in emergency plan does this twice per year.	Assessment / Familiarization: Review of table top performance	Staff removed from routine activities. Building operational.	50x2x2= 200p.hr/yr No loss of building functionality.
[6]	Full-scale Egress Drill	Entire staff / occupant does this yearly.	Assessment – Procedure: Effectiveness to meet safety objectives given scenario examined.	Staff and occupant population removed from routine activities. Loss of Building Functionality.	5000x1x1=5000p. hr.yr
[7]	Hot Debrief	Entire staff / occupant does this yearly.	Familiarization – Individual / Procedure: Review of drill conducted in small groups.	Groups of staff / occupants removed from routine activities.	5000x1x1=5000p. hr/yr No loss of building functionality.

				Building operational.	
[8]	Egress Simulation Tool	Entire staff / occupant reviews results produced by simulation yearly.	Assessment – Individual / Procedure: Examine different scenarios to assess robustness / effectiveness of procedure in place. Feedback suggests procedural enhancements and scenarios to be drilled.	Small number of staff removed from routine activities. Building operational.	2x50x1=100p.hr/yr No loss of building functionality.
[9]	Immersive VR	Entire staff / occupant does this yearly.	Individual Training / Assessment: Staff/occupants take part in range of virtual drills to train and assess under different scenarios (informed by [Step 8])	Groups of staff removed from routine activities. Building operational.	5000x1x1=5000p.hr/yr No loss of building functionality.
[10]	Cold Debrief	Entire staff / occupant does this yearly.	Training - Procedural: Assessment of performance in actual and virtual environments.	Staff removed from routine activities. Building operational.	5000x1x1=5000p.hr/yr No loss of building functionality.

If the building was required to be evacuated four times per year (Approach B), as is currently the case in some instances, then the equivalent 'cost' for the same building would be 40,000p.hr/yr (5000 people involved for two hours four times per year), with a loss of building functionality four times per year (each of the four times that it is performed). In addition, if we assume that the drills were conducted in accordance with the regulatory frameworks highlighted earlier, the assessment would typically be made at the procedural level (with fewer requirements to assess individual performance during each of the trials) and fewer individual-level training elements to enhance this performance. In this example, Approach A is less costly in terms of person-hours per year lost during the emergency planning events, it has fewer interruptions to building operations and enables emergency planners to assess / enhance performance under a larger number of incident scenarios and at a greater number of levels (individual, procedural, etc.).

This is, of course, a simplification. In addition, there are a number of other approaches (i.e. various combinations of the models highlighted) that might provide sufficient training and assessment whilst making more significant person/hour savings over multiple full-Approach B and reducing building functionality loss. However, it is hoped that this example illustrates the potential of employing complementary approaches and of quantifying the relative costs of doing so.