# Humanities & Social Sciences Communications



# **ARTICLE**

Check for updates

1

https://doi.org/10.1057/s41599-021-00993-6

OPFN

# Impacts of school shooter drills on the psychological well-being of American K-12 school communities: a social media study

Mai ElSherief<sup>1,5</sup>, Koustuv Saha <sup>2,5</sup>, Pranshu Gupta<sup>3</sup>, Shrija Mishra<sup>3</sup>, Jordyn Seybolt<sup>3</sup>, Jiajia Xie<sup>3</sup>, Megan O'Toole<sup>4</sup>, Sarah Burd-Sharps <sup>4</sup> & Munmun De Choudhury <sup>3⊠</sup>

The toll from gun violence in American K-12 schools has escalated over the past 20 years. School administrators face pressure to prepare for possible active shootings, and often do so through drills, which can range from general lockdowns to simulations, involving masked "shooters" and simulated gunfire, and many variations in between. However, the broad and lasting impact of these drills on the well-being of school communities is poorly understood. To that end, this article applies machine learning and interrupted time series analysis to 54 million social media posts, both pre- and post-drills in 114 schools spanning 33 states. Drill dates and locations were identified via a survey, then posts were captured by geo-location, school social media following, and/or school social media group membership. Results indicate that anxiety, stress, and depression increased by 39–42% following the drills, but this was accompanied by increases in civic engagement (10–106%). This research, paired with the lack of strong evidence that drills save lives, suggests that proactive school safety strategies may be both more effective, and less detrimental to mental health, than drills.

<sup>&</sup>lt;sup>1</sup> University of California, San Diego, CA, USA. <sup>2</sup> Microsoft Research, Montreal, QC, Canada. <sup>3</sup> Georgia Institute of Technology, Atlanta, GA, USA. <sup>4</sup> Everytown for Gun Safety, New York, NY, USA. <sup>5</sup> Authors conducted this research while affiliated to Georgia Institute of Technology. <sup>™</sup> email: munmund@gatech.edu

# Introduction

s firearm fatalities surge as a major public health problem in the United States (Grinshteyn and Hemenway, 2019), the country has also been experiencing more shootings in K-12 school premises than nearly all other nations (Grabow and Rose, 2018). Although school shootings are still rare compared to daily gun violence (Nekvasil et al., 2015)—a cause that leads to about 103 daily gun deaths in the U.S. (Ludwig, 2017; Resnick et al., 2017)—current statistics show that school shootings are happening more frequently (Bonanno and Levenson, 2014; Everytown, 2019). Each of these incidents is alarming to both local communities and the nation as a whole (Rygg, 2015), particularly because schools are intended to be safe spaces for children to grow and learn (Williamson, 2019).

Likewise, recent years have seen an increase in calls for improved school safety and preparedness, particularly related to active-shooter incidents (Madfis, 2016; Schildkraut et al., 2020; Terrades and Khan, 2018). According to the U.S. Education Department's National Center for Education Statistics, one of the notable responses to make schools safer has been drill implementation (Diliberti et al., 2017). Since the 1999 Columbine shooting, school shooter drills have proliferated in America's school systems at an exponential rate. By the 2015-16 school year -3 years after the Sandy Hook shooting-95 percent of public schools drilled students on lockdown procedures, with at least 40 states requiring these drills today (Musu-Gillette et al., 2018). However, limited guidance exists on what these drills should look like, or their impacts in light of these many variations. One example of this can be seen in the fact that these drills are often referred to as both "lockdown drills," and "active-shooter drills." While researchers, school safety and mental health organizations, and drill programs (e.g., A.L.I.C.E.) themselves have made strides to standardize these definitions (i.e., lockdown drills prepare students and teachers for general danger, and active-shooter drills specifically address an armed assailant), the general public and school administrators often conflate and/or use the two terms synonymously (see ALICE, 2021; NASP and NASRO, 2017; Schildkraut and Nickerson, 2020). A New York Times headline in 2019 illustrates this lack of clarity on the distinction between these two terms: "How Do You Feel About Active-Shooter Drills in Schools? Nearly every American public school now conducts lockdown drills. Do they make you feel more safe?" Similarly, state statutes on this type of drill are often vague and leave the nature, content, and identification of who participates up to the interpretation of school administrators. As a result, students are required to participate in drills that vary dramatically across America's schools. Some of these drills require students and school staff to remain in hiding in a designated area and practice specific emergency procedures, such as staying quiet, locking the door, and turning off lights, but can also feature tactics such as fighting back, distracting the shooter, and evacuating. In some instances, the drills are unannounced and some present "masked gunmen" actors, simulated gunfire, and fake blood (Gubiotti, 2015).

At the same time, extant research on school shooter drills is mixed and methodologically limited. Ethical limitations and (fortunately) low base rates of active-shooter incidents in schools make it particularly difficult to study lives saved by implementing drills (Jonson et al., 2020). In the absence of this, studies have relied primarily on behavioral observations or surveys to assess drill instruction compliance and/or related perceptions and emotions. Some results suggest that drills improve students' abilities to perform lockdown instructions (e.g., improvements in turning lights off and locking doors), while others do not (e.g., continued difficulties with hiding and remaining silent) (Dickson

and Vargo, 2017; Schildkraut and Nickerson, 2020). Similarly, some suggest that participants report feeling more prepared and less anxious right after completing drills, while others suggest that they feel less safe, more scared, and more concerned that could-be school shooters—given that most are current or former students —now have insight into emergency response strategies (Peterson et al., 2015; Peterson and Densley, 2019; Schildkraut and Nickerson, 2020; Schildkraut et al., 2020; Zhe and Nickerson, 2007). Notably, all of these studies are conducted in individual schools/districts, each of which implements its own drill protocols, thus limiting generalizability of the results to the many diverse drills implemented across the country today. Further, only two of the aforementioned drills explicitly referenced active shooters—at least within study protocols—suggesting that some of the emotional harms caused may be even more heightened under different conditions.

Also missing from the extant research is an exploration of the long-term mental health impacts of drills. This is a particularly important inquiry, because while not every student will experience an active-shooter incident-nor is the evidence on their effectiveness in these situations clear-cut—at this point, most will experience a drill. For students, schools constitute formative years of their life, when their brains and coping skills are still developing—thereby making them react to stress in varied ways (King and Bracy, 2019). Psychologists and health professionals have argued that these tactics can developmentally uninformed (King and Bracy, 2019), and in some cases terrorize already anxious students, increase student fear and anxiety about a shooting occurring, and even risk inducing trauma (Jonson et al., 2020). School teachers, counselors, and parents may experience similarly adverse effects, especially as they assume responsibility for students' well-being and post-drill follow-up, feel pressure to comply with drill guidelines, and experience their own, often unattended, adverse emotional reactions (Goodman-Scott and Eckhoff, 2020). Parents especially also report feeling stress and anxiety over navigating how to explain school safety threats to their children (Kubicek et al., 2008). However, empirical research has not yet assessed whether these impacts sustain over time, or manifest in behaviors and symptoms beyond self-reports.

As a way to close this gap, this article provides empirical evidence on both the long term and widespread mental health impacts of school shooter drills on affected school communities. It does so by applying rigorous, evidence-based machine learning and interrupted time series analysis of mental health-relevant phrases in social media posts made by local school communities both before and after drills, comparing these results with a control group, and further triangulating them through a final set of focus groups. In particular, the findings are based on data that spans 114 K-12 schools and drills, 54 million social media posts, and focus groups of 34 K-12 students, parents, and teachers.

# **Data and methods**

Our approach builds on a growing body of evidence in three complementary research directions. First, the past decade of computational social science research, which has repeatedly showcased how social media postings can provide rich insights about many real-world happenings, whether political, economic, social, or about health and well-being (Golder and Macy, 2011; Lazer et al., 2009, 2020). Specifically, studies in psycholinguistics and crisis informatics have found promising evidence that the content shared on social media can help us to study mental health responses to crises, ranging from understanding how communities cope with protracted wars (Mark et al., 2012), community

violence (De Choudhury et al., 2014; Saha and De Choudhury, 2017), terrorism (Hoffman, 2018), homicides and mass shootings (Glasgow et al., 2014; Jones et al., 2017; Lin and Margolin, 2014). Second, with the growing adoption of social media among K-12 school communities including students, teachers, school administrators, and parents (Kimmons et al., 2018), social media constitutes a promising opportunity to study psychological states unobtrusively and passively. Finally, a growing body of work has been appropriating social media data for observational research, to establish causal relationships between interventions and outcomes, especially in contexts where randomized experimentation may be infeasible or unethical (De Choudhury and Kiciman, 2017; Saha et al., 2019b; Tian and Chunara, 2020). Our approach aligns with these efforts where we circumvent the limitations of existing approaches by leveraging self-initiated, voluntary expressions shared by various school stakeholders on social media. Notably, this study is the most comprehensive investigation of the impact of school shooter drills in school communities in the United States thus far. Supplement S1 gives an extended rationale behind the use of social media data in this research.

**Sources of data**. This research uses responses from a survey and a variety of public social media data.

Data on school drill events. K-12 schools in the United States conduct lockdown and active-shooter drills at various frequencies and using a variety of techniques and simulations. To date, the majority of research on the impacts of these drills has been narrowly defined to individual school models and short term outcomes, thus limiting our ability to capture the reality that countless variations on these models exist in practice, and may have widespread and lasting impacts. There is no comprehensive resource currently available that includes information on these drills at scale or with precision and completeness, however. In fact, schools rarely maintain historical records of drills, as they are often unannounced, unregulated, and do not follow a precise periodic schedule, making it difficult to systematically archive dates on which they occurred.

In the absence of a national dataset, we first identified a sample of drill dates and locations through a survey (Table S1) completed by student, teacher, and parent volunteers of a grassroots gun violence prevention organization. Notably, while volunteers shared a similar goal of ensuring school safety, their views on the appropriateness of drills varied, as some had survived actual shooter events, others experienced trauma following drills, and others still experienced drills that were mostly discussion-based, etc. Likewise, a strength of this sample came not from a uniformity in stances or experiences with drills, but a heightened attention to gun violence prevention strategies, and a related increase in the likelihood of their writing down or recalling the dates of drills.

The survey was fielded through a variety of electronic channels such as emails, electronic flyers, and postings in online groups, and remained active between November 2019 and February 2020. A total of 153 individuals responded and provided information regarding the dates of the school shooter drills occurring between April 2018 and December 2019, the school grade (elementary, middle, or high school), the school name and location (geographical address, city, state, and zipcode). No personal identifiers or social media usernames were collected through the survey, nor was an active social media profile or participation in online discussions about drills a requirement for participation, as the survey's purpose was to identify a sample of drill dates and places, not specific individuals or their perceptions of the effectiveness or impact of the drills, to be featured in this study.

A summary of the key questions in the survey are included in Table S1. Figure S1a gives the geographical distribution of the 138 schools and corresponding drills, spanning 37 states, considered in this work, while Fig. S1b depicts the temporal distribution of these drills in our dataset. 13% of schools in our sample were geographically located on the west coast, 40% in the central U.S., and 47% on the east coast. Of the 138, 63 were elementary schools, 35 were middle schools, 18 were high schools, while 21 were uncategorized. 130 schools were public, public-charter or pre-K, and 8 were private.

Social media datasets. We then assembled a diverse set of public social media posts associated with survey-identified K-12 schools spanning a 90-day period before (~3 months) and a 90-day period after (~3 months) after each drill identified above. We adopted three procedures to build out this dataset: (1) Identifying Twitter posts by individuals who follow the official Twitter account of one of the schools identified to have had a drill event in the survey; (2) Identifying Twitter posts by individuals whose self-reported geo-location on Twitter lies with the officially defined district of one of the drilled schools; and (3) Identifying Reddit posts shared within communities associated with one of the school names. The first approach uses the homophily principle that structures network ties such as those relating to work, support, and other relationships (McPherson et al., 2001). Based on cursory observations, usually, parents, teachers, students, and other members of a school community tend to follow the Twitter account of a school. This can be due to multiple reasons, such as to stay updated with the content posted by the school or engage in discussions with other members of the community. School followers are also more likely to share posts that might be relevant to the school, the events related to the school or the neighborhood. The second approach leveraged the fact that those likely to be directly impacted by the drills will reside within associated school districts' geographic boundaries. The third and final approach leveraged the observation that K-12 related discussions often happen in a variety of Reddit communities (or "subreddits"), ranging from dedicated school subreddits, to the neighborhood or city in which the schools are located. In all, we analyzed 27.8M social media posts shared between January 2018 and March 2020 by 542.27K unique users for 114 schools (out of the 138 extracted from the survey responses) that were present in one of the three data collection strategies and located in 33 states. Detailed information on this data collection are given in Supplement S2.1, S2.2, S2.3, Table S2, and Fig. S11. We also gathered an equal size sample of control postings. Specifically, we selected a random sample of about 27M posts from the 1% sample of Twitter's public stream whose geo-location (or the self-reported location of the post author) lied outside the 114 school districts. These control posts spanned a 90-day period before (~3 months) and a 90-day period after (~3 months) each of the drills identified in the above survey, thus making their timeframe the same as our social media data on the drills, and thus providing us a total of 54 million posts on which the ensuing analyses were conducted.

We note here that we intentionally did not filter the social media posts to only look at those with explicit mentions of the drill events. This is because literature indicates that well-being concerns may manifest in different ways in an individual's interaction with others, sometimes with explicit references to the underlying causes and sometimes implicitly embedded in other conversations (Kícíman et al., 2018). Moreover, psychological state, are associated with non-content words, such as articles and pronouns, which are rarely consciously regulated by the individual and may span a variety of topics directly or indirectly related to a crisis, whether offline (Cohn et al., 2004) or online

(Saha and De Choudhury, 2017). Analyzing all temporally and school-community relevant longitudinal social media data mitigates issues of partial observability of well-being outcomes that may result from focusing on drill-specific postings.

Machine-learning methods. Stress, anxiety, and depression are often fueled by unconscious factors that people are unable to pinpoint. Similarly, they manifest in many aspects of our lives, including our general thought processes (Palen and Anderson, 2016; Tausczik and Pennebaker, 2010). Language can reflect these affective, cognitive, perceptual, and other psychological processes of individuals, including changes over time around specific events (school shooter drills in our case) (Pennebaker et al., 2003). In our case in particular, anyone who experienced a drill could experience these impacts, positive or negative, without being consciously aware of what was driving them. Therefore, our analysis was not limited to just individuals who recalled the date and location of a drill enough to report it in our survey, or people who posted their thoughts or experiences about drills online rather we analyzed all social media postings of individuals in the school community datasets compiled above, regardless of their stance on the drills or participation in our initial survey. Motivation behind our analytic approach is further borrowed from prior research in crisis informatics (Palen and Anderson, 2016; Reuter and Kaufhold, 2018), where it has been observed that, from a psycholinguistics perspective, for both positively and negatively affected individuals, the effects of a certain event are likely to spill into their linguistic expressions across a wide range of topics, not just those relating to the event (De Choudhury et al., 2014; Mark et al., 2012). We also note that similar approaches have been used in prior crisis informatics research: De Choudhury et al. (2014) analyzed Twitter postings from the entire community to understand the affective responses, both positive and negative, to the Mexican Drug War, and Saha and De Choudhury (2017) looked at Reddit posts of the entire college community to assess stress levels following incidents of violence on campus.

We adopted a two-dimensional approach to quantify psychological impacts as observed via social media posts: mental health symptoms and psycholinguistic expressions, described below.

Mental health symptomatic expressions. We quantified symptomatic expressions of stress, anxiety, and depression using machine-learning classifiers that were built and validated in prior work (Saha et al., 2019b). These are essentially n-gram based binary support vector machine (SVM) models trained using transfer learning methodologies (Pan and Yang, 2009)—transfer learning is a method where a model developed for a task is reused as the starting point for a model on a second task. Therefore, the main idea here is to infer mental health outcomes in an unlabeled data by transferring a classifier trained on a different labeled dataset, as first introduced by Bagroy et al. (2017). The positive class of the training datasets in this approach come from appropriate Reddit communities (r/depression for depression, r/anxiety for anxiety, r/stress for stress), and the negative class of training datasets comes from non-mental health-related content on Reddit—a sample of 20M posts, gathered from 20 subreddits (such as r/AskReddit, r/aww, r/movies, and others) that appeared in the landing page of Reddit during the same period as the mental health subreddit posts. These classifiers have been found to yield a high-performance accuracy of approximately 0.90 on average and transfer well on Twitter with an 87% agreement between machine-predicted labels and expert appraisal (Saha et al., 2019b). Details of the validity of these classifiers are in Supplement S4. Given the high likelihood of comorbidity of stress

and anxiety in crisis contexts (O'Donnell et al., 2004), our analyses combined these expressions for the ease of exposition, that is, any post that expresses either or both of *High* stress or anxiety (according to our classifiers) are labeled as high symptomatic expression.

Psycholinguistic expressions. People's affective, cognitive, perceptual, and other psychological processes can be reflected in their language and its changes around specific crisis events (active-shooter drills in our case) (Pennebaker et al., 2003). As we seek to understand psycholinguistic expressions, we employed the well-validated psycholinguistic lexicon, Linguistic Inquiry and Word Count (LIWC) (Tausczik and Pennebaker, 2010). This tool is known to work well with short text and social media data, as revealed in a large body of social computing and crisis informatics literature (Lin and Margolin, 2014; Mark et al., 2012; Saha and De Choudhury, 2017). We used the following non-affective psycholinguistic expression categories given in the LIWC dictionary: "cognition and perception", "interpersonal focus", "temporal references", "lexical density and awareness", and "social and personal concerns."

Temporal analytic technique. We investigated two types of temporal change in expressions of psychological well-being, drawing on the interrupted time series (ITS) literature (Tian and Chunara, 2020): immediate change (IC) and a longer-term change (LC). ITS is a quasi-experimental study adopted as an alternative to randomized control trial research designs, typically used to measure the causal effect of an intervention, by controlling for confounds, and primarily through its control over regression to the mean. ITS is an analysis of a single time-series data before and after the intervention (Bernal et al., 2017; McDowall et al., 2019). This method also contains a strong inferential power and has wide applications in epidemiology, medication research, and program evaluations in general (Chandrasekharan et al., 2017; Linden, 2013). The immediate change (IC) in outcomes is calculated by measuring the difference in z-scores of the daily measures immediately after and before the drill. On the other hand, longer-term change LC measures the relative change, between the After and Before periods, in terms of the average proportion of social media posts that are associated with a certain outcome. First, for all time series corresponding to a mental health or psycholinguistic expression, we calculated the immediate change or IC by performing an ITS analysis (Saha and De Choudhury, 2017; Tian and Chunara, 2020). We first normalized all of the time series corresponding to each well-being outcome, and spanning the whole time duration under examination (180 days), using the statistical standard score (z-score). The z-score measure allows comparison of trends across time and data types because of a reliance on proportions, rather than raw frequencies. For this, aggregated across all schools and their respective drills, we fitted a linear function for the 90 days of data (mental health symptomatic or psycholinguistic expression) before the drill (Before) and a linear function similarly for the 90 days after the drill (*After*). Hence, *IC* in each of the outcomes of psychological well-being (e.g., stress/anxiety, depression, etc.) was calculated as the interrupted change at time zero—the day of the drill event—based on the difference in intercepts of the After linear fit with respect to Before—essentially capturing differences in levels surrounding the event. Contrastively, the long-term relative change measure, LC gives how each of the outcomes were manifested over the entire 90-day duration following the drills, in comparison to that preceding them. For all outcomes, to assess whether the IC and LC differences between the Before and After periods were statistically significant, we performed Welch t-test,

followed by Benjamini-Krieger-Yekutieli False Discovery Rate (FDR) correction, given multiple comparisons. Further details are given in Supplement S6.

**Establishing causality**. Although interrupted time series analysis is a fairly strong quasi-experimental design (Cousens et al., 2011), recent research has shown that it can either fail to identify the effects of external factors on the time series, resulting in a false causal attribution, or conversely, confuse the causal interpretation when a directionally correct change in the time series also occurs prior to the intervention (Linden, 2018). To reduce bias and strengthen causal interpretation in interrupted time series analysis studies, the treated unit's outcomes should be contrasted to those of a "control" group that is comparable on observed characteristics (including, at a minimum, the baseline level and trend of the outcome). For this reason, various robustness checks are often recommended to determine if treatment effects persist under various data and model specifications, and to individuals outside of the sample used for model estimation (Kiciman and Sharma, 2019; Linden, 2018). Consequently, to determine if the change we see spanning from 90 days before the drill and until 90 days after the drill, is really a causal effect of the active-shooter drill, we employed four methods, as four types of robustness checks. These approaches are motivated from prior work (Saha and De Choudhury, 2017; Saha et al., 2018), and they essentially minimize confounding effects of social media expressions such as changes due to seasonal, local, and other coincidental factors on our outcomes of interest.

Temporal alignment based on offsets from drills dates and isolating impacts of other events. In the **first method**, while examining the changes for the various psychological well-being outcomes, we aggregated all the posts associated with all of the schools in our dataset spanning the *Before* and *After* periods. That is, we temporally aligned the *Before* and *After* datasets for all schools across their respective drill dates by using the date offset of each social media post with respect to the associated school's drill date (day zero). As the drills happened at various different times of the year (ref. Fig. S1b), this alignment of the *Before* and *After* periods would minimize the confounding effects of other stressors that are seasonal or subject to school-specific events.

In addition, we examined the impact of other events that might confound the changes in well-being outcomes. In particular, since gun violence incidents may also trigger similar patterns of responses, we isolated *temporally* and *spatially* those schools impacted by such an event during our study period. Should the patterns of changes in well-being outcomes hold with the temporal and spatial exclusion of these events, it would be logical to ascribe the changes to the drills.

While there is no broad consensus on what constitutes a gun violence incident on a school campus, we ground our definition with that of mass shooting incidents derived from the Congressional Research Service and the Gun Violence Archive; this definition considers an incident to be a mass shooting if four or more people shot (injured or killed) (Bagalman et al., 2013; Gun Violence Archive, 2020). We implemented two different settings: (1) Temporal overlap mitigation: removing all schools that had drill dates that coincide within the same month as when a gunfire mass shooting occurred. For this analysis, we considered the following six mass shooting events occurring in Santa Fe, TX, Charlotte, NC, Highlands Ranch, CO, Atlanta, GA, Santa Clarita, CA and Edgard, LA in May 2018, April, May, August, November and December 2019, respectively. (2) Spatial overlap mitigation: removing all schools whose locations were in the same state where a gunfire mass shooting events occurred. For this analysis, we considered the entire duration of our analysis (90 days before the earliest drill date and 90 days after the latest drill date) to pinpoint these gunfire events. This led to the identification of eight mass shooting events in nine states including the six incidents from the temporal analysis in addition to two mass shooting events in Benton, KY and Parkland, FL. These states were Kentucky, Florida, Texas, North Carolina, Colorado, Georgia, California, and Louisiana.

Comparison with the counterfactual. Per the ITS framework, one way to infer causality is to estimate what the outcome variable could look like, if there were no intervention (Cousens et al., 2011); in our case, this would translate to inferring the future well-being outcomes based on the past outcomes, assuming that the drill (an intervention) has not occurred. To do this, we train an an autoregressive integrated moving average (ARIMA) model (a statistical model for analyzing and forecasting time series (Hyndman and Athanasopoulos, 2015)) on the 3-month period prior to the drill events for stress/anxiety and depression (Before period). We then use this trained model to forecast the future values of the two well-being outcomes in the aftermath of the events (the 3-month long After period). Our goal here is to understand what the time series would look like without the drill occurring. If significant differences were to be found between the predicted time series and the actual time series in the *After* period, with the predicted time series showing lower values for the wellbeing outcomes, we can conclude that the increase in stress/ anxiety and depression during the After period is due to the occurrence of the drills (the intervention).

*Comparison with a synthetic control.* The **third method** tested for the validity and robustness of our causal claims—that the activeshooter drills impacted mental health and community outcomes —by using permutation tests conducted on the (treatment) time series data on our mental health symptomatic expressions of anxiety/stress, depression, and the six LIWC-based psycholinguistic categories associated with community outcomes. This examination aimed to rule out the possibility of observing the temporal changes by chance (Anagnostopoulos et al., 2008), and serves as a comparison of the actual (treatment) time series changes with a "synthetic time series". Essentially, for each outcome measure, we generated 1000 synthetic time series, which followed similar non-parametric data distribution as the treatment (actual) time series data. Then, we compared the relative change in each outcome around a placebo drill date in control (mid-point of the time series) against the actual drill date in our treatment data. That is, for each synthetic time series, we measured the LC (synthetic LC) and record if it is larger than the actual LC. We measured the probability (p-value) that the synthetic LC is greater than actual LC, which helps to quantify the statistical significance of our observations against chance or random observations. This method emulates permutation test frameworks applied in the prior work (Das et al., 2020; Saha, 2019a), and tests for the null hypothesis that outcome change around a randomly generated drill date is comparable to outcome change around actual drill date. If this *p*-value is found to be zero or significantly low (e.g., p < 0.05), then we can deduce that the treatment LC is indeed attributed to the effects of the activeshooter drill.

Comparison with an actual control. The **fourth method** involved comparing our temporal trends (treatment), specific to the schools and their respective drill dates, with a suitably chosen control time series in the same timeframe (De Choudhury and Kiciman, 2017). Specifically, as described above, we use the random sample of 27M posts from the 1% sample of Twitter's public

stream (control) that covered the *Before* and *After* timeframes for our school dataset, and whose geo-location lied outside of the school (districts) under considerations. We then proceeded to calculate the date offsets of each post in the control data, from the school drills, and measured the various psychological well-being outcomes of interest: mental health symptoms like anxiety/stress and depression, and the six LIWC-based psycholinguistic expressions for each Twitter post. By aggregating the results for each offset, we were able to construct a control time series and compute the temporal trends for each outcome for our control dataset.

#### Focus group triangulation

Rationale. Finally, the research team conducted semi-structured interviews in the form of focus groups with students, parents, and K-12 school teacher volunteers at a grassroots gun violence prevention organization, to provide a school community-centered contextualization, to triangulate our results, and to ensure that our findings mimicked and were informed by lived experiences. As referenced above, volunteers shared a common commitment to ensuring school safety, but different stances on, and experiences with, both drills and actual active-shooter incidents, rendering them a relatively diverse set of focus group participants.

This supplemental triangulation methodology was guided by (Boyd and Crawford, 2012)'s provocation, "in this computational turn, it is increasingly important to recognize the value of 'small data"-a mixed methods approach enables us to represent the voices of those directly impacted by drills in two different ways in our investigation. First it allows for a large-scale, generalizable analysis of the students, parents, and teachers naturalistically shared social media data and juxtaposes it with a sample of their retrospective interpretation of those incidents. These interpretations may not be apparent in an individual's online content, but stem from offline factors that serve as motivations and intentions behind online behaviors. Essentially, our approach allows us to more accurately uncover the social meanings and offline implications of online articulations of the psychological impacts of the drills. Second, in the absence of any prevailing knowledge or "ground-truth" about the psychological impacts of drills, our approach enables aligning machine-learning inferences and estimations of well-being outcomes of the drills, with the selfreported lived experiences of the stakeholders to gain concurrent validity in the quantitative findings. Likewise, other recent research has used qualitative data, such as that from field work and domain experts, to enhance the validity or ground truth of computational linguistic analyses (Stuart et al., 2020). Like Fine (2006), we adopt this term, which had first appeared in meteorological research to denote the practice of checking weather forecast models against direct observations of weather conditions occurring "in the real-world" (e.g., verifying that a tornado has in fact touched down at the geographic coordinates indicated by remote sensing). Here, ground truth refers specifically to the correspondence between social media signals of the well-being impacts of the school shooter training programs and the trauma that school communities experience when they undergo these drills. Summarily, contextualization and validity via adequate representation of the voices of school community members, as supported by our chosen methodology, are key to ensure that the outcomes of this project accurately depict reality and are not misconstrued as provisions for school safety are evaluated and implemented in K-12 schools.

Nevertheless, it is important to note from McDonald et al. (2019) that, qualitative interviews cannot be expected to provide the same type of quantitative validity as large-scale data would, because the qualitative and quantitative approaches

fundamentally differ epistemologically and ontologically (Hunt, 1991). Qualitative data can be useful to unpack the stories behind quantitative observations through triangulation, with the intention to increase or deepen understanding of the study phenomenon (Hussein, 2009). In this work, our focus is on social and interpreted, rather than quantifiable, phenomena and we aimed to discover and describe rather than to test and evaluate. Therefore, we deemed the latter to be appropriate for this study, which served as rationale behind our focus group study design, described next.

Focus group study design. Our focus group based validation approach interviewed 34 stakeholders through 6 focus groups. Before recruiting and engaging with participants. Individuals 15 years or older were invited to participate in 1-h interviews. Participants were eligible if they engaged (teachers and students) in at least one school shooter drill in a U.S. based school in 2018–2019 or if they had a child who participated in at least one school shooter drill in a U.S. based school in 2018–2019.

Approach. We conducted 6 1-h long focus group interviews with 21 parents, 11 teachers, and 2 students. All except one participant were female. All focus groups were conducted by three coauthors over the teleconferencing software Zoom; an in-person format for these discussions was not possible at the time due to the geographical spread of the participants as well as social distancing/travel restrictions imposed by the Coronavirus Disease 2019 (COVID-19) pandemic (Organization et al., 2020). Discussions were audio-recorded with permission. Participants were asked to freely share their experiences with and stances on drills—their efficacy, benefits, harms, and engage in semi-structured group discussions based on several key results from the social media data-driven results of this study particularly around the observed psychological well-being impacts in the school communities. Guiding questions are included in Table S4.

We then transcribed the audio recordings, along with removal of any personally identifiable information, and stored them in secure, two-factor authenticated, encrypted servers. Transcribed data was analyzed using an inductive and iterative semi-open coding approach (Mayring, 2004). Open coding is common in the analysis of qualitative research and is an established approach in Grounded Theory (Charmaz, 2014). With open coding, a first step is breaking up the data into discrete parts and creatingcodes— by hand to label them. The purpose of breaking up the data and labeling them with codes is to enable the researcher to continuously compare and contrast similar events in the data. This is done by collating all pieces of data (such as quotes) that were labeled with a particular code. In this work we have focused on social and interpreted, rather than quantifiable, phenomena and aimed to discover, interpret, and describe rather than to test and evaluate; therefore this analytic approach was deemed appropriate (McDonald et al., 2019). We used a primary coder with a background in Psychology and a secondary coder with a background in Computing for the tasks; both coders are authors of this paper. Through mutual and iterative discussion, the coders relied on our quantitative methodology and well-being outcomes of interest (Cohen and Wills, 1985; Sullivan, 1996). Our final list consisted of 15 codes (Fig. S5). Finally, the researchers used this codebook to code all of the transcripts and identify interpretive broader themes that aligned with lived experiences of the school stakeholders around the drills. Paraphrased quotes from the participants are inter-dispersed among the discussion of results when appropriate throughout the main manuscript, along with attributions of specific quotes to specific participants—the 21 parents are referred to as P1-P21, the 11 teachers as T-T11, and the 2 students as S1 and S2.

# Results

Impacts on mental health. We first discuss the aggregated temporal patterns of mental health symptomatic expressions before and after the school-specific drill events. Based on Fig. 1a, we observe a notable increase in stress or anxiety and depression markers. To quantify the extent of change, we first measure the mean proportion of posts that are indicative of high stress or anxiety and depression for Before and After-drill periods (the measure LC). We find that the mean proportion of posts indicative of stress or anxiety to be 0.281 for the Before period and 0.399 for the After period; an LC of 42.1% increase (t = 19.1,  $p < 10^{-15}$ ). As for depression, we find that the mean proportion of posts indicative of depression is 0.125 and 0.173 for the Before and After-drill periods, respectively. This change in depression constitutes an LC of 38.7% (t = 10.13,  $p < 10^{-15}$ ). While we not only observe a statistically significant increase in the levels of stress/anxiety and depression in terms of LC, when we calculate the trends for the Before and After time series, as shown in Fig. 1b, c using a linear fit (see Supplement S6), we observe immediate changes (IC) as well. The After trends seem to be sustained (in the case of anxiety or stress, the slope is 0.00144) or show an increase (in the case of depression, the slope is 0.0053). We additionally measure the IC for stress/anxiety and depression for the z-score distribution and find them to be 0.936 and 0.545, respectively. Examples of anxiety/stress and depression-indicative

posts shared by different school communities in the aftermath of the drills are given in Table S6.

To give richer context to our analyses, we next examine the linguistic markers (n = 1-grams) present in posts classified as indicative of *High* stress/anxiety or depression using a commonly used lexical analytic generative model known as SAGE (Eisenstein et al., 2011) (see Supplement S5 for a description of the technique). Table 1 has two parts corresponding to the anxiety/ stress and depression expressions: on the left, it shows the top salient words that uniquely characterize the Before time period but not the After time period and vice versa. Based on Table 1, positive words like proud, grateful, best and great exhibit saliency in the Before period and a decreased usage after the drills among the High stress or anxiety posts. In contrast, home, school, kids, community, and help show increased use in the High stress/ anxiety posts after the drills. This is indicative of people sharing concerns about kids, classrooms, and schools in the aftermath of the school drills. The same trends are sustained for High depression posts. We notice positive emotions and thoughts such as excited, beautiful, care, best, and amazing to be salient before the drills and calls for support and help afterward (hope, love, help, support, need, young, thank, family).

Based on the above, our results indicate that school shooter drills can negatively impact the well-being of school communities

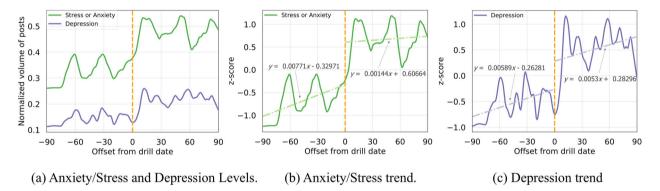


Fig. 1 a Aggregated temporal variation of mental health symptomatic expressions. Also shown are trends corresponding to (b) anxiety/stress and (c) depression in a 90-day period before and after school shooter drills. Solid lines in (b) and (c) denote a 2-week moving average of the normalized volume of posts.

Table 1 Top salient n-grams (n = 1), identified using SAGE (Eisenstein et al., 2011), in *High* stress/anxiety and depression posts before and after school shooter drills.

| Anxiety or Stress Expressions |              |          |        |           |             |          |      | Depression Expressions |              |          |        |             |             |          |      |
|-------------------------------|--------------|----------|--------|-----------|-------------|----------|------|------------------------|--------------|----------|--------|-------------|-------------|----------|------|
| Before Drill                  |              |          |        |           | After Drill |          |      | Before Drill           |              |          |        | After Drill |             |          |      |
| n-gram                        | SAGE         | n-gram   | SAGE   | n-gram    | SAGE        | n-gram   | SAGE | n-gram                 | SAGE         | n-gram   | SAGE   | n-gram      | SAGE        | n-gram   | SAGE |
| casebiloxi                    | -1.54        | im       | -0.43  | free      | 0.91        | need     | 0.33 | jan                    | -1.12        | ago      | -0.37  | help        | 0.75        | way      | 0.36 |
| lol                           | -1.02        | taking   | -0.40  | morning   | 0.68        | team     | 0.29 | welcome                | -1.04        | hear     | -0.36  | support     | 0.65        | say      | 0.34 |
| shit                          | -0.84        | man      | -0.39  | school    | 0.67        | time     | 0.25 | listen                 | -0.87        | text     | -0.36  | stay        | 0.63        | love     | 0.32 |
| god                           | -0.83        | best     | -0.37  | check     | 0.66        | thank    | 0.25 | grade                  | -0.67        | amazing  | -0.33  | twitter     | <b>0.61</b> | thank    | 0.30 |
| job                           | -0.80        | dont     | -0.30  | read      | 0.64        | ready    | 0.25 | education              | -0.64        | taking   | -0.33  | right       | 0.61        | work     | 0.29 |
| education                     | -0.71        | link     | -0.26  | help      | 0.64        | today    | 0.25 | hate                   | -0.54        | status   | -0.32  | time        | 0.58        | today    | 0.29 |
| service                       | -0.68        | got      | -0.26  | week      | <b>0.64</b> | right    | 0.22 | winter                 | -0.52        | dont     | -0.32  | school      | <b>0.57</b> | college  | 0.27 |
| thats                         | -0.65        | people   | -0.25  | kids      | 0.55        | new      | 0.22 | teacher                | -0.46        | action   | -0.31  | want        | 0.54        | need     | 0.24 |
| thing                         | -0.52        | grateful | -0.25  | high      | 0.46        | family   | 0.19 | important              | -0.44        | hard     | -0.27  | reading     | 0.48        | student  | 0.23 |
| oh                            | -0.48        | proud    | -0.24  | make      | <b>0.43</b> | awesome  | 0.19 | best                   | -0.41        | trying   | -0.24  | hope        | ■ 0.47      | doing    | 0.22 |
| teachers                      | -0.47        | years    | -0.19  | year      | 0.38        | working  | 0.19 | thats                  | -0.41        | know     | -0.23  | tonight     | 0.43        | good     | 0.22 |
| fpsctpic                      | -0.47        | board    | -0.16  | love      | 0.36        | tomorrow | 0.18 | beautiful              | -0.41        | excited  | -0.23  | story       | 0.40        | tell     | 0.20 |
| meeting                       | <b>-0.46</b> | know     | -0.15  | want      | 0.35        | home     | 0.18 | thing                  | <b>-0.40</b> | care     | ■-0.23 | week        | ■ 0.39      | tomorrow | 0.18 |
| great                         | -0.46        | winter   | -0.14  | community | 0.33        | students | 0.16 | person                 | -0.39        | building | -0.22  | community   | 0.37        | morning  | 0.16 |
| forward                       | -0.43        | learning | I-0.14 | game      | ■ 0.33      | broken   | 0.16 | yes                    | -0.37        | high     | -0.22  | better      | ■ 0.36      | family   | 0.16 |

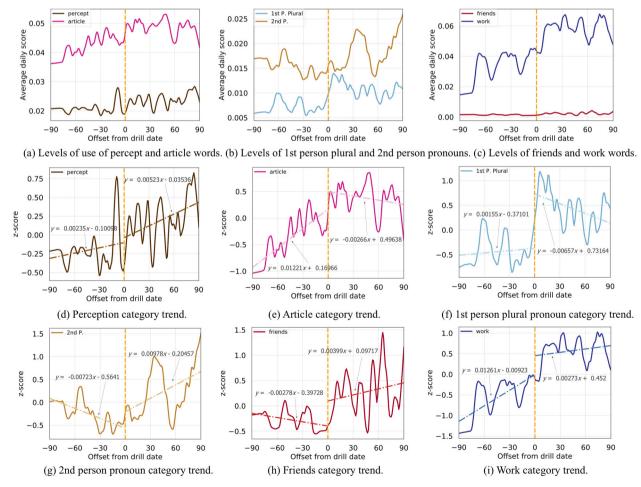
SAGE outputs words (and their associated saliency scores) that are more uniquely and distinctively likely to be used in posts following the drills, in comparison to those in posts prior to the conduct of the drills. Positive scores (**Teal** bars) indicate greater saliency in the *After* data, and Negative scores (**Golden** bars) indicate greater saliency in the *Before* data. Bar sizes represent SAGE magnitude.

7

over prolonged periods of time. The focus groups provide further support and convergent validity to these observations. For instance, parents and teachers noted that drills can be triggering. Many students were "texting their parents, praying, crying" (teacher participant #2 or T2) because they thought "they were going to die," (T1) and this caused many of them to remain nervous long after the drill was over, even prompting extreme reactions such as panic attacks and "downright fear" (P1) in response to other unrelated, innocuous situations such as "a fire alarm going off" (P2). The drills caused even seasoned teachers to "break down at recess" (teacher participant #7 or T7) on the day of the drills. Other reactions included avoiding talking about the school drill experience as a result of being desensitized; "It was like nothing happened. It was the same thing as breaking a pencil (P3)" and "It's just kind of part of their norm. She's been doing it ever since she was in preschool (P9)".

Community outcomes. Next, we explore the psycholinguistic aspects related to community outcomes, such as the LIWC categories of perception, article, second person pronouns, first-person plural pronouns, friends, and work, as given in Fig. 2. These categories have been noted to be the most salient to understand social media dynamics around crisis events (Lin and Margolin, 2014; Saha and De Choudhury, 2017). Results pertaining to the rest of the non-affective psycholinguistic expressions can be

found in Table S7. Based on Fig. 2a, we find that words that invoke perceptual processes (e.g., see, hear, and feel), summarized in the LIWC category of perception as well as those in the lexical density and awareness LIWC category article (e.g., a, an, the) show an increase in terms of LC; 10.78% (t = 3.6,  $p < 10^{-4}$ ) and 16.3% increase (t = 9.5,  $p < 10^{-15}$ ), respectively. In terms of IC (difference in intercepts of the linear fits between the After and Before periods), we observe that both of these LIWC categories show a positive change: 0.084 and 0.353, respectively. These increases, which are often associated with first-person accounts of unanticipated incidents as well as greater awareness of and attention to one's surroundings (Tausczik and Pennebaker, 2010), together indicate that, to the students, parents, and teachers, the drills impacted their cognitive mechanisms in a way that traumatizing crises would, instilling fear and confusion. They also show that, in the 90 days following drills, social media conversations featured significantly more words that show attributions, attempts to make sense of why something occurred, and reflections on the experience itself and what feelings it evoked. The focus group interviews provided more context and credence to these observed cognitive changes. A teacher (T10) said that she could not "really shake the feeling of [being faced with an active shooter]" although she knew "who was rattling the door" and "what was going on." To a student (S2) participating in the focus groups, a similar acute response accompanied the drills-he felt convinced there was a shooter on campus" whereas a parent



**Fig. 2 Aggregated temporal variation of psycholinguistic expressions.** The figures show trends in the form of LIWC categories *perception, article, 1st person plural pronouns, second person pronouns, friends,* and *work,* in a 90-day period before and after school shooter drills. Solid lines in **a-c** denote a 2-week moving average of the daily values. **d-i** represent the z-score trends computed on the time series in **a-c. d-i** also include linear fits in the form of dotted lines, corresponding equations indicating the nature of the fit.

(parent participant #3 or P3) said that their child reported "[hearing] someone going and jiggling the doorknobs in the classrooms" well after the drills were over.

In terms of the LC of interpersonal focus (Fig. 2b), our analysis shows a 47.79% increase  $(t=8.1,p<10^{-15})$  and a 10.21% increase (t=3.1,p<0.01) in the usage of first-person plural pronouns and second person pronouns, respectively, for the *After* period compared to *Before*. For outcomes related to personal and social concerns such as *friends* and *work*, per Fig. 2c, we observe a 33.7%  $(t=3.5,p<10^{-4})$  and a 106.18%  $(t=18.3,p<10^{-15})$  increase, respectively, in the *After* period. For all these categories (Fig. 2d-i), the IC is also positive (0.423-1.175), indicating that overall, the levels of psycholinguistic expressions became more prominent or heightened right after the drills. These results summarily indicate an increased sense of solidarity and engagement among members of school communities, despite the mental health impacts noted above.

This is echoed by teachers in our focus groups, where one teacher (T3) mentioned that she "felt more comfortable bringing this up with like-minded people who are concerned..or if we know somebody understands that that was like a stressful event" in a mom's meeting or other kinds of gatherings. A parent (P8) similarly said that, after the events happened in her child's school, she felt the need to "bring [the issues around school shooter drills] up" in her conversations with other moms, who she noted felt overwhelmed, but did not "necessarily [have] the time to process it, think it through, [or] follow up with their kids." A student (S1) confirmed that the drills spark conversations in her school about [preparation for an active shooter should it happen...a lot of the teachers would ask like do you think that was helpful in preparing you for this event?]. The student elaborates that the main response is that [it stresses students out and that it can cause sort of a collective worry among my classmates]. Together, these results suggest that school shooter drills in schools may fuel collective action and community advocacy.

Additional analyses to examine causality. To what extent can we causally attribute the changes observed in the trends of various psychological well-being outcomes, such as anxiety/stress, depression, and the psycholinguistic categories, to the drill events? To answer this, we adopted a multi-prong causal analysis approach outlined in above in the Data and Methods section.

Isolating impacts due to other events. Recall that, beyond drills, mass shootings themselves can lead to an increase in stress and anxiety-inducing online conversations; thus potentially affecting overall measurements of psychological well-being in online discourse. To account for these incidents and their potential impact on the well-being of the school communities, we report on how well-being outcomes changed from our analyses, when drills that overlapped with these events were excluded spatially or temporally. First, the temporal overlap mitigation led to the exclusion of 31 drill events corresponding to approximately 1.5M data points/ posts from our analysis. Figure S6a-S6c show the results for this analysis. We note that this resulted in an (LC) of 43.12% increase in stress/anxiety (p < 0.001) and a 39.58% increase in depression (p < 0.001). Next, the spatial overlap mitigation led to the exclusion of 41 drill events corresponding to approximately 10.7M data points/posts from our analysis. Figure S6d-S6f show the results for this second analysis. We observe the following results: an LC of 30% increase in stress/anxiety (p < 0.001) and a 29.6% increase in depression (p < 0.001). Together, these results show that even when accounting for large-scale mass shooting events (LC range [29.6-43.12%] with changes in the range [-12.1-1%] in comparison to the original results), the changes

in well-being outcomes still persist and could be attributed to the actual real-world intervention of conducting a drill.

Comparison with the counterfactual. Based on the approach described in "Data and Methods", we fitted two ARIMA models for stress/anxiety and depression, respectively, based on Akaike's Information Criterion (AIC) score, and using data from the 90day Before period. The best performing models were ARIMA models with an order of (0, 0, 0) and (1, 0, 0) for stress/anxiety and depression, respectively. Upon using these fitted ARIMA models to forecast values in the 90-day After period, we found the Root Mean Square Error (RMSE) between the forecasted values and the actual time series to be 0.14 and 0.09 for stress/anxiety and depression, respectively. Figure S7a and S7b show the predicted time series based on the ARIMA models (predicted) versus the actual time series based on our well-being measurements during the After period. On average, we find that the predicted time series exhibits lower levels (26.3% and 50% lower than the actual levels) of stress/anxiety and depression for the After prediction (p < 0.001). This shows that the two time series (actual versus predicted) are significantly distinct, indicating that the interventions (drill events) likely caused a notable change in the expressions of depression and anxiety/stress, as demonstrated in the social media posts of the respective school communities.

Comparison with an synthetic control. In addition, we conducted the robustness test of causality using permutation tests, and across 1000 permutations of synthetic time series for each outcome; Figure S8 shows our results. In all the cases, the relative change in the After period compared to Before in the synthetic time series, around placebo drills, was significantly smaller than that in the treatment (actual) data. For the various synthetic time series, we find that the probability (p-value) that a placebo drill leads to a greater relative change than actual drill is extremely low for all the measures: p = 0 for stress/anxiety, depression, *article*, first-person plural, second person, and work, and p = 0.021 for *perception*, p = 0.035 for *second person*), and p = 0.013 for *friends*. This indicates statistical significance in all our observations at the p < 0.05 level, suggesting that changes observed in our study can indeed be attributed to the actual (treatment) school shooter drills, and are not by chance.

Comparison with an actual control. Finally, we compared the time series of normalized posts associated with all of the outcomes: anxiety/stress, depression, and the LIWC categories, with an equal duration control time series capturing these same outcomes in the same timeframe of the drills, in non-school-specific Twitter data. As per Fig. 3 and Table S5, the control time series show a change (LC) of only -0.375-2.5% (and corresponding 0-lag cross correlation of 0.593-0.988) across all outcomes in their second 90-day period, compared to their first, assuming the mid-point of the control time series as the placebo drill. This change is statistically significantly lower than the change (LC) of 10.2-106.18% for the treatment (actual) time series whose 0-lag cross correlation between *Before* and *After* periods is much lower (-0.009 to 0.043), allowing us to associate the changes in the treatment to the drill events. Additionally, many of the changes we observe for the control time series were found to be statistically insignificant  $(0.36 \le p \le 0.77)$  for all outcomes with the exception of stress/anxiety (LC = 0.3%, p < 0.05), work (LC = 2.5%, p < 0.05), and perception (LC = 0.43%, p < 0.05).

**Subgroup analyses**. We additionally investigate changes in the outcomes broken down by three different community groups who are likely to be directly experiencing school drills—teachers,

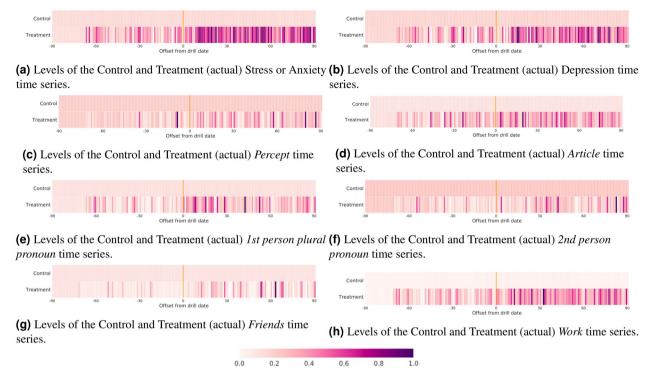


Fig. 3 Causality analysis comparing treatment and control trends for (a) Stress/Anxiety, (b) Depression, and the LIWC categories (c) Percept, (d) Article, (e) 1st person plural pronoun, (f) 2nd person plural pronoun, (g) Friends, and (h) Work. Figures show comparison between (1) normalized levels of mental health and psycholinguistic expressions around a 90-day period before and after school shooter drills (treatment); and (2) an equal duration control time series for the same mental health outcomes and psycholinguistic expressions, non-specific to the school districts and communities with the drills. Individual heatmap values are scaled between 0 and 1 for visual comparability purposes.

students, and parents, as well as school grade types— elementary, middle, and high. These results are outlined in Supplement S7 and S8. We find the general patterns observed for the aggregate to hold true in these subgroups, with high school communities and teachers showing the largest changes in terms of worsened mental health ( $^24.3-55.1\%$ ; p < 0.001 for anxiety/stress and depression) and heightened community outcomes ( $^5.9-27.2\%$ ; p < 0.001 for the LIWC categories).

# Discussion

**Interpreting the results**. The crisis literature posits that external reality, for example, traumatic events, can have profound effects on an individual's psyche, and can be considered to be the cause of emotional upheaval and stress (Vernberg et al., 2008). For example, persons exposed to mass shootings frequently report sleep disruption, due to feelings of grief over loss and anxiety about disaster re-occurrence and ongoing threats or due to symptoms of depression or post-traumatic stress. Notably, when children experience trauma leading to chronic stress or fear in the early childhood years, they are more vulnerable to behaviors associated with anxiety, which can also inhibit their ability to engage in higher-level thinking (Mulvihill, 2005). While the school shooter drills are not designed to be crises or to cause trauma—on the contrary, they are intended to better prepare students, staff, and teachers to face a traumatizing event involving an actual shooter on campus—this research empirically reveals, for the first time in such a large and diverse sample the negative psychological impacts of currently unregulated drills on school

Our results on mental health symptoms expressed in social media indicate in terms of broader longitudinal trends, that trauma and collective worry experienced by school stakeholders, increased by 42% for anxiety/stress, and 39% for depression, following drills. This trend sustains at least 90 days following drills and spans across diverse school districts, drill tactics, and the times of the year when the drills were conducted. In addition, the absence of such a change in the control data, that spanned the same time period in non-school-specific communities, further indicates a causal relationship between the drills and the wellbeing outcomes; in other words, the observed changes in the treatment data are not a consequence of seasonal or concurrent events. The focus groups further contextualized the computational insights, with some students asking their teachers and parents if they could be exempt from participating in these programs in the future, and others internalizing the reality of drills. Furthermore, the lexical analyses of the social media posts associated with high anxiety/stress and depression as well as the psycholinguistic outcomes reveal shifts in cognitive capacity and vocabulary changes, as well as people's perception and thinking in the aftermath of the drills. Overall, this study echoes and adds rigorous and scalable statistical support to the lived experiences of many American students, parents, and teachers who have expressed reservations with the inconsistent, unregulated, and at times traumatizing nature of K-12 school drills (Hamblin, 2018). Notwithstanding these negative impacts, the finding around increased collective action and advocacy resonates with the trend whereby school shooting incidents and the fear surrounding them propels many to join the broader movement of gun violence prevention (Everytown, 2019).

Methodological and practical implications. Our work bears important methodological implications. As noted above, existing research on school shooter drills is limited and fraught with methodological challenges. Given the nature of active-shooter

incidents, it is nearly impossible to study the direct impacts of drills on safety during these events. Therefore, most researchers in the past have opted to observe compliance with emergency procedures during drills and self-reported perceptions of safety and anxiety immediately following them (Dickson and Vargo, 2017). Moreover, due to a lack of well documented information on such training programs in various campuses and to circumvent potential for harm in gathering direct self-reports, some scholars have used lab-based experimental study to expose students to recordings of a simulated incident on campus, and then measure change in their psychological response (Peterson et al., 2015). However, these approaches lack ecological validity and suffer from the problem of small, non-representative sample. Together, this has hampered the development of scientific consensus as findings often conflict, as described above. To this end, our observational study approach contributes rigorous, scalable, and comprehensive scientific evidence to this debate. In other words, in this work, we observed how various school communities responded to the drill events based on their social media feeds, without us controlling who undergoes or experiences a drill at a school. We leveraged social media data as a natural experiment to provide a temporal pre- and post- (drill) event comparison with respect to a community's psychological health. Since we focused on an intra-community comparison (the changes in well-being outcomes noted in the paper are actually the changes in the same set of individuals in the three months preceding and succeeding the drills) and a sizable pre-post drill comparison period (~3 months before the drills, representing pre-drill baseline levels of the well-being measures), our approach inherently accounts for implementation, socio-cultural-political, and geographic differences across schools. Moreover, since the drills considered in the 114 schools were distributed through the school year, albeit not uniformly, our approach allows us to rule out the effects of local, national, or global events, seasonal trends, that might impact people's mental health or psycholinguistic expressions on social media (see Fig. S6 for additional analyses that repeated our analysis of the trends of stress/anxiety and depression excluding schools who experienced a gun violence incident as well as excluding time periods when such an incident was reported at one of the 114 schools under consideration). Our work thus advances existing approaches used to study the potentially causal impacts of school shooter drills.

Altogether, our findings raise important questions regarding the unseen costs of these school safety strategies that so far have not been evidence-based. In our focus groups, while parents and teachers noted that school shootings are "not an if, but when" (P19) and many were "just waiting for the next one to happen," (T11) at the same time, they argued that a trade-off must be navigated between preparedness and the psychological ramifications of these programs. In view of the lasting impact on psychological well-being we observed in our data, alternative, proactive school safety measures should be considered. These alternatives reduce the risk of harmful impacts and are supported by stronger evidence on their effectiveness. These include placing priority on school climate—creating a trusting school environment where all students are treated equally and bullying is not tolerated at any level. Research has shown that school climate is a strong predictor of school violence (Hurford et al., 2010). A second important proactive measure is student access to quality mental health resources, particularly for students in crisis or who are encountering social or emotional difficulties (NASP, 2021). Finally, school safety experts strongly support the use of threat assessment programs, multidisciplinary teams that intervene early when threats of violence come to light in a school. In 2018, the US Department of Homeland Security concluded that addressing these threats and behavioral issues are more effective than

physical plant improvements for school safety (US Department of Homeland Security, 2018).

Schools that do choose to implement shooter drills as their evidence-base continues to be assessed should prioritize trauma-informed trainings and excluding students, given their unique developmental stage. A 2020 report by Everytown for Gun Safety, American Federation of Teachers, and the National Education Association outlines several recommendations for drills, should they be implemented, which may help to ensure that they are trauma-informed and less detrimental to mental health (Everytown for Gun Safety, 2020). These include strategies like notifying school communities in advance; avoiding realistic simulations; ensuring developmental appropriateness; consulting with mental health professionals; and tracking data on their efficacy and effects.

Limitations, future research, and conclusion. Our study has limitations some which also open up opportunities for future directions. Although our classifiers of mental health symptomatic outcomes were built with expert appraisal and clinically validated instruments, we caution against making clinical inferences. As with other observational studies, we recognize that we cannot infer true causality (Watts, 2014). However, our work includes statistical rigor in minimizing various confounds, including additional triangulation from interviewing focus group participants; therefore the work provides insights beyond correlational analyses. Further, albeit longitudinal, our work only looks at the effects at a limited period of time (3 months). Future studies that include multiple follow-ups would allow for an enhanced understanding of the processes that lead to the observed trends and potential for the symptoms to turn chronic over time. In addition, to the best of our knowledge, the majority of school shooter drills are conducted on a quarterly to annual basis (Education Commission of the States, 2019). Therefore, drills occurring at least 90 days apart—not the whole 6 months—did not overlap with the pre- or post-periods and thus were unlikely to impact results. Since also we offset every day in comparison to the drill date (that is, the drill dates were spread across various times of the year), the probability of having multiple drills happen at the same school more than once in the period under investigation is low, allowing us to further isolate the effect of a single drill. That said, future work can explore additional causal inference techniques, such as that involving alternate hypothesis testing, to tease out situations where multiple drills may happen in quick succession.

We also acknowledge that our quantitative analyses may suffer from self-selection bias, as it is restricted to only those who are on social media and choose to post on it (Olteanu et al., 2019). Such data likely leaves out particularly young students at the schools where these drills were conducted, because Twitter's Terms of Service prevents individuals under the age of 13 to have accounts on the platform. Students' perspectives were somewhat supplemented with the focus groups as well as limited Twitter data from older students; however deeper analysis, with careful ethical considerations (Anderson, 2005; Farrell, 2005) can involve younger children more actively. On a related note, recall that our study focused on all individuals on social media potentially connected to the 114 school communities that experienced a drill during the study period, not just the direct victims. Parents were intentionally included in this sample, as research and anecdotal evidence suggests that they feel pressured to allow their children to participate in drills, are often the adults that must help them cope with any emotions in the aftermath, and have their own fears and anxieties around how to teach their children about gun threats at school (Kubicek et al., 2008). Moreover, the broader sample we focused on is likely to involve other family members and caregivers of students, teachers, and staff, or bystanders of school communities who may be indirectly impacted by the drills or have an interest in the topic. Crisis psychology literature advocates

considering both the direct and indirect victims of crisis events, to comprehensively understand the impacts in the aftermath (Beaton and Murphy, 1995; Harvey, 1998). Future research may expand on the diversity of psychological impacts in people who experience the drills directly as well as indirectly. Finally, our survey did not collect information about the type nor implementation of the school shooter drill. Future research can assess the differential impact of various protocols and their differing implementations to understand which, if any, pose fewer mental health risks.

In conclusion, by analyzing social media data in this study, we were able to quantify hard-to-observe impacts of school shooter drills; assess whether they are sustained over time and in comparison to suitable controls; and draw conclusions that uphold across different schools, drills, geography, communities, and time of the year. We provide the first empirical evidence that school shooter drills—in their current, unregulated statenegatively impact the psychological well-being of entire school communities, indicating that those who are affected are in need of continued support to process their aftermath, and that school systems need to rethink the design and utility of these approaches, against alternative gun violence prevention measures.

# **Data availability**

We used public online data (Twitter, Reddit), whose sharing in the raw form is restricted by the official API policies of the respective platforms. In addition, the paper also uses information on active-shooter drills gathered through a survey and focus group insights from school stakeholders. Deidentified version of this data can be requested by interested parties through appropriate data use agreements.

# **Code availability**

Custom code that supports the findings of this study is available from the corresponding author upon reasonable request.

Received: 20 October 2020; Accepted: 11 November 2021; Published online: 08 December 2021

# References

- ALICE (2021) ALICE's Standard of Care (K-12 Schools). https:// www.alicetraining.com/our-program/alice-training/k12-education/. [Online; accessed 28-May-2021]
- Anagnostopoulos A, Kumar R, Mahdian M (2008) Influence and correlation in social networks. In: Proceedings of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining. ACM. pp. 7–15
- Anderson WA (2005) Bringing children into focus on the social science disaster research agenda. Int J Mass Emerg Disasters 23(3):159
- Bagalman E, Caldwell SW, Finklea KM, McCallion G (2013) Public mass shootings in the united states: Selected implications for federal public health and safety policy. Congressional Research Service
- Bagroy S, Kumaraguru P, De Choudhury M (2017) A social media based index of mental well-being in college campuses. In Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems, ACM, pp. 1634–1646
- Beaton RDMurphy SA,(1995) Working with people in crisis: Research implications. In: Compassion fatigue: coping with secondary traumatic stress disorder in those who treat the traumatized vol. 1. Brunner/Routledge, pp. 1–51
- order in those who treat the traumatized. vol. 1. Brunner/Routledge, pp. 1–51
  Bernal JL, Cummins S, Gasparrini A (2017) Interrupted time series regression for the
  evaluation of public health interventions: a tutorial. Int J Epidemiol 46(1):348–355
- Bonanno CM, Levenson Jr RL (2014) School shooters: History, current theoretical and empirical findings, and strategies for prevention. Sage Open 4(1):2158244014525425
- Boyd D, Crawford K (2012) Critical questions for big data: Provocations for a cultural, technological, and scholarly phenomenon. Inform Commun Soc 15(5):662–679
- Chandrasekharan E, Pavalanathan U, Srinivasan A, Glynn A, Eisenstein J, Gilbert E (2017) You can't stay here: the efficacy of reddit's 2015 ban examined through hate speech. Proceedings of the ACM on Human-Computer Interaction, 1(CSCW), pp. 1–22

- Charmaz K (2014) Constructing grounded theory. Sage
- Cohen S, Wills TA (1985) Stress, social support, and the buffering hypothesis. Psychol Bull 98(2):310
- Cohn MA, Mehl MR, Pennebaker JW (2004) Linguistic markers of psychological change surrounding september 11, 2001. Psychol Sci 15(10):687–693
- Cousens S, Hargreaves J, Bonell C, Armstrong B, Thomas J, Kirkwood B, Hayes R (2011) Alternatives to randomisation in the evaluation of public-health interventions: statistical analysis and causal inference. J Epidemiol Commun Health 65(7):576–581
- Das Swain V, Saha K, Reddy MD, Rajvanshy H, Abowd GD, De Choudhury M (2020) Modeling organizational culture with workplace experiences shared on glassdoor. In: CHI. ACM
- De Choudhury M, Kiciman E (2017) The language of social support in social media and its effect on suicidal ideation risk. In: Proceedings of the International AAAI Conference on Web and Social Media, AAAI, vol 11, no 1
- De Choudhury M, Monroy-Hernandez A, Mark G (2014) "narco" emotions: affect and desensitization in social media during the mexican drug war. In Proceedings of the SIGCHI conference on Human Factors in Computing Systems, ACM, pp. 3563–3572
- Dickson MJ, Vargo KK (2017) Training kindergarten students lockdown drill procedures using behavioral skills training. J Appl Behav Anal 50(2): 407–412
- Diliberti M, Jackson M, Kemp J (2017). Crime, violence, discipline, and safety in us public schools: Findings from the school survey on crime and safety: 2015-16. first look. National Center for Education Statistics
- Education Commission of the States (2019) 50 State Comparison | K-12 School Safety. https://reports.ecs.org/comparisons/k-12-school-safety-03. [Online; accessed 28-May-2021]
- Eisenstein J, Ahmed A, Xing EP (2011) Sparse additive generative models of text. In Proceedings of the 28th International Conference on Machine Learning, ICML, pp. 1041–1048
- Everytown (2019) Gunfire on school grounds in the united states. Everytown for Gun Safety. https://maps.everytownresearch.org/gunfire-in-school/#16734
- Everytown for Gun Safety (2020) Keeping Our Schools Safe | A Plan for Preventing Mass Shootings and Ending All Gun Violence in American Schools. https://everytownresearch.org/report/preventing-gun-violence-in-american-schools/. [Online; accessed 28-May-2021]
- Farrell A (2005) Ethical research with children. McGraw-Hill Education (UK)Fine GA (2006) Ground truth: Verification games in operational meteorology. JContemp Ethnogr 35(1):3–23
- Glasgow K, Fink C, Boyd-Graber JL (2014) "our grief is unspeakable": Automatically measuring the community impact of a tragedy. In: ICWSM
- Golder SA, Macy MW (2011) Diurnal and seasonal mood vary with work, sleep, and daylength across diverse cultures. Science 333(6051):1878–1881
- Goodman-Scott E, Eckhoff A (2020) School counselors' experiences with lockdown drills: A phenomenological investigation. J Counsel Dev 98(4):435–445
- Grabow C, Rose L (2018) The us has had 57 times as many school shootings as the other major industrialized nations combined. CNN. https://www.cnn.com/2018/05/21/us/school-shooting-us-versus-world-trnd/ index.html
- Grinshteyn E, Hemenway D (2019) Violent death rates in the us compared to those of the other high-income countries, 2015. Prevent Med 123:123–20
- Gubiotti M (2015) Opposing viewpoints: preparing students, teachers, and the community for school shootings: Saving lives with active shooter simulations. Child Legal Right J 35(3):254
- Gun Violence Archive (2020) General methodology. https://www.gunviolencearchive. org/methodology. [Online; accessed 28-May-2021]
- Hamblin J (2018) What are active-shooter drills doing to kids? The Atlantic. https://www.theatlantic.com/health/archive/2018/02/effects-of-active-shooter/554150/
- Harvey MR (1998) Crisis assessment and interventions with victims of violence Hoffman Y (2018) Who is afraid of isis? isis anxiety and its correlates. Stress Health 34(1):84–92
- Hunt SD (1991) Modern marketing theory: critical issues in the philosophy of marketing science. South-Western Pub
- Hurford DP, Lindskog R, Cole A, Jackson R, Thomasson S, Wade A (2010) The role of school climate in school violence: a validity study of a web-based school violence survey. J Educ Res Policy Stud 10(1):51–77
- Hussein A (2009) The use of triangulation in social sciences research: Can qualitative and quantitative methods be combined. J Comp Soc Work 1(8):1–12
- Hyndman RJ, Athanasopoulos G (2015). 8.9 seasonal arima models. Forecasting: principles and practice. oTexts. Retrieved, 19
- Jones NM, Thompson RR, Schetter CD, Silver RC (2017) Distress and rumor exposure on social media during a campus lockdown. Proc Natl Acad Sci USA 114(44):11663–11668
- Jonson CL, Moon MM, Hendry JA (2020) One size does not fit all: Traditional lockdown versus multioption responses to school shootings. J School Viol 19(2):154–166
- Kícíman E, Counts S, Gasser M (2018) Using longitudinal social media analysis to understand the effects of early college alcohol use. In ICWSM, pp. 171–180

- Kiciman E, Sharma A (2019) Causal inference and counterfactual reasoning (3hr tutorial). In Proceedings of the Twelfth ACM International Conference on Web Search and Data Mining, ACM, pp. 828–829
- Kimmons R, Carpenter JP, Veletsianos G, Krutka DG (2018) Mining social media divides: an analysis of k-12 us school uses of twitter. Learn Media Technol 43(3):307–325
- King S, Bracy NL (2019) School security in the post-columbine era: Trends, consequences, and future directions. J Contempo Crim Just 35(3):274–295
- Kubicek K, Ramirez M, Limbos MA, Iverson E (2008) Knowledge and behaviors of parents in planning for and dealing with emergencies. J Commun health 33(3):158–168
- Lazer D, Pentland A, Adamic L, Aral S, Barabasi A-L, Brewer D, Christakis N, Contractor N, Fowler J, Gutmann M et al. (2009) Social science. computational social science. Science (New York, NY) 323(5915):721–723
- Lazer DM, Pentland A, Watts DJ, Aral S, Athey S, Contractor N, Freelon D, Gonzalez-Bailon S, King G, Margetts H et al. (2020) Computational social science: obstacles and opportunities. Science 369(6507):1060–1062
- Lin Y-R, Margolin D (2014) The ripple of fear, sympathy and solidarity during the boston bombings. EPJ Data Science 3(1):31
- Linden A (2013) Assessing regression to the mean effects in health care initiatives. BMC Med Res Methodol 13(1):119
- Linden A (2018) A matching framework to improve causal inference in interrupted time-series analysis. J Eval Clin Pract 24(2):408–415
- Ludwig J (2017) Reducing gun violence in america. Proc Natl Acad Sci USA 114(46):12097–12099
- Madfis E (2016) "it's better to overreact": School officials' fear and perceived risk of rampage attacks and the criminalization of american public schools. Crit Criminol 24(1):39–55
- Mark G, Bagdouri M, Palen L, Martin J, Al-Ani B, Anderson K (2012) Blogs as a collective war diary. In Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work, ACM, pp. 37–46
- Mayring P (2004) Qualitative content analysis. A companion to qualitative research 1(2004):159–176
- McDonald N, Schoenebeck S, Forte A (2019) Reliability and inter-rater reliability in qualitative research: Norms and guidelines for cscw and hci practice. Proc ACM on Hum-Comput Interact 3(CSCW):1–23
- McDowall D, McCleary R, Bartos BJ (2019) Interrupted time series analysis.

  Oxford University Press
- McPherson M, Smith-Lovin L, Cook JM (2001) Birds of a feather: Homophily in social networks. Ann Rev Sociol 27(1):415–444
- Mulvihill D (2005) The health impact of childhood trauma: An interdisciplinary review, 1997-2003. Issue Compre Pediatr Nurs 28(2):115–136
- Musu-Gillette L, Zhang A, Wang K, Zhang J, Kemp J, Diliberti M, Oudekerk B (2018) Indicators of school crime and safety: 2017. National Center for Education Statistics, US Department of Education
- NASP (2021) Preventing mass violence requires access to mental health services and reduced inappropriate access to firearms. https://bit.ly/2SQRoAJ. [Online; accessed 28-May-2021]
- NASP and NASRO (2017). Best practice considerations for schools in active shooter and other armed assailant drills. National Association of School Psychologists
- Nekvasil EK, Cornell DG, Huang FL (2015) Prevalence and offense characteristics of multiple casualty homicides: Are schools at higher risk than other locations? Psychol Viol 5(3):236
- Olteanu A, Castillo C, Diaz F, Kiciman E (2019) Social data: Biases, methodological pitfalls, and ethical boundaries. Front Big Data 2:13
- Organization, W. H. et al. (2020) Coronavirus disease 2019 (covid-19): situation report, 72. WHO
- O'Donnell ML, Creamer M, Pattison P (2004) Posttraumatic stress disorder and depression following trauma: understanding comorbidity. Am J Psychiatr 161(8):1390–1396
- Palen L, Anderson KM (2016) Crisis informatics—new data for extraordinary times. Science 353(6296):224–225
- Pan SJ, Yang Q (2009) A survey on transfer learning. IEEE Trans Knowl Data Eng 22(10):1345–1359
- Pennebaker JW, Mehl MR, Niederhoffer KG (2003) Psychological aspects of natural language use: Our words, our selves. Ann Rev Psychol 54(1):547–577
- Peterson J, Sackrison E, Polland A (2015) Training students to respond to shootings on campus: Is it worth it? J Threat Assess Manag 2(2):127
- Peterson JK, Densley JA (2019) The violence project database of mass shootings in the united states, 1966–2019. Saint Paul, MN: The Violence Project
- Resnick S, Smith RN, Beard JH, Holena D, Reilly PM, Schwab CW, Seamon MJ (2017) Firearm deaths in america: can we learn from 462,000 lives lost? Ann Surgery 266(3):432
- Reuter C, Kaufhold M-A (2018) Fifteen years of social media in emergencies: a retrospective review and future directions for crisis informatics. J Conting Crisis Manag 26(1):41–57
- Rygg L (2015) School shooting simulations: At what point does preparation become more harmful than helpful. Child Legal Right J 35:215

- Saha K, De Choudhury M (2017) Modeling stress with social media around incidents of gun violence on college campuses. Proc. ACM Hum.-Comput. Interact., 1(CSCW)
- Saha K et al (2019a) Imputing missing social media data stream in multisensor studies of human behavior. In Proceedings of International Conference on Affective Computing and Intelligent Interaction (ACII 2019), IEEE, pp. 178–184
- Saha K, Sugar B, Torous J, Abrahao B, Kícíman E, De Choudhury M (2019b) A social media study on the effects of psychiatric medication use. Proc Int AAAI Conf Web Soc Media 13(01):440–451
- Saha K, Weber I, De Choudhury M (2018) A social media based examination of the effects of counseling recommendations after student deaths on college campuses. In: ICWSM
- Schildkraut J, Nickerson AB (2020) Ready to respond: effects of lockdown drills and training on school emergency preparedness. Victim Offender 15(5):619–638
- Schildkraut J, Nickerson AB, Ristoff T (2020) Locks, lights, out of sight: Assessing students' perceptions of emergency preparedness across multiple lockdown drills. J School Viol 19(1):93–106
- Stuart F, Riley A, Pourreza H (2020) A human-machine partnered approach for identifying social media signals of elevated traumatic grief in chicago gang territories. PLoS ONE 15(7):e0236625
- Sullivan CF (1996) Recipients' perceptions of support attempts across various stressful life events. Commun Res Rep 13(2):183–190
- Tausczik YR, Pennebaker JW (2010) The psychological meaning of words: LIWC and computerized text analysis methods. J Lang Soc Psychol 29(1):24–54
- Terrades V, Khan SK (2018) Will it ever end: Preventing mass shootings in florida & the us. Suffolk Univ Law Rev 51:505
- Tian Y, Chunara R (2020) Quasi-experimental designs for assessing response on social media to policy changes. In Proceedings of the International AAAI Conference on Web and Social Media, AAAI, vol 14, pp. 671–682
- US Department of Homeland Security (2018) K-12 school security: a guide for preventing and protecting against gun violence. https://www.cisa.gov/sites/ default/files/publications/K12-School-Security-Guide-2nd-Edition-508.pdf. [Online; accessed 28-May-2021]
- Vernberg EM, Steinberg AM, Jacobs AK, Brymer MJ, Watson PJ, Osofsky JD, Layne CM, Pynoos RS, Ruzek JI (2008) Innovations in disaster mental health: Psychological first aid. Profess Psychol: Res Pract 39(4):381
- Watts DJ (2014) Common sense and sociological explanations. Am J Sociol 120(2):313–351
- Williamson E (2019) When active-shooter drills scare the children they hope to protect. New York Times. https://www.nytimes.com/2019/09/04/us/politics/ active-shooter-drills-schools.html.
- Zhe EJ, Nickerson AB (2007) Effects of an intruder crisis drill on children's knowledge, anxiety, and perceptions of school safety. School Psychol Rev 36(3):501–508

#### **Acknowledgements**

This research was supported through a grant from Everytown for Gun Safety to Georgia Institute of Technology; PI: Munmun De Choudhury.

# **Author contributions**

MOT, SBS, and MDC designed the research; ME, KS, and MDC conceptualized and developed the analytic techniques; ME, KS, PG, SM, JX, MOT, SBS, and JS gathered and analyzed the data; ME and MDC interpreted the results; ME, KS, and MDC drafted the paper; and MOT and SBS read, edited, and provided feedback on the paper.

# **Competing interests**

The authors declare no competing interests.

# Ethics approval

This focus group study was approved by the Institutional Review Board at the lead authors' institution, Georgia Tech. As noted in the manuscript, the focus group interview study constituted human subjects research, for which IRB approval was obtained at the lead author's institution, Georgia Tech, and participants were recruited following informed consent. Since historical, publicly accessible social media data from Twitter and Reddit were used for the computational aspects of this research and it involved no interaction with the authors of the social media postings, this part did not constitute human subjects research and therefore ethics approval was not required

#### **Informed consent**

Eligible individuals 18 years or older signed a consent form if they agreed to participate, while those under the age of 18 but above 15 provided their assent with consent from their legal guardian. All participants were provided with a plain language statement of the study and approved personnel—three of the paper's authors—conducted the interviews. As the drills could be emotionally triggering or distressing to some, especially those with first-hand experiences of a mass/school shooting, the consent form provided pointers to help, support, and counseling resources and participants were told that they could stop participating at any time.

# **Additional information**

**Supplementary information** The online version contains supplementary material available at https://doi.org/10.1057/s41599-021-00993-6.

Correspondence and requests for materials should be addressed to Munmun De Choudhury.

Reprints and permission information is available at http://www.nature.com/reprints

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, dis-

tribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <a href="http://creativecommons.org/licenses/by/4.0/">http://creativecommons.org/licenses/by/4.0/</a>.

© The Author(s) 2021