Safety Science

Where Drills Differ from Evacuations: A Case Study on Canadian Buildings

--Manuscript Draft--

Manuscript Number:	SAFETY-D-20-01950R1
Article Type:	Case study
Keywords:	Archival study; drill; fire evacuation; egress
Corresponding Author:	Max Kinateder, PhD National Research Council Canada ON CANADA
First Author:	Max Kinateder, PhD
Order of Authors:	Max Kinateder, PhD
	Chunyun Ma, PhD
	Steven Gwynne, PhD
	Martyn Amos, PhD
	Noureddine Bénichou, PhD
Manuscript Region of Origin:	North America
Abstract:	Planned egress drills are required by building codes around the world, and are commonly used to both train occupants and assess evacuation procedures. However, capturing the idea of a "successful" drill is often difficult. Data from both drills and unplanned evacuations are often incomplete and unreliable, which raises a key question: How well-matched are planned egress drills and unplanned evacuations in terms of their properties and outcomes? That is, are drills a good model of evacuation? In this paper, we compare 93 planned egress drills and 23 unplanned evacuations, which occurred in Canadian office buildings over a span of four years. Our two main findings are that (1) planned egress drills differ from unplanned evacuations in terms of frequency, timing, and outcome (e.g., reported total evacuation time), and (2) the reported number of occupants correlates with total evacuation time. These findings motivate a discussion of the strengths and weaknesses of the current approach to data reporting, and we highlight potential implications for (and limitations of) the current drill model.
Response to Reviewers:	

Abstract

Planned egress drills are required by building codes around the world, and are commonly used to both train occupants and assess evacuation procedures. However, capturing the idea of a "successful" drill is often difficult. Data from both drills and unplanned evacuations are often incomplete and unreliable, which raises a key question: How well-matched are planned egress drills and unplanned evacuations in terms of their properties and outcomes? That is, are drills a good model of evacuation? In this paper, we compare 93 planned egress drills and 23 unplanned evacuations, which occurred in Canadian office buildings over a span of four years. Our two main findings are that (1) planned egress drills differ from unplanned evacuations in terms of frequency, timing, and outcome (e.g., reported total evacuation time), and (2) the reported number of occupants correlates with total evacuation time. These findings motivate a discussion of the strengths and weaknesses of the current approach to data reporting, and we highlight potential implications for (and limitations of) the current drill model.

Keywords: Archival study, drill, fire evacuation, egress

Title Page (with Author Details)

Where Drills Differ from Evacuations: A Case Study on Canadian Buildings

Kinateder, M.[a*+], Chunyun Ma[a*], Gwynne, S.[b,c], Amos, M.[d], & Bénichou, N.[a]

[a] National Research Council Canada, 1200 Montreal Road, Ottawa ON, K1A0R6, Canada

[b] Movement Strategies, 31-35 Kirby Street, Farringdon London EC1N 8TE, United Kingdom

[c] Lund University, P.O. Box 118, 22100 Lund (Sweden)

[d] Northumbria University, Ellison Place, Newcastle upon Tyne, NE1 8ST, United Kingdom

[*] Equal contribution

[+] Corresponding author; email: max.kinateder@nrc-cnrc.gc.ca

1. Introduction

Planned egress drills[1] (hereafter referred to as "drills") are used to train individuals in emergency evacuation procedures, and to assess the performance of both protocols and evacuees [1]. Training is necessary, since the responses of both active (e.g., assisting) staff and evacuating occupants are fundamental to the outcome of an evacuation in an emergency incident. Assessment is equally important, as we must evaluate whether or not performance targets are met. Drills are the most common model of evacuation and are required by many building codes around the world [1]. They are often seen as a reliable means of simulating evacuation scenarios in order to meet various objectives. However, addressing the question of what constitutes a *successful* drill is not straightforward.

Data collection from drills is often inconsistent and/or partial. Data collection during *unplanned emergency evacuations* (hereafter referred to as "evacuations") is generally even more limited, and this is often due to a lack of systematic and objective data collection methods [2]. This information gap limits both our understanding of drills and evacuations, and our ability to compare them. More fundamentally, it limits our ability to assess *whether or not drills constitute a valid model of evacuation*. If we rely on drills to predict occupants' performance in evacuations, or have them serve as a proxy for evacuations within the regulatory process, then this question must be addressed.

In this paper, we share an array of new data from drills and evacuations that may be used by safety engineers and researchers to characterize evacuee performance. We present an analysis of a dataset derived from reports of 116 events (both drills and evacuations) involving 49 Canadian office buildings over a four-year time-span. We begin by defining the notion of a "drill", which informs subsequent discussion of their key objectives (and how these can be met). These objectives naturally suggest a number of parameters that may be used to assess both drills and evacuations. We then describe the results of our analysis, which examines 93 drill reports and 23 evacuation reports, and interrogates them in terms of outcomes and conditions. We find that evacuations differ from drills in two key aspects: (1) the total evacuation time, and (2) the time of day at which they occur. We conclude with a discussion of both our methodology and our findings, offer some suggestions as to how the reporting of drills and evacuations could be improved, and suggest some future lines of research in this area.

1.1 Drill Definitions

In this section, we summarize a number of existing definitions of drills, in order to both contextualize what follows and extract a number of core objectives. The National Fire Protection Association (NFPA)

[1] Egress and evacuation are used interchangeably throughout.

offers two definitions in its standards: NFPA 101 (Life Safety Code) states that the purpose of emergency egress and relocation drills is

> "To educate the participants in the fire safety features of the building, the egress facilities available, and the procedures to be followed. Speed in emptying buildings or relocating occupants, while desirable, is not the only objective. Prior to an evaluation of the performance of an emergency egress and relocation drill, an opportunity of instruction and practice should be provided." [3]

The code elaborates on specific cases concerning different occupancy (e.g., hospitals, office buildings, commercial buildings, etc.), the importance of unannounced drills, and the risks of insufficient implementation of a drill. NFPA 101 further states that drills are to be held "at sufficient frequency" and

> "At both expected and unexpected times and under varying conditions. As fire is always unexpected, if the drill is always held in the same way at the same time, it will lose its value and effectiveness." [3]

NFPA 600 (Standard on Facility Fire Brigades) defines a drill as

> "An exercise involving a credible simulated emergency that requires personnel to perform emergency response operations for the purpose of evaluating the effectiveness of the training and education programs and the competence of personnel in performing required response duties and functions." [4]

Although it focusses on emergency response, this definition still emphasizes that the simulated emergency should be credible and that performance should be evaluated. The NFPA high-rise emergency plan guide adopts this definition, but also expands on the purpose of fire drills with regards to evaluation:

> "Tests should be conducted to evaluate the preparedness and capabilities of building occupants and life safety staff (e.g., via fire drills)." [5]

This explicitly differentiates the roles of staff and occupants.

The Canadian National Fire Code (NFC) does not define a drill *per se*, but alludes to the idea that drills encompass a range of activities:

> "A fire drill … is at least a review of the fire safety plan by supervisory staff." [6]

This provides a minimum threshold for determining the adequacy of the plan in place.

Finally, Gwynne, *et al*. define an egress drill as:

> "A preplanned simulation of an emergency evacuation for a specific incident scenario … used to improve the performance of the occupant population and/or staff present and active during an emergency." [1]

This definition reflects regulatory frameworks from around the world, including those in the UK, Italy, Spain, US and Canada [1]. Additionally, the authors suggest that any improvement of evacuee performance may either be *direct* (through acquired learning during the drill) or *indirect* (through assessment and subsequent instruction).

1.2 Drill Objectives

If we consider these definitions together as a whole, we may extract the following common core objectives of drills [2]:

1. *Performance assessment* of evacuees and the evacuation system.
2. *Training* of staff and occupants in evacuation procedures.
3. *Fulfillment* of regulatory requirements (e.g., by documenting the drill outcome).

These are all clearly important objectives. However, we cannot assume that they are always achieved [7], and sometimes objectives are co-dependent (this is particularly true for Objectives 1 (Assessment) and 2 (Training)). Fundamentally, we need to ask how these objectives may be used to define a "successful" drill. Is a successful drill one in which all participants meet the training goals, or one that identifies weaknesses in the evacuation procedures?

If we are to use drills as the basis for judging a building's preparedness for emergency evacuations (Objective 1), then an exercise should have elements that are reflective of a real emergency (albeit within limits of available resources), while considering practical and ethical issues. To this end, a drill should be conducted under a range of representative circumstances and evacuation scenarios, both ideal (i.e., best case conditions) and realistic (i.e., requiring the evacuees to adapt their response in some form) [2]. In reality, it may be convenient to run a drill on a slow business day, or when the weather is favourable. However, a real emergency could happen at any moment.

In order to fully meet Objective 1, drills should provide measures of evacuation performance, on both an *individual* level (e.g., did a specific occupant follow the evacuation procedure correctly?) and a *system* level (e.g., were there any bottlenecks on an egress route?). Metrics against which a drill could be measured include:

- *Total evacuation time:* What is the overall egress time for different buildings / populations / evacuation scenarios? Do egress times differ between drills and evacuations?

- *Scheduling:* Are drills conducted at representative times of the day (morning, evening, off-hours, etc.) and seasons of the year (e.g., summer vs. winter)?
- *Building/facility preparedness*: Are drills being held under idealised vs. realistic conditions? For example, do drills simulate potential evacuation scenarios (e.g., loss of egress routes)?
- *Occupancy load:* How does the number of occupants relate to egress performance?

In order to meet Objective 2, drills should train building occupants and staff in evacuation procedures (e.g., embed knowledge of egress routes and muster points). For any training to be considered successful, rehearsed procedures should be executed correctly during an exercise. The training effectiveness of drills may be measured using the following metrics:

- *Performance:* Do trainees meet specific requirements (e.g., evacuate in time and in an orderly manner)? Does overall performance improve after training (e.g., improved evacuation times)? How does performance compare between drills and evacuations (e.g., in terms of evacuation time)?
- *Documentation:* Has the drill been documented so that critical information may be recorded and acted upon (e.g., highlighting potential pain points in the evacuation procedures)? This metric is critical, as it forms the basis of the assessment process.

In order to meet Objective 3, drills need to fulfill the regulatory requirements of their jurisdiction. Codes typically require that buildings conduct drills periodically (e.g., in some jurisdictions, this is at least once every 12 months), depending on the occupancy type [6]. Typically, codes only provide general guidance on drill implementation, and leave it to the discretion of building management to decide how a specific drill should be conducted. In addition, the variation in requirements for drills highlighted above shows the potential for inconsistencies to arise (e.g., in how drills are implemented and documented) [2], [8]. At the very least, records should be kept of each drill in order to verify and document that regulatory requirements were met [3].

The objectives/criteria listed above are certainly not exhaustive. We select them because they are the most commonly-used, and are therefore deemed (by a broad community) to be useful and necessary benchmarks against which drills may be evaluated [9].

1.3 Motivation for the study

Many studies have documented various important aspects of drills (e.g., [7], [10]–[14]) and evacuations (e.g., [15]–[18]). Most of these rely on relatively complex and sophisticated methods of recording data (e.g., through interviews or video analysis) in an attempt to meet *research* objectives, as opposed to meeting more routine needs. However, in practice, most drills and evacuations are evaluated using less complicated means (e.g., evacuation reports completed by those conducting the drill/evacuation). These generally require fewer resources, but may come at the cost of losing rigor. The goal of the work we present here is to establish what can be learned from such reports of routine evacuations, rather than from more refined research-oriented observations. To this end, we describe a case study comprising a sample of standard handwritten reports on both drills and evacuations. Each report contains information about the outcome (total evacuation time) and conditions (e.g., scheduling, number of occupants in the building, use of evacuation scenarios) of the drill/evacuation. Informed by the three common objectives mentioned above, we pose a series of questions of the data, in order to facilitate comparisons between reported evacuations and drills:

- *Total evacuation time*: Are there overall differences in reported total evacuation time *between* drills and evacuations? How does the total evacuation time relate to the *number of occupants* in a building?
- *Scheduling of drills*: When do drills and evacuations take place in terms of the time of day and season of the year?
- *Evacuation scenarios*: How often do drills employ evacuation scenarios in order to simulate unplanned incidents?
- *Frequency of drills*: How often are drills conducted?

In the following sections, we describe the dataset in detail and then present the results of our analysis, which is framed in terms of these questions. In each subsection, we explain the motivation behind the corresponding question, present our analysis, and use it to assess how well drills meet the main objectives listed earlier.

2. Analysis and results

2.1 Data description

We analyze 116 reports (93 drill reports and 23 evacuation reports) obtained from data gathered on 49 office buildings over a period of four years (2016 – 2019), as shown in Table 1 (see also Appendix A for the de-identified raw dataset). The buildings are located across Canada, and the majority of these buildings are multi-story office buildings, with reported occupancies ranging from 3 to 417 occupants at the time of reporting (mean = 135 occupants, median = 100 occupants). All drills and evacuations were followed up with a debriefing (usually on the same day, or (at most) a few days after the event; the content and form of the debriefing is not known). The total number of occupants across all buildings is

15,387 (12,344 for drills and 3,043 for evacuations). In all instances, drills and evacuations were facilitated and documented by designated building staff who were specifically required to record the outcome.

Because of their routine nature, there naturally exist more reports on drills than on unplanned evacuations in each of the four years. In the case of drills, the evacuation is typically initiated by a staff member manually triggering the fire alarm. However, in the cases of the unplanned evacuations, it is not always clear from the reports what exactly triggered the fire alarm. The most commonly stated reason, however, is a false alarm (e.g., a fire alarm triggered during maintenance work).

All reports use the same one-page hard-copy form and were completed by the supervisory staff responsible for conducting the drill/evacuation at a specific building. The data is anonymized, so that neither the building nor those individuals involved may be identified. We extract the following information from the reports:

- Total evacuation time
- Time and date of the drills/evacuation
- Time and date of the report
- Method by which the fire alarm was triggered
- Number of occupants involved

Table 1 Number of reports for drills and evacuations over the analyzed time period.

Year	**Drill**	**Evacuation**	**Total**
2016	16	3	18
2017	32	11	43
2018	26	9	35
2019	19	0	19
TOTAL	93	23	116

The report template allows for additional observations (e.g., descriptions of an evacuation scenario) and some subjective assessments. We do not subject these comments to a formal analysis, but we consult them in order to provide additional context.

We acknowledge some immediate limitations in the data. First, the data were donated to the authors, and we do not know whether it represents all drills/evacuations that occurred during the period under study. In addition, the authors were not present during the events or during the collection of the data, and were therefore unable to either influence or observe this first-hand. Secondly, the authors had no control over the design of the report template and the method of collection of the raw data (e.g., how the number of occupants in a building was assessed or how evacuation times were measured).

We now consider each of the questions posed of the dataset.

2.2 Total evacuation time

We define total evacuation time as the elapsed clock time between the triggering of the alarm and the completion of the drill/evacuation (i.e., all occupants have been accounted for). Despite its limitations as a descriptor of evacuation dynamics [19]–[21], total evacuation time is an important benchmark of performance for both assessment and training purposes [22]. In addition, this metric may also allow practitioners/safety managers to estimate the Required Safe Egress Time (RSET, [19]) for a given building/scenario, and to design interventions if necessary.

A Wilcoxon rank-sum test shows that *total evacuation time was faster in drills than in evacuations*, $W = 526$, $p < .001$ (see Figure 1A and Table 2; note that the requirements for parametric testing were not met since the data were not normally distributed, as shown by a Shapiro-Wilk test, $W = 0.84$, $p < .001$, and the variances of the two samples were not homogeneous, as indicated by Bartlett's test, $k^2 = 18.055$, $p < .001$; also see marginal histograms in Figure 1B). On average, drills are reported to be 124 seconds faster than evacuations.

Table 2. Summary descriptive statistics for total evacuation time in seconds (also see Figure 1A)

	Mean	Median	Standard Deviation	Minimum	Maximum
Drills	268	248	123	76	660
Evacuations	393	300	250	120	1020

Using multiple linear regression, we test the model in which evacuation time is a function of occupancy size for both types of egress (drill vs. unplanned evacuations). Again, regardless of the occupancy size, drills are faster than unplanned evacuations by approximately two minutes on average ($\beta = 129.30$, $p = .021$). Perhaps as expected, egress tends to last longer as the number of occupants increases (see Figure 1B). On average, for every additional 100 occupants present, the total evacuation time increases by one minute ($\beta = 0.61$, $p < .001$). We see that this is true for both drills and unplanned evacuations.

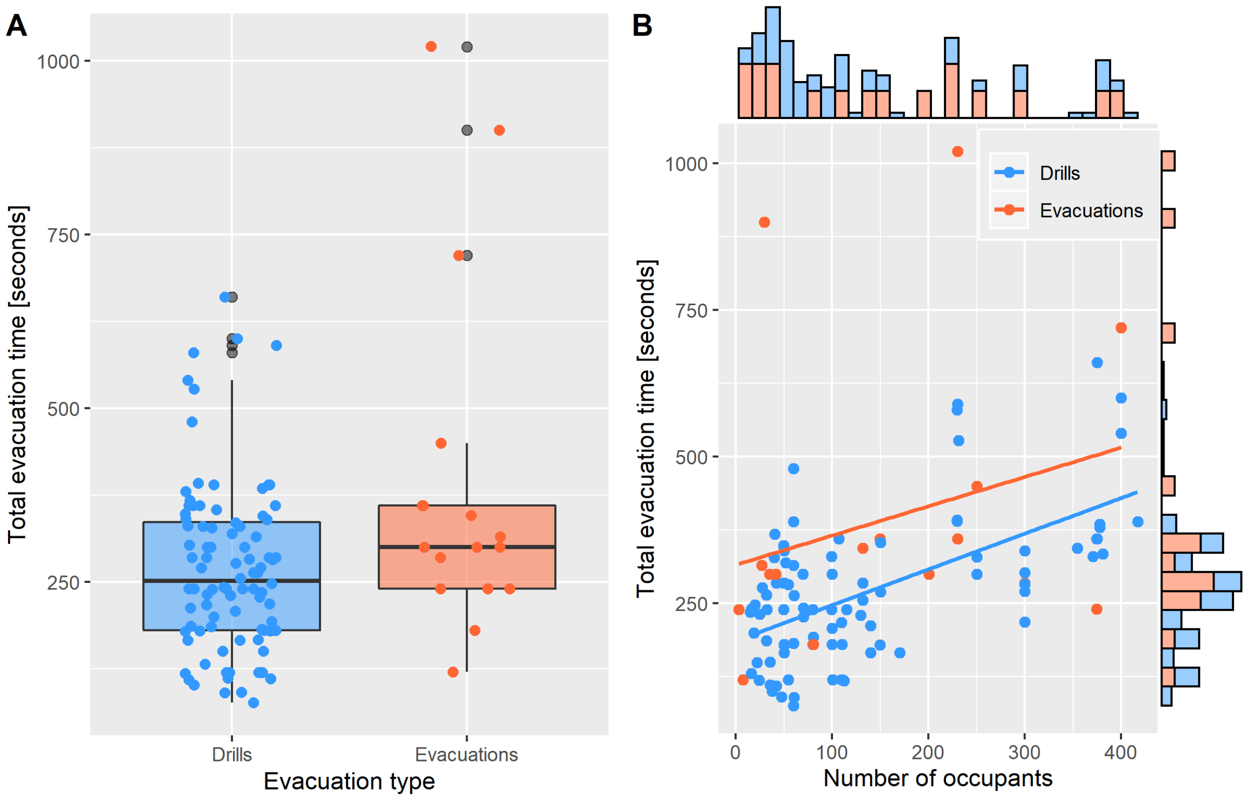

Figure 1. A: Total evacuation time in drills vs. evacuations. B: Relationship between evacuation time and number of occupants in drills vs. evacuations.

We observe three outliers in the evacuation group, with total evacuation times of at least 720 seconds (see Figure 1A). Further inspection of the written reports corresponding to these three events yields no evidence of unusual behaviour. However, in one case a smaller-than-usual occupant number was reported, as the evacuation happened after 5 pm (when most people had already left the building). In fact, this same building had two additional unplanned evacuations, but evacuations were terminated before completion because the triggers were immediately identified and deactivated (the total evacuation times for these cases are not reported here). It is reasonable to speculate that the unplanned evacuations in the current dataset are under-represented from a statistical point of view, and that the identified outliers are a side-effect of this small sample size.

Total evacuation time was reported in documentation as “X minutes and Y seconds”. Analyzing the frequency distribution of the reported seconds (i.e., the Y value), reveals how accurately evacuation times were measured (this analysis is inspired by an approach used in forensic data analysis that reveals how often values repeat within a dataset in order to detect ‘number bunching’ [23]). If evacuation reports are accurate to the second, these values should be relatively uniformly distributed. Figure 2 plots the last minute of every reported total evacuation time, and shows a ranked histogram (sorted from most to least frequently observed number). For both drills and evacuations, the bulk of total evacuation times were reported as either full or half minutes. This suggests that most observers either had no access to more precise measurements, rounded their measured time, or simply estimated the egress time.

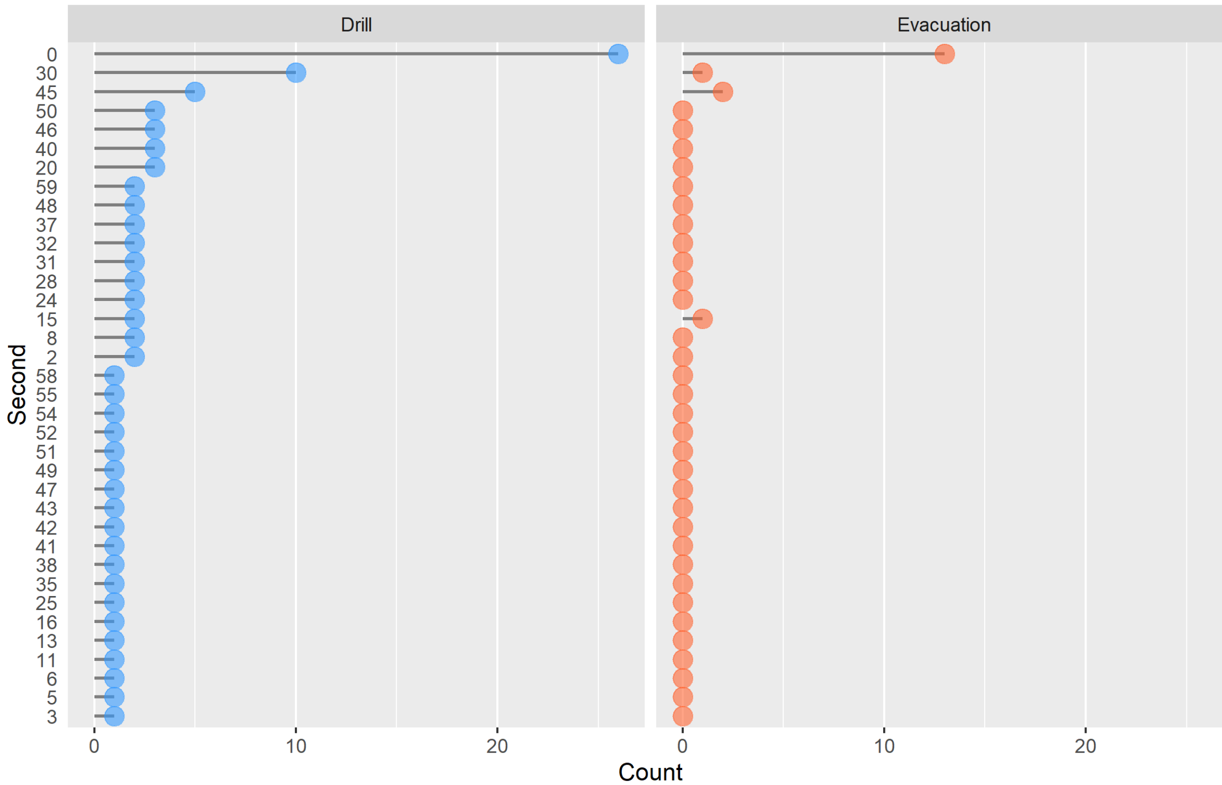

Figure 2. Seconds in the last minute reported sorted by frequency of reported values.

2.3 Scheduling of drills

The *scheduling* of drills affects how representative they are of actual evacuations [3]. It also influences the range of evacuation scenarios available to organizers of drills, the number of occupants in the building, and even their levels of alertness. We first analyze how drills are scheduled in terms of the time of day, and then in terms of seasonality (Figure 3).

We find that drills tend to be held during regular working hours, excluding lunch breaks (Figure 3A), whereas evacuations are more evenly distributed throughout the day. Fisher's Exact Test for Count Data confirms this observation ($p < .01$). Some reports on unplanned evacuations during early working hours described significant challenges; in one case, for example, not all occupants (or staff responsible for evacuation) had arrived at work when the alarm was triggered, and occupants arriving for work were unaware that an evacuation was happening and tried to enter the building.

Similarly, drills are more likely to be held in October than in any other month (Figure 3B). In fact, drills conducted in October account for 40.4% of all drills. This is in sharp contrast to evacuations, which seem to peak in frequency during January and February. Fisher's Exact Test for Count Data also confirms that the seasonality of drills is significantly different from that of unplanned evacuations ($p < .01$). The reason for the seasonality of drills is unclear, but we might speculate that these relate to the specific context of the buildings involved in the drills. Perhaps rather more speculatively, the National Fire Protection Association organizes its National Fire Prevention week (in the US and Canada) annually at the beginning of October, which may explain the cluster of planned drills during that month).

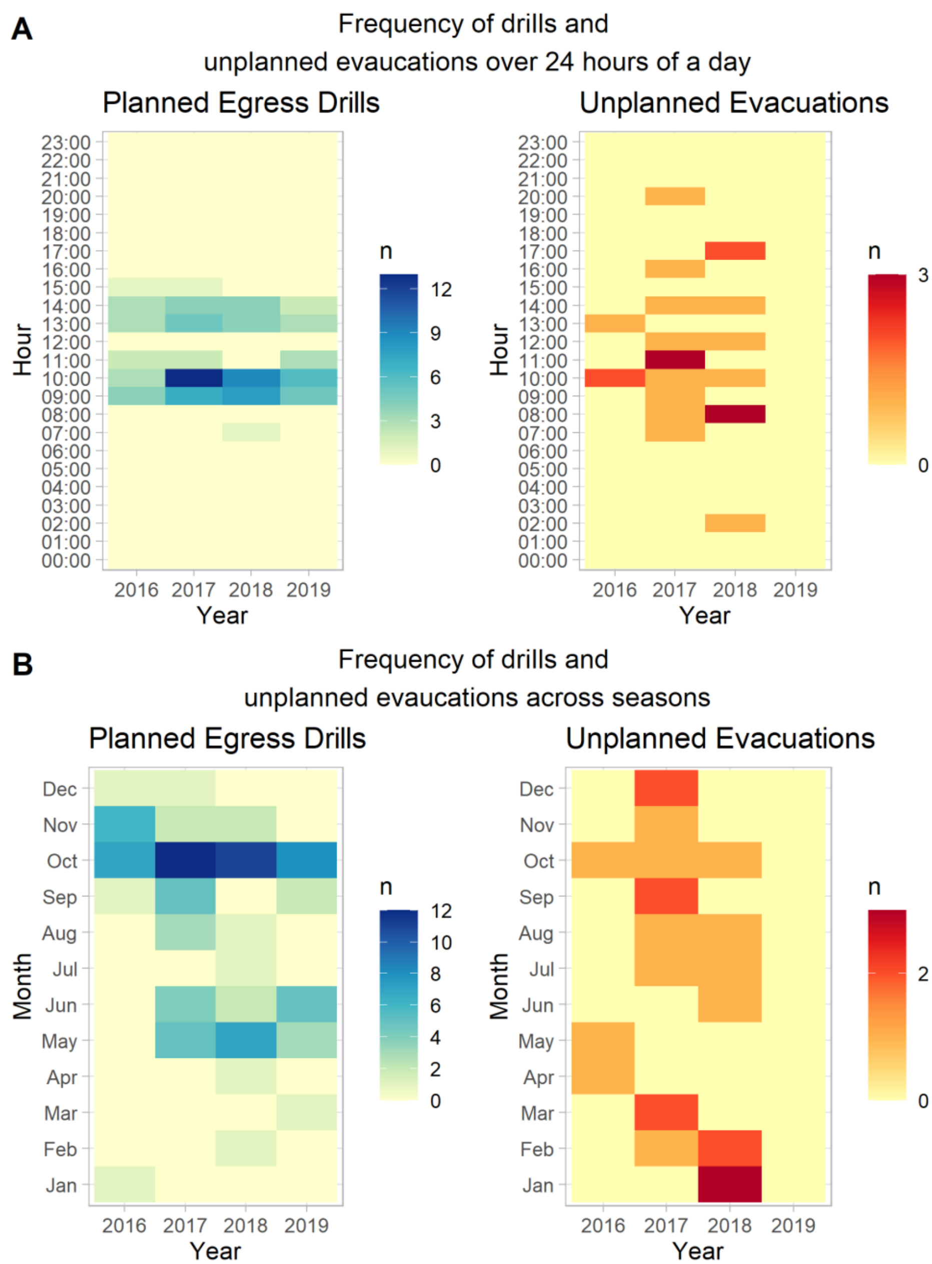

Figure 3. Frequency of drills vs. unplanned evacuations a) over 24 hours of a day (top panel), and b) across seasons (bottom panel).

2.4 Evacuation scenarios in drills

Deliberately introducing hypothetical conditions or restrictions into a drill, such as the loss of an egress route, is a technique to simulate a variety of conditions that might occur during an actual emergency evacuation. This is considered an effective strategy for training building staff and occupants, so that buildings are not only prepared for the benign situations, but also for more representative (and potentially hazardous) situations [2]. It is also an effective method to test for the existence of “pain points" in an evacuation procedure. The most frequently reported intervention in our dataset was the blocking of an egress route (typically the main entrance of a building). One report described an evacuation scenario in which one of the staff responsible for conducting the drill placed him/herself in front of an exit door and held up a sign saying "I am fire".

Ten of the 93 drills included an evacuation scenario that simulated some kind of incident (e.g., by blocking certain egress routes). Although this ratio is too small to discern any systematic difference between drills with and without evacuation scenarios in terms of their total evacuation time, we find some evidence that mean and median evacuation times were longer in drills that had implemented an evacuation scenario vs those without such a scenario (mean difference = 35 s; median difference = 34 s). Note that the 1:10 ratio could be an underestimation, because evacuation scenarios are not explicitly interrogated in the report. We are only able to identify drills with scenario-based interventions if such details are explicitly mentioned in the free text comments.

2.5 Frequency of drills

In general, most regulations require that drills be conducted on a regular or periodic basis (see e.g., [3]). However, in the present dataset we see only a single building for which drills are reported for each of the four years covered in this analysis. In turn, four buildings that had reported at least one evacuation reported no drills during the period. Figure 4 shows whether each building in the dataset conducted a drill at least once a year. Assuming that buildings in our sample had not conducted drills other than the ones represented in our dataset, then not every building in the current study met the typical regulatory requirement. However, we acknowledge that additional unreported drills may have been conducted, with documentation lost or omitted from the donated database for unknown reasons.

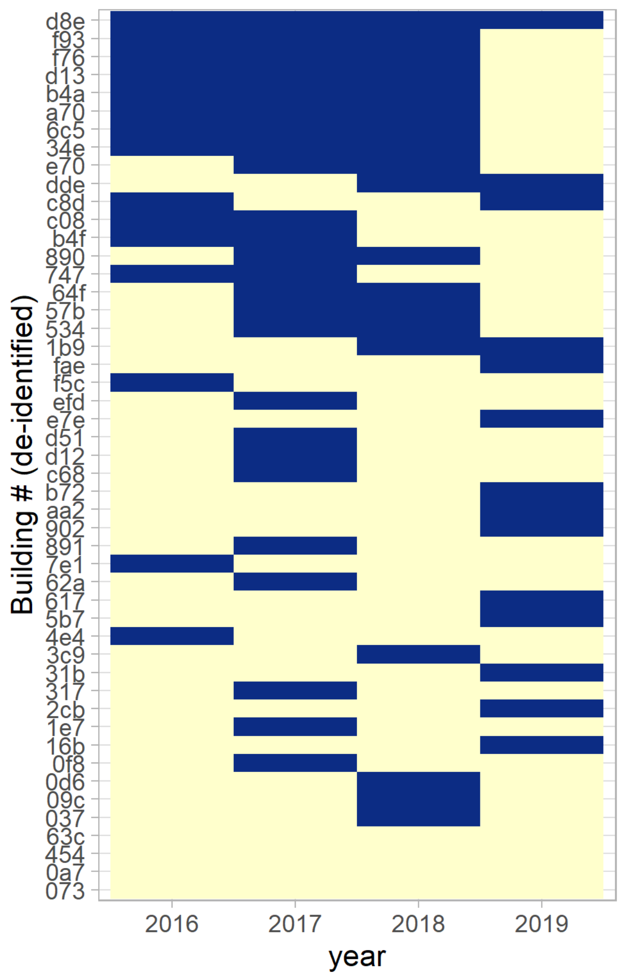

Figure 4. Whether a building reported at least one drill (blue) or none (yellow) from 2016 to 2019. Each row represents one building in the dataset. Note that the y axis shows a de-identified random label for each building.

3. Discussion

We analyzed 116 reports from drills and evacuations, and found that regular fire drills differ from evacuations in terms of total evacuation time and the time at which they occur. In this section, we (1) discuss the degree to which each of the three objectives (assessment, training, fulfilling regulatory requirements) was met in the present data, (2) consider the limitations of the present study, and (3) provide some recommendations for future work and briefly discuss potential alternatives to the *status quo* drill model. While these results have potential implications on how to improve evacuation drills, the intention here is not to question the usefulness of drills *per se*.

3.1 Performance assessment (Objective 1)

Conducting a drill makes the implicit assumption that results from the exercise may be extrapolated to a real emergency (i.e., that a drill is a valid model of an evacuation) – or at least that it is a consistent indicator of evacuation performance. In the current study, we tested this assumption by contrasting the circumstances under which drills were conducted to those under which evacuations occurred. Compared to evacuations, the drills in the current sample were scheduled under favorable conditions including daytime, agreeable weather, and mostly without simulating less than optimal scenarios. This provides a potential explanation for why total evacuation time in drills was around two minutes faster than in evacuations. Importantly, it also means that *results from drills may not provide an accurate representation of a building's evacuation performance*.

Although there are many possible explanations for the observed differences between drills and evacuations in the present data (see discussion of limitations below), these discrepancies do serve to raise the question of whether or not drills (as they are currently performed) may be used with confidence to assess and predict evacuation performance . For example, would drills conducted during a temperate fall season fully inform building managers about what to expect during evacuations in an extreme Canadian winter?

The correlation between the number of building occupants and the observed total evacuation times is not surprising. Buildings with higher occupancy loads tended to be larger and, as a consequence, occupants had to travel longer distances during egress. This finding is also consistent with other observations in the literature concerning evacuee density and evacuation speed. The more occupants there are in a building, the more likely we are to encounter high densities of pedestrians on egress routes. For instance, Pauls report a correlation between the number of occupants per square meter on a staircase and total evacuation time (assuming that the number of occupants in an evacuation is positively correlated with the area of the stairway) [19] (Figure 5).

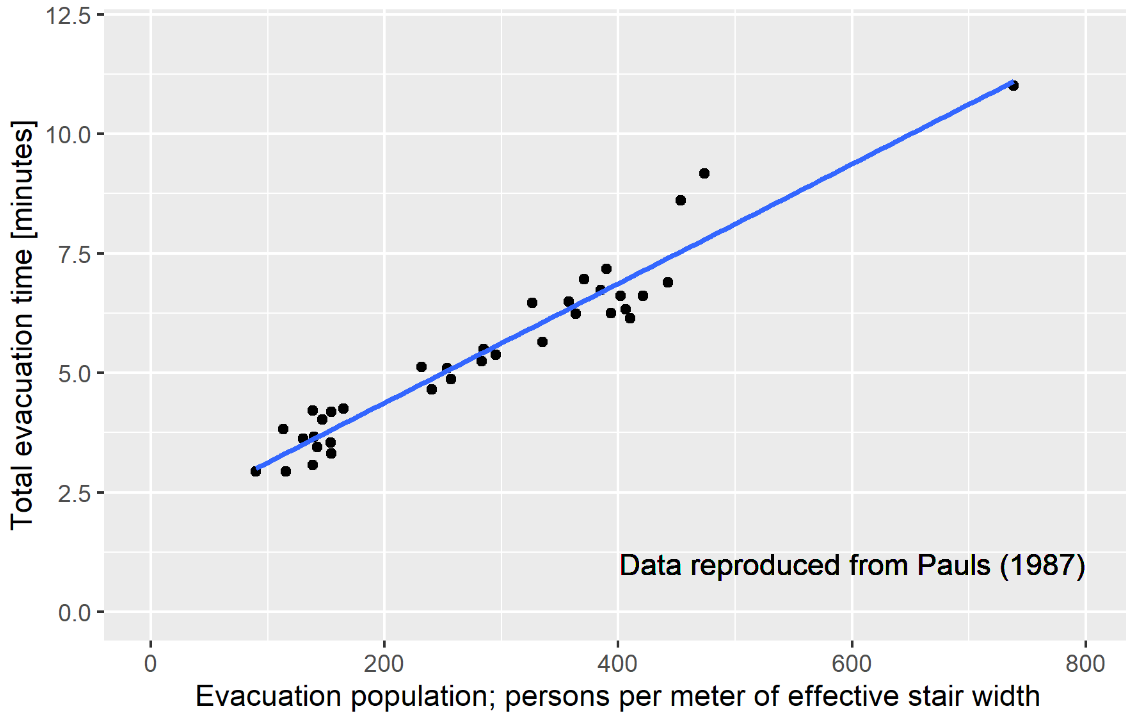

Figure 5. Data from Pauls 1987 (Figure 7) digitally reproduced for comparison [19]. Only the "took no coats" data reprinted here. Data reproduced using the R Digitize package. The blue line represents a linear model fitted to the digitized data.

A valid assessment of a drill or evacuation fundamentally relies on a reliable measurement of relevant outcome variables (e.g., total evacuation time). We observe indications in our data that total evacuation time was measured with differing levels of precision (see Figure 2). The required level of precision for the measurement of evacuation time is still not clear. However, regardless of the precision level, this measure should be reliable and consistent across reports.

3.2 Training (Objective 2)

One important factor in the success of training is the transfer of knowledge and expertise to participants, such as the process of finding the appropriate egress route(s). Restricted use of evacuation scenarios in drills, corroborated by an overall faster evacuation time in drills, indicates some limitations of training. Beyond this, the current dataset does not report any formal diagnostic data for training success, so we cannot draw any firm conclusions on this. However, some reports included descriptions of non-optimal occupant behavior (e.g., re-entry prior to completion of the drill), comparisons with prior drills (e.g., improvement of total evacuation times), and elaborated on mitigating measures. In addition, each report indicated that there was a post-evacuation debriefing, providing an opportunity for feedback and analysis.

3.3 Fulfill regulatory requirements (Objective 3)

In the current dataset, not all buildings reported the consistent performance of a drill every year (Figure 4). Assuming that this a typical regulatory requirement, we cannot categorically state that all buildings in our study fulfilled this..

3.4 Suggestions for improvement/alternatives

Our conclusions do not question the general value of drills. In fact, we identify several practical strengths of drills and their documentation; for example, the report form used for all buildings in our study – a single page template – is quick and easy to complete. A reporting format that is standardized within an organization has the additional benefit of allowing comparisons both between buildings, and across multiple exercises performed in the same building.

However, given the costs and risks associated with drills, we offer here some general suggestions on drill implementation and documentation:

- **Systematic implementation of evacuation scenarios**: Evacuation scenarios (that is, interventions to simulate specific situations that may arise during an evacuation) were only employed in a small number of drills; however, as described above, they can be relatively easy to implement, and broaden the training base of participants. If, for example, every other drill includes an evacuation scenario, a comparison between baseline performance and performance under non-optimal conditions is made possible for individual buildings. This also diversifies the conditions under which drills are performed, thus increasing participant exposure to a variety of challenges that might occur during incidents (a training benefit), and providing opportunities to test protocols and performance in response to those conditions (an assessment benefit).
- **Random scheduling** could improve the representativeness of drills. For example, drills could be scheduled at random times during working hours and across the year, and not just when it is most convenient or comfortable. . We did not see this in our dataset, which indicated a tendency for drills to be mostly performed at convenient times.
- **Standardized documentation of key drill/evacuation outcomes in an electronic database** has several immediate benefits. Data from previous drills and evacuations may be made readily available so that information may be easily extracted. This allows trends to be identified more easily (e.g., improvement or deterioration of evacuation performance over time, or differences between drills and evacuations) and lessons learned from drills across

different structures over time. Documentation may include narrative descriptions of likely explanations for egress performance and provide useful insights for emergency planners.

- **More detailed standardized reporting**. In addition to total evacuation time, other measures such as the orderliness of the evacuation, route usage or pre-evacuation behavior may be reported. This may involve relatively simple additions to documentation, such as Likert scale ratings of orderliness of the evacuation procedure, the time the first and the last evacuee reach a muster point, and reports on timing and types of pre-evacuation behavior. The latter aspect seems particularly relevant, as the literature suggests that pre-movement time dominates total evacuation time [21], [24]. Perhaps more importantly, the degree to which the emergency procedure was actually followed during the drill should be reported. This allows us to determine whether the results produced (e.g., the overall evacuation time) were due to the effectiveness of the planned emergency procedure, or to some other *ad hoc* procedure employed or chance event occurring during the drill.
- **Considerations in engineering calculations:** Required Safe Egress Time (RSET) calculations are based on data that may have been collected under drill-like conditions. Our data suggest that drills underestimate total evacuation time. If this is the case, we would propose that additional safety margins be included in RSET calculations should they rely on drill data alone.
- **Definition of specific goals for each drill**. If we require those responsible for drills to (a) decide, ahead of time, whether the primary focus of the exercise is assessment or training, and (b) document associated benchmarks of success, we may “nudge” planners to implement their drills more effectively by clarifying the objective and associated performance requirements in advance.

3.5 Limitations and future work

The main purpose of this study was to explore whether drills, as reported, meet the objectives commonly cited in the literature and required by many regulatory codes. Our results show that this is surprisingly hard to evaluate, due to a number of limitations. The most important limitation is the ambiguity / variability in reporting practices that exist in the evacuation reports themselves (for example, the precision of measurement/reporting of the total evacuation time)..Also, we do not currently have available more detailed reporting of in-depth features of exercises, such as pre-evacuation behavior and movement conditions, which would help to establish the underlying dynamics that contribute to a total evacuation time. This limits the diagnostic value of the dataset, both from a scientific and practical perspective. We believe that including such information in future evacuation reports would be extremely useful, as it could provide a more detailed picture of the evacuation process. Potential new insights that could be generated include whether or not differences in total evacuation time are due to unusual scenario conditions, extended pre-evacuation times, or flow constraints. Future work should explore how such data may be included in evacuation reports without overburdening rapporteurs with this additional task. However, given the costs and risks of drills (and the additional resources and expertise required to extract more refined data), establishing standardized reporting methods for drills and evacuations appears to be highly desirable.

Another important statistical limitation of the data centres on the fact that we observed many more drills (n = 93) than evacuations ($n = 23$). In particular, the small number of evacuations limits the representativeness of the data. Future research should obtain larger samples of drill and evacuation data.

Finally, the causes of many of the observed differences between drills and evacuations are not clear. Future research in more controlled conditions (e.g., through behavioral experiments or evacuation modelling) could shed some light on the types of discrepancy we see in our data.

4. Conclusions

We analyzed 116 evacuation reports that covered planned drills as well as unplanned evacuations, based on donated data that was extracted from a short single-page report template. Our study revealed that planned drills (as currently conducted and reported on) differ in several important aspects from unplanned evacuations (e.g., total evacuation time, scheduling). This limits the degree to which drills achieve their objectives of assessing and training building occupants' evacuation performance, and the extent to which they can reasonably act as a proxy for actual evacuation conditions. In reality, drills might be employed for training and for performance assessment – both of which rely on the drill as a proxy, but which provide different insights. The degree to which a drill is representative of actual evacuation conditions might therefore have a twin-pronged effect on our understanding of emergency preparedness and performance.

One important question remains open: What is a successful drill? Does a good drill mean a relatively quick evacuation time (and how fast is good enough?), that the procedure was followed, or that lessons were learned? In the reports discussed here, we sometimes found comments such as "good drill" provided by the observers. These reports were usually those in which the drill appeared to have gone smoothly and no unusual events were noted. In other reports, informal yet detailed observations of non-optimal occupant behaviors were described as well as the actions taken to improve the issues. It appears that much can be learned from such non-optimal cases.

References

[1] S. M. V. Gwynne *et al*., "Enhancing egress drills: Preparation and assessment of evacuee performance," *Fire Mater*., vol. 0, no. 0, 2017, doi: 10.1002/fam.2448.

[2] S. Gwynne *et al*., "The future of evacuation drills: Assessing and enhancing evacuee performance," *Saf. Sci*., vol. 129, p. 104767, Sep. 2020, doi: 10.1016/j.ssci.2020.104767.

[3] National Fire Protection Association (NFPA), *NFPA 101: Life Safety Code*. NFPA, 2018.

[4] National Fire Protection Association (NFPA), *NFPA 600: Standard on Facility Fire Brigades, High-rise emergency plan guide*. NFPA, 2020.

[5] High Rise Building Safety Advisory Committee, NFPA, "Developing Emergency Action Plans for All-Hazard Emergencies in High Rise Apartment Buildings.," White paper, 2018. Accessed: May 22, 2020. [Online]. Available: https://www.nfpa.org/-/media/Files/White-papers/WhitePaperHighRiseApartmentBuildings.pdf.

[6] National Research Council of Canada (NRCC), *National Fire Code of Canada*, 2015th ed., 2 vols. Ottawa, Ontario, 2015.

[7] L. Catovic, C. Alniemi, and E. Ronchi, "A survey on the factors affecting horizontal assisted evacuation in hospitals," *J. Phys. Conf. Ser*., vol. 1107, p. 072001, Nov. 2018, doi: 10.1088/1742-6596/1107/7/072001.

[8] S. M. V. Gwynne *et al*., "Enhancing egress drills: Preparation and assessment of evacuee performance," *Fire Mater*., vol. 0, no. 0, doi: 10.1002/fam.2448.

[9] M. Amos, N. Bénichou, S. Gwynne, and M. Kinateder, "NEED: networking activities for enhanced evacuation drills: roadmap for enhanced evacuation drills," National Research Council of Canada. Construction, Sep. 2019. doi: 10.4224/40001242.

[10] G. Proulx, "Evacuation time and movement in apartment buildings," *Fire Saf. J*., vol. 24, no. 3, pp. 229–246, Jan. 1995, doi: 10.1016/0379-7112(95)00023-M.

[11] M. Kobes, I. Helsloot, B. de Vries, J. G. Post, N. Oberijé, and K. Groenewegen, "Way finding during fire evacuation; an analysis of unannounced fire drills in a hotel at night," *Build. Environ*., vol. 45, no. 3, pp. 537–548, Mar. 2010, doi: 10.1016/j.buildenv.2009.07.004.

[12] T. J. Shields and K. E. Boyce, "A study of evacuation from large retail stores," *Fire Saf. J*., vol. 35, no. 1, pp. 25–49, Jul. 2000, doi: 10.1016/S0379-7112(00)00013-8.

[13] G. Proulx and R. F. Fahy, "The Time Delay To Start Evacuation: Review Of Five Case Studies," *Fire Saf. Sci*., vol. 5, pp. 783–794, 1997.

[14] D. Nilsson and A. Johansson, "Social influence during the initial phase of a fire evacuation—Analysis of evacuation experiments in a cinema theatre," *Fire Saf. J*., vol. 44, no. 1, pp. 71–79, Jan. 2009, doi: 10.1016/j.firesaf.2008.03.008.

[15] R. F. Fahy and G. Proulx, "Collective common sense : a study of human behavior during the World Trade Center evacuation," *NFPA J*., vol. 89, no. 2, pp. 59–67, Mar. 1995.

[16] J. Ouellette, S. Gwynne, R. Brown, and M. Kinateder, "Spatiotemporal mapping of evacuee response," *Fire Mater*., vol. n/a, no. n/a, doi: 10.1002/fam.2842.

[17] H. W. Fischer, G. F. Stine, B. L. Stoker, M. L. Trowbridge, and E. M. Drain, "Evacuation behaviour: why do some evacuate, while others do not? A case study of the Ephrata, Pennsylvania (USA) evacuation," *Disaster Prev. Manag. Int. J*., vol. 4, no. 4, pp. 30–36, Jan. 1995, doi: 10.1108/09653569510093414.

[18] S. Horiuchi, Y. Murozaki, and A. Hukugo, "A Case Study Of Fire And Evacuation In A Multi-purpose Office Building, Osaka, Japan," *Fire Saf. Sci*., vol. 1, pp. 523–532, 1986.

[19] J. Pauls, "Calculating evacuation times for tall buildings," *Fire Saf. J*., vol. 12, no. 3, pp. 213–236, Dec. 1987, doi: 10.1016/0379-7112(87)90007-5.

[20] R. Lovreglio, E. Kuligowski, S. Gwynne, and K. Boyce, "A pre-evacuation database for use in egress simulations," *Fire Saf. J*., vol. 105, pp. 107–128, Apr. 2019, doi: 10.1016/j.firesaf.2018.12.009.

[21] D. A. Purser and M. Bensilum, “Quantification of behaviour for engineering design standards and escape time calculations,” *Saf. Sci.*, vol. 38, no. 2, pp. 157–182, Jul. 2001, doi: 10.1016/S0925-7535(00)00066-7.
[22] J. Pauls, “The movement of people in buildings and design solutions for means of egress,” *Fire Technol.*, vol. 20, no. 1, pp. 27–47, Feb. 1984, doi: 10.1007/BF02390046.
[23] “[77] Number-Bunching: A New Tool for Forensic Data Analysis,” *Data Colada*, May 25, 2019. http://datacolada.org/77 (accessed May 16, 2020).
[24] M. Forssberg, J. Kjellström, H. Frantzich, A. Mossberg, and D. Nilsson, “The Variation of Pre-movement Time in Building Evacuation,” *Fire Technol.*, vol. 55, no. 6, pp. 2491–2513, Nov. 2019, doi: 10.1007/s10694-019-00881-1.

Appendix A

De-identified raw data; the data are organized in a table containing the following columns:

1. Building: a unique randomly generated identifier for each building in the dataset
2. Num_occupants: The number of occupants reported during the event
3. Date_time_event: The date and time of the event
4. Event_total_sec: The reported total evacuation time in seconds
5. Drill_or_evac: Denotes if the data was from a drill or an unplanned evacuation
6. Scenario: Denotes if a drill included an evacuation scenario (yes/no) of it the report refers to an evacuation
7. False_alarm [string]: Identifies false alarms (no = no false alarm; yes = false alarm)

building	**num_occupants**	**date_time_event**	**evac_total_sec**	**drill_or_evac**	**scenario**	**false_alarm**
073	208	2016-05-14 10:15:00	NA	Evacuation	Evacuation	no
d13	NA	2016-12-19 13:47:00	360	Drill	no	no
d12	27	2016-04-18 13:40:00	315	Evacuation	Evacuation	no
d8e	381	2016-11-03 10:01:00	335	Drill	no	no
c8d	42	2016-10-26 11:00:00	285	Drill	no	no
f93	50	2016-11-17 11:30:00	180	Drill	no	no
b4f	60	2016-10-12 10:40:00	90	Drill	no	no
b4f	60	2016-11-09 14:02:00	76	Drill	no	no
a70	80	2016-10-28 09:00:00	193	Drill	yes	no
6c5	400	2016-10-31 10:46:00	540	Drill	no	no
747	32	2016-10-11 09:30:00	240	Drill	no	no

34e	70	2016-09-21 15:00:00	300	Drill	yes	no
f5c	132	2016-10-14 09:45:00	285	Drill	no	no
f5c	132	2016-10-14 10:10:00	345	Evacuation	Evacuation	no
b4a	170	2017-10-20 10:25:00	166	Drill	yes	no
b4a	130	2016-10-20 09:50:00	230	Drill	no	no
7e1	50	2016-11-30 13:30:00	285	Drill	NA	NA
57b	300	2017-10-18 09:55:00	271	Drill	no	no
57b	300	2017-02-10 16:50:00	285	Evacuation	Evacuation	no
57b	150	2017-07-10 12:00:00	360	Evacuation	Evacuation	no
e70	110	2017-05-10 14:00:00	120	Drill	no	no
e70	100	2017-09-13 14:00:00	120	Drill	no	no
d13	110	2017-06-29 13:28:00	180	Drill	no	no
454	3	2017-03-01 20:40:00	NA	Evacuation	Evacuation	no
0a7	42	2017-03-02 11:03:00	300	Evacuation	Evacuation	no
d12	32	2017-10-24 13:05:00	264	Drill	yes	no
c68	16	2017-09-27 10:00:00	240	Drill	no	no
c08	50	2017-05-30 10:03:00	348	Drill	yes	no

c08	40	2017-10-27 13:15:00	328	Drill	no	no
d8e	354	2017-10-31 09:30:00	345	Drill	no	no
1e7	24	2017-08-23 10:50:00	119	Drill	no	no
4e4	22	2016-11-08 14:00:00	150	Drill	no	no
534	NA	2017-10-15 09:45:00	185	Drill	no	no
b4f	48	2017-05-09 09:00:00	91	Drill	no	no
62a	54	2017-11-17 10:18:00	120	Drill	no	no
0f8	20	2017-09-14 10:30:00	248	Drill	no	no
891	60	2017-08-17 14:03:00	480	Drill	no	no
a70	50	2017-06-12 09:14:00	340	Drill	no	no
a70	50	2017-11-03 09:00:00	179	Drill	no	no
6c5	400	2017-05-30 10:45:00	600	Drill	no	no
6c5	400	2017-09-18 11:42:00	720	Evacuation	Evacuation	no
d51	100	2017-10-12 10:20:00	330	Drill	no	no
037	35	2017-12-20 10:23:00	300	Evacuation	no	no
f93	80	2017-12-19 13:30:00	240	Drill	no	no
6c5	375	2017-10-11 10:30:00	360	Drill	no	no

b4f	8	2017-11-08 07:05:00	120	Evacuation	Evacuation	no
64f	38	2017-06-15 15:22:00	109	Drill	no	no
64f	36	2017-10-16 11:01:00	111	Drill	no	no
f76	230	2017-12-06 08:45:00	1020	Evacuation	Evacuation	no
317	100	2017-08-31 10:41:00	240	Drill	no	no
317	100	2017-09-06 14:15:00	NA	Evacuation	no	yes
34e	70	2017-06-22 11:31:00	242	Drill	no	no
34e	60	2017-10-13 09:00:00	NA	Evacuation	no	yes
890	57	2017-08-31 11:25:00	NA	Evacuation	no	yes
890	60	2017-10-19 10:21:00	315	Drill	yes	no
890	132	2017-09-13 13:38:00	255	Drill	no	no
b4a	100	2018-02-28 09:55:00	208	Drill	yes	no
57b	300	2018-10-17 09:50:00	218	Drill	no	no
e70	110	2018-06-07 13:30:00	120	Drill	no	no
d13	115	2018-02-21 10:03:00	240	Evacuation	no	yes
d13	100	2018-05-29 09:10:00	180	Drill	no	no
d13	100	2018-11-30 13:30:00	300	Drill	no	no

63c	3	2018-06-27 02:09:00	240	Evacuation	no	no
534	19	2018-11-21 10:30:00	200	Drill	no	no
f93	80	2018-05-09 10:30:00	180	Drill	no	no
a70	50	2018-05-29 09:00:00	166	Drill	no	no
6c5	375	2018-05-02 10:32:00	360	Drill	no	no
6c5	375	2018-10-23 10:45:00	660	Drill	yes	no
0d6	40	2018-10-23 14:30:00	368	Drill	no	no
d8e	417	2018-04-25 09:49:00	390	Drill	no	no
d8e	378	2018-10-30 09:58:00	380	Drill	no	no
037	42	2018-10-25 10:30:00	110	Drill	no	no
f93	80	2018-10-10 08:54:00	180	Evacuation	no	yes
3c9	250	2018-08-14 12:48:00	450	Evacuation	no	no
3c9	250	2018-10-23 10:00:00	330	Drill	no	no
3c9	250	2018-10-17 07:20:00	NA	Drill	no	no
6c5	375	2018-07-30 08:15:00	240	Evacuation	no	no
1b9	15	2018-10-10 10:30:00	235	Drill	no	no
efd	115	2017-10-13 10:00:00	240	Drill	yes	no

747	32	2017-09-19 09:30:00	186	Drill	no	no
64f	38	2018-06-13 14:57:00	101	Drill	no	no
f76	30	2018-01-24 17:50:00	900	Evacuation	no	no
f76	230	2018-01-02 14:50:00	360	Evacuation	no	no
f76	200	2018-01-31 08:45:00	300	Evacuation	no	yes
f76	230	2018-05-16 13:40:00	390	Drill	yes	no
f76	231	2016-01-11 14:00:00	527	Drill	no	no
f76	300	2018-05-01 10:05:00	303	Drill	no	no
f76	230	2017-10-11 14:00:00	590	Drill	no	no
34e	70	2018-05-08 10:00:00	228	Drill	no	no
09c	150	2018-10-10 14:02:00	270	Drill	no	no
09c	150	2018-08-30 14:39:00	354	Drill	no	no
890	60	2018-07-04 09:36:00	182	Drill	no	no
890	8	2018-02-08 17:15:00	NA	Evacuation	no	yes
890	60	2018-10-12 09:37:00	390	Drill	no	no
c08	50	2016-11-30 13:30:00	285	Drill	no	no
1b9	16	2019-10-11 11:00:00	131	Drill	no	no

d8e	371	2017-05-04 10:01:00	330	Drill	no	no
b72	230	2019-06-13 13:50:00	580	Drill	no	no
b72	230	2019-10-16 13:35:00	392	Drill	no	no
31b	60	2019-06-13 09:36:00	264	Drill	no	no
31b	55	2019-10-08 10:24:00	282	Drill	no	no
fae	140	2019-09-26 14:00:00	166	Drill	no	no
fae	140	2019-05-06 14:00:00	212	Drill	no	no
e7e	28	2019-10-07 11:10:00	277	Drill	no	no
e7e	25	2019-05-22 09:40:00	232	Drill	no	no
2cb	300	2019-06-12 09:38:00	283	Drill	no	no
aa2	35	2019-06-12 10:00:00	150	Drill	no	no
617	150	2019-10-07 11:05:00	180	Drill	no	no
c8d	50	2019-03-19 10:30:00	240	Drill	no	no
902	107	2019-06-27 10:00:00	360	Drill	no	no
5b7	52	2019-10-04 10:45:00	320	Drill	no	no
2cb	300	2019-10-10 09:45:00	340	Drill	no	no
d8e	378	2019-05-09 10:00:00	385	Drill	no	no

dde	112	2018-10-11 13:30:00	118	Drill	no	no
dde	110	2019-10-08 13:30:00	217	Drill	no	no
16b	NA	2019-09-06 09:30:00	NA	Drill	no	no