Journal of Experimental Criminology (2019) 15:551–567
https://doi.org/10.1007/s11292-019-09380-3

Policing people with mental illness: experimental evaluation of online training to de-escalate mental health crises

Robyn L. Hacker[1] · John J. Horan[2]

Published online: 7 September 2019

Abstract

Objectives This study was conducted to complete a proof of concept for a brief online training designed to improve the policing of people with mental illness. The training, positioned within a stress inoculation framework, is scalable worldwide at minimal cost. Our primary intent was to effect improvements in law enforcement officers' ability to effectively respond to and de-escalate mental health crises.

Method Participants were randomly assigned to either DEFUSE, the online de-escalation training, or a delayed treatment control condition. DEFUSE was evaluated with the full array of measures also used to assess CIT, the most comprehensive mental health training available to law enforcement in the USA. Additionally, DEFUSE was evaluated using a randomized controlled trial design, including additional measures assessing performance competence, knowledge, and satisfaction with the training.

Results Participants initially receiving DEFUSE showed significant improvement in performance competence, assessed from their responses to simulated mental health crises. Multivariate and univariate analyses indicated that DEFUSE produced significant beneficial effects on all CIT measures (empathy, stigma, self-efficacy, and behavioral self-report) with strong effect sizes; control participants obtained comparable benefits after DEFUSE training.

Conclusions DEFUSE appears to be a promising tool for teaching law enforcement officers how to intervene in mental health crises. It provides consistent training at the convenience of the learner and is engaging, time and cost effective, and easily scalable worldwide. Furthermore, participants found it interesting and the skills easy to understand and remember. Future studies will evaluate DEFUSE with a larger law enforcement sample and consider the possibility of collecting longitudinal data on level of force and resolution, referral, and arrest.

Keywords De-escalation · Training · Law enforcement · Police officers · Mental illness · Mental health crises

The commonly accepted roles of law enforcement are to serve and protect. These generic responsibilities include a wide array of specific duties such as preventing dangerous driving, solving heinous crimes, halting violence, and suppressing the sale

of narcotics. Identifying individuals with mental illness has historically not been included in these duties, yet people involved with the criminal justice system often have comorbid mental illness.

Indeed, police contact involving people with mental illness regularly occurs as a result of professional regulation as well as chance. Mental health providers are mandated by law and required by their ethics codes (e.g., American Psychological Association, American Counseling Association, American Psychiatric Association) to contact law enforcement when a patient reports a plan to harm him or herself or someone else. Historically, 7% of all police contacts in US cities with more than 100,000 people have involved an individual with mental illness (Deane et al. 1999). More recently, Tinney and Rosenbaum (2015) reported that mental illness was the primary reason for police contact in 33% of calls in the USA. Other surveys indicate that 60 to 90% of officers respond to at least one call per month involving a subject with mental illness (Borum et al. 1998; Gillig et al. 1990). Furthermore, officers respond to an average of six calls involving subjects with mental illness per month.

The most frequent scenarios involving these subjects, according to one US study (Peck 2003), include concerned persons (e.g., family member, friend, other) contacting the police during a psychiatric emergency, suicidal individuals calling the police for help, police encountering someone behaving inappropriately in public, police being contacted due to an individual feeling threatened by unusual behavior, and persons seeking help from the police due to imagined threats. The prevalence of calls involving mental illness persists internationally as well. For example, the College of Policing reports an increasing number of incidents involving individuals with mental illness in England and Wales and estimates that 15–20% of all police incidents are linked to mental illness (College of Policing 2015).

At a minimum, a person with mental illness is killed during 25% of all fatal police encounters. Studies both in the USA and internationally (i.e., Australia, UK, Canada) report even higher rates, closer to 50%. Furthermore, a person with untreated mental illness is 16 times more likely to be killed during a police incident (Fuller et al. 2015). These statistics are surely impacted by the fact that law enforcement is often the first professionals to have contact with individuals experiencing symptoms of mental illness and they are not adequately prepared for this task.

To date, law enforcement recruits rarely receive sufficient training on how to effectively manage subjects experiencing mental health crises. In the USA, for example, 45,000 recruits begin basic training each year in any of 664 different state and local academies monitored by the Bureau of Justice Statistics (BJS) of the U.S. Department of Justice (Reaves 2016). On average, recruits receive 21 weeks (840 h) of training before beginning patrol duties, an uptick of about 2 weeks since the last BJS report (Reaves 2009), but still conspicuously less than what is required for a barber (900 h) or a cosmetologist (1500 h). The 2009 BJS report further indicates that recruits receive an average of 123 hours of training dedicated to weapons/self-defense, strategies not conducive to effectively managing mental health symptoms, whereas the topic of mental illness was neither covered nor queried in the data collection survey. The 2016 BJS update indicated that 95% of academies now offer some training on mental illness; however, it is unclear what this means as only a single yes/no item was used to obtain this information. If given appropriate training, law enforcement could divert

individuals into treatment rather than the criminal justice system (Teller et al. 2006) and decrease the number of fatal police encounters.

Despite numerous training options being available worldwide to help law enforcement better understand mental illness—Crisis Intervention Team (CIT) Training in the USA (Compton 2008), Training and Education About Mental Illness for Police Organisations (TEMPO) in Canada (Coleman and Cotton 2010, 2014), the MHIT program in Australia (Herrington and Pope 2014), and the Connect Trial in England (Scantlebury et al. 2017)—it remains unclear how to best deliver such training. Training design varies and delivery of content is inconsistent. Additionally, the existing evidence evaluating these trainings for law enforcement is of such poor quality that conclusions on their effectiveness cannot be confidently made (Booth et al. 2017). Only three randomized controlled trials have been attempted (Rafacz 2012; Teagardin et al. 2012; Scantlebury et al. 2017), and none have looked at observable behavioral change in officer interactions.

In the USA, the most studied approach to training law enforcement on mental illness is the CIT or Memphis model (Dupont et al. 2007). Developed in Tennessee in 1988, the model spread rapidly and now has a presence in all but three USA states (http://cit.memphis.edu/citmap/), despite limited evidence of its effectiveness (Compton 2008; Booth et al. 2017). Furthermore, the full CIT model contains 10 different elements (see Dupont et al. 2007 for details of each element), one of which is a 40-hour comprehensive training for patrol officers that emphasizes mental–health–related topics, de-escalation strategies, and access to community-based services. This training incorporates didactics, exposure components, and scenario-based skills training and allows officers to take on a specialist role within their department. Police departments are generally responsible for underwriting costs of this training (El-Mallakh et al. 2014) and motivating patrol officers to participate (Compton et al. 2014a), after varying periods of service. Implementing CIT is resource intensive, training delivery is inconsistent, and its full implementation presents sustainability challenges, particularly for small and rural departments (Compton et al. 2010). Furthermore, it is rarely offered to all officers within a department.

DEFUSE was developed to remedy these concerns. It is an interactive online training program developed to specifically train law enforcement on mental illness and effective de-escalation strategies for handling mental health crises. DEFUSE is organized within a classic stress inoculation framework, offering education, skills training, and rehearsal (Meichenbaum and Deffenbacher 1988). This framework has previously been applied and shown successful with law students and nurses (Sheehy and Horan 2004; West et al. 1984). Classic law school curriculum is reportedly one of the most stressful of all graduate training programs, and acute care nurses routinely face horrific injuries, perform CPR, and must tell parents their child has died. Law enforcement officers face similar daily stressors but also the ever-present reality of possible injury or death. Figure 1 provides an overview of DEFUSE organized within the stress inoculation framework. It further depicts a logic map linking the components of stress inoculation, to the learning objectives of DEFUSE and the dependent variables assessed in this study. The components (education, skills training, and rehearsal) are delivered in a consistently individualized, interactive manner, and the online format ensures that DEFUSE can scale to virtually every law enforcement officer, not only in the USA but worldwide.

We conducted a proof of concept using a randomized controlled trial of DEFUSE. Our primary intent was to effect officers' ability to respond to crisis situations involving

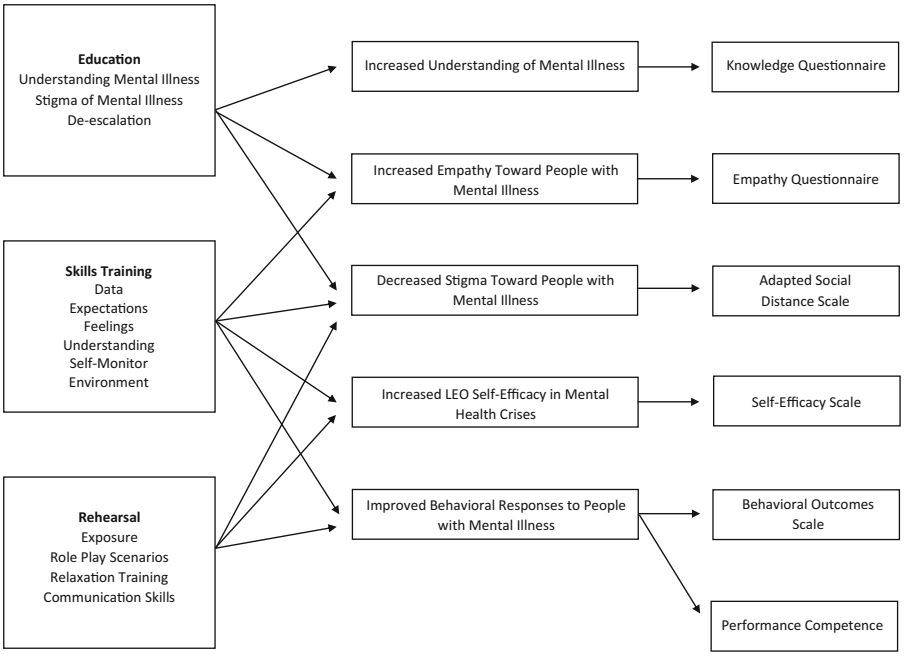

Fig. 1 A priori logical relationships among stress inoculation components, DEFUSE training objectives, and dependent variables

individuals with mental illness. These evaluative data were derived from videotapes of participant responses to simulated mental health crises. Similar to evaluators of CIT (Compton et al. 2014a), we were interested in whether DEFUSE would increase officers' knowledge of mental illness, empathy toward individuals with mental illness, self-efficacy in responding to calls involving mental health crises, and behavioral self-report while decreasing their stigma toward mental illness in general. We supplemented the battery used by these researchers with additional measures tapping specialized knowledge unique to DEFUSE and satisfaction with the training. Furthermore, our evaluation design included a built-in replication, namely a delayed treatment control condition in which participants eventually received the DEFUSE training and an attenuated assessment battery. We hypothesized improvements on all measures for all participants following training with DEFUSE. Although this study did not seek to assess DEFUSE's impact on level of force and resolution, referral, or arrest, research indicates that training on mental illness has led to increases in the use of verbal de-escalation strategies (Compton et al. 2014b) and referral of subjects with mental illness to treatment rather than arrest (Compton et al. 2008, 2014b; Scantlebury et al. 2017).

Method

Participants

For this proof of concept, a sample of 24 volunteers from 11 states of the USA (7 women and 17 men) were randomly assigned to either DEFUSE training (*M* age =

34.92) or the delayed treatment control (M age = 39.92) conditions; each group contained 12 volunteers. Participants ranged in age from 23 to 60, and 87.5% of the sample self-identified as Caucasian. Exclusionary criteria included being under 21, CIT trained, or a mental health professional. Twenty-one percent of the participants worked in law enforcement with experience ranging from 3 to 20 years. See Table 1 for additional demographic information by participant treatment condition.

Recruitment A priori power analyses conducted using G*Power indicated that a sample of 72 participants would be sufficient to detect a small effect size (power = 0.80, alpha = 0.05, 2 groups). Participant recruitment occurred through numerous avenues over a 6-month period. Police academies and individual departments throughout the USA were directly contacted via email and postal mail. In response to these efforts, numerous departments and individual officers expressed interest in participating; however, departmental training structures, concerns about the technology involved, and bureaucratic policies often prevented participation. To advance the proof of concept, it was decided to open recruitment to non-law enforcement participants through online marketing efforts (e.g., websites, advertisements). In theory, DEFUSE would be offered to new officers early in their training; therefore, non-law enforcement participants would be similar to officers at this stage of training. Participants were offered a $25 Amazon Gift Card for their time.

Measures

Demographic information was collected via a questionnaire and used for sample description and screening purposes. All data were collected via Qualtrics (www.qualtrics.com), an online survey software program at pre- and post-testing, except for a satisfaction survey administered after completion of DEFUSE.

Performance competence Our primary objective, namely improving the de-escalation skills of participants, was assessed through videotaped responses to simulated stressful encounters with individuals displaying symptoms of mental illness. Six unique escalating scenarios including an infant death, a suicide threat, and an escalated subject with mental illness were selected from a video library of scenarios developed by The Behavioral Personnel Assessment Device Group (B-PAD 2014). The B-PAD group specializes in the development of video-testing technology that presents real-life situations to assist law enforcement agencies with hiring decisions; this technology captures the applicants' spontaneous behavioral response to selected scenarios and allows hiring managers to assess whether applicants perform desired competencies.

Different sets of three scenarios were streamed to study participants at pre- and post-testing through their computers; the same six scenarios were delivered in the same order to each participant due to technological limitations prohibiting counterbalancing. Participants were instructed to engage in role plays as if they were the officer on the scene, and their responses were captured by web cameras on their computers and uploaded for rating by trained independent masters-level judges, blind to treatment condition. Judges were selected for their willingness to be trained and complete the task of scoring each video on the presence or absence of 24 specific behaviors indicative of

DEFUSE competencies (e.g., attempt to gather data, introduction of self to subject, asking open-ended questions). Individual scores were the total number of skills displayed on each set of three pre- and post-scenarios; scores ranged from 0 to 24 for each video. Interrater reliabilities were strong, $k = 0.98$, $p < 0.000$.

CIT measures Secondarily, to allow for potential comparison, we were interested in assessing the same outcomes as researchers evaluating CIT. The CIT measures have all demonstrated good internal consistency and test–retest reliability as noted in prior studies (e.g., Bahora et al. 2008; Broussard et al. 2011; Compton et al. 2006, 2014a).

The *Empathy Questionnaire* is an 11-item measure adapted from Levy et al. (2002). Two items are rated on an 11-point Likert-type scale ranging from *Not at All* to *Definitely So.* These items ask participants whether they have ever imagined how people with mental illness feel about having a mental illness and whether they have ever considered how living with a mental illness would affect their own life. The remaining items ask participants to indicate how much they feel nine different emotions (e.g., compassion, hostility) in relation to people with mental illness. Each of these items is rated on an 11-point Likert-type scale ranging from *Not at All* to *Extremely.* Negative emotions (e.g., disgust, hostility, suspicion) were reverse scored; therefore, higher scores indicate greater ability to empathize with individuals experiencing symptoms of mental illness. Pre-test internal consistency was high at 0.879 as was test–retest reliability calculated on the controls, $r = 0.837$, $n = 12$, $p = 0.001$.

Participants responded to the remaining CIT measures, as described below, after reading a vignette scenario about someone with a serious mental illness. The vignette was developed by Broussard et al. (2011) and represented a real-world interaction commonly faced by on duty officers. The name of the individual in the vignette was changed to "Jordan," from "David," to make it more neutral in regard to gender and race.

The *Adapted Social Distance Scale* (*ASDS*), adapted from Bogardus (1925), is a 9-item self-report measure designed to assess participants' social distance, or stigma, toward individuals displaying symptoms of mental illness. Items are rated on a 4-point Likert-type scale ranging from *Very Unwilling* to *Very Willing*. Scores range from 9 to 36 with higher scores indicative of lower stigma or less social distance. Sample items include *Six months from now, when Jordan is not in crisis, how willing would you be to sit next to him/her on the bus?* and *Six months from now, when Jordan is not in crisis, how willing would you be to rent an apartment in your basement to him/her?* Pre-test internal consistency for the ASDS was high at 0.846; test–retest reliability calculated on the control subjects was also good, $r = 0.746$, $n = 12$, $p = 0.005$.

The *Self-Efficacy Scale* (*SES*), modified from Bahora et al. (2008), is a 16-item measure designed to assess participants' perceived ability to handle interactions with someone exhibiting symptoms of mental illness. Items are rated on a 4-point Likert-type scale ranging from *Not at all Confident* to *Very Confident*; scores range from 16 to 64 with higher scores indicative of greater confidence in interactions with someone displaying symptoms of mental illness. Sample items include *How confident would you feel in your ability to effectively communicate with someone like Jordan?* and *How confident would you feel in your ability to effectively de-escalate a mental health crisis involving someone like Jordan?* Pre-test internal consistency for the SES was excellent at 0.916 as was test–retest reliability calculated on the control subjects, $r = 0.905$, $n = 12$, $p = 0.000$.

The *Behavioral Outcomes Scale* (*BOS*) is a 16-item self-report measure developed to assess de-escalation and referral decisions or reported behavioral outcomes. Eight items correspond to each construct. Items are rated on a 4-point Likert-type scale ranging from *Very Negative* to *Very Positive*; eight items required reverse scoring. Scores range from 4 to 64 with higher scores indicative of good de-escalation skills and positive referral decisions. Sample items include *Having your hand on your baton or gun when speaking with Jordan* and *Contacting a mobile crisis unit to take Jordan to a mental health facility.* Consistent with prior studies (Broussard et al. 2011), reliabilities were moderate on this measure; pre-test internal consistency for the full BOS was 0.627, and test–retest reliability calculated on the control participants was $r = 0.695$, $n = 12$, $p = 0.012$.

Supplemental DEFUSE measures In addition to performance competence, the primary outcome criterion of this investigation, two self-report measures were relevant to DEFUSE. *Knowledge* proffered by the DEFUSE program was assessed by an 18-item true/false test. Eight of the items assessed information derived from the module on symptoms of mental illness and recovery; the other 10 items assessed six de-escalation skills taught in the second module of the training. Scores ranged from 0 to 18 with higher scores indicating greater knowledge of the measured content. Test–retest reliability calculated for the no-treatment control subjects was moderate, $r = 0.629$, $n = 12$, $p = 0.028$.

Finally, a *Satisfaction Survey* was used to capture feedback about the participants' experience completing the DEFUSE training course; it was administered following completion of the online training. The survey contained five 4-point Likert-type items, ranging from *Very Dissatisfied* to *Very Satisfied*, asking about the ease of navigating the program, the program's ability to maintain interest, the breadth of the curriculum, the simplicity of the acronym, and overall satisfaction. Two open-ended qualitative items asked what participants specifically liked about the program and how the program could be improved.

Procedures

After completion of the pre-test battery, participants were randomly assigned to either the delayed treatment control or DEFUSE condition. Recruitment of participants to the study was gradual, permitting ongoing monitoring of potential outcome effects. When it became apparent that DEFUSE was achieving success, recruitment ceased due to time and funding constraints.

Delayed treatment control condition Approximately 7 days after pre-testing, the 17 participants in the control condition received the post-test measures, excluding the satisfaction survey. Following the completion of these measures, control participants were provided a personalized login and password to access the DEFUSE course, permitting a quasi-experimental replication of effects attributable to DEFUSE. Immediately following completion of DEFUSE, a final post-test battery, including the satisfaction survey but excluding the performance competence measure, was presented. A third administration of the performance competence measure was prohibited by cost constraints and technological limitations.

DEFUSE condition The 15 participants in the DEFUSE condition received a personalized login and password to access the DEFUSE course approximately 7 days after pretesting. Following the completion of the course, they were immediately presented with the post-test battery, including the satisfaction survey.

DEFUSE is a 2-hour, interactive online training that contains two modules, developed in consultation with multiple law enforcement officers: MENTAL ILLNESS AND DE-ESCALATION SKILLS. The mental illness module opens with education about the topic and attempts to break down negative schema people often have about mental illness and people who suffer from it. Mental illness in general and mental illness as the result of the stresses inherent in law enforcement are covered. This portion of the training underscores the fact that recovery can occur with proper treatment, reminds officers that they are often the first professionals to have contact with individuals exhibiting acute symptomology, and reinforces that officers are not immune to mental illness. Moreover, it introduces the critical role law enforcement has in diverting people from the criminal justice system into treatment. These topics are covered through didactics, video, and graphic portrayals of real people living with mental illness. Research indicates that both education and contact have positive effects on reducing stigma for adults (Corrigan et al. 2012) and increasing empathy (Kalisch 1971).

The next portion of the mental health module educates learners about six common symptoms of mental illness: sadness, anxiety, anger, mania, delusions, and hallucinations. Being able to distinguish between various mental illnesses is beyond the scope of this introductory course and beyond the role of law enforcement; diagnosis and treatment are advanced skills of mental health providers. Law enforcement officers can, however, benefit greatly from being able to recognize symptoms commonly experienced by people suffering from a variety of mental illnesses. This knowledge can aid them in determining what skill set will work best to keep themselves safe and minimize fatal police encounters. Each symptom is introduced didactically based on DSM 5 criteria. Learners are then presented with a demonstrative video clip or activity; many of the videos show excerpts of law enforcement officers displaying the symptoms as portrayed in well-known television shows. The section on hallucinations requires the learner to complete a word search review of the learned symptoms while experiencing auditory hallucinations. A knowledge check follows the demonstrations, reinforcing the most critical information to retain.

The de-escalation module teaches learners six important skills: gather data and document, set expectations, figure out feelings of the subject, demonstrate understanding, self-monitor, and use the environment. Each of these skills is taught to entry-level mental health counselors as foundational to more advanced clinical intervention due to their effectiveness in managing mental health crises. The course name, DEFUSE, also serves as an acronym to aid in recall of each of these skills. The structure of this module facilitates increased officer self-efficacy. Learners are initially introduced to each skill through simple and clear language. Next, they observe *Officer Fuller* model the skill through interactions with *Mary*, a subject displaying delusional thinking and anger. The presented information is reviewed, and the learner is asked to explain the skill in his/her own words. Following this, learners have the opportunity to rehearse the skill by role-playing with *David*, a new subject displaying anger and sadness; suggestions are offered before transitioning to the subsequent skill. Self-monitoring is structured differently as the focus is on the learner being aware of his/her own body, emotions,

and limitations. Following psychoeducation on these topics, the learner is invited to practice deep breathing strategies, mindfulness, and meditation exercises, and shown how to use visualizations and positive self-talk. Throughout the de-escalation module, learners are reminded of the importance of staying calm and being patient, repeating their message, and not taking anything stated by a subject personally; this module has a heavy emphasis on effective communication strategies.

Results

Preliminary analyses

Pre-treatment analyses Univariate (ANOVA) and multivariate analyses of variance (MANOVA) were conducted on the pre-test scores of DEFUSE and control participants to determine whether any differences between groups appeared at pre-test. An ANOVA on total score for the simulations used to determine performance competence was insignificant $F(1, 10) = 0.001$, $p = 0.978$, partial $\eta^2 = 0.000$, as was a MANOVA on the battery of all CIT measures along with the supplemental DEFUSE knowledge scale, Wilk's $\lambda = 0.770$, $F(5, 18) = 1.076$, $p = 0.406$, partial $\eta^2 = 0.230$, indicating that random assignment resulted in no pre-treatment differences on any measure. Pre-test and post-test scores and confidence intervals for all participants on each measure are presented in Table 2.

Attrition Complete pre-/post-data were obtained from 22 participants on all self-report measures in the CIT and supplemental DEFUSE assessment batteries. Half of the participants provided data on the performance competence measure, whereas the remainder cited equipment failure and noncompliance due to privacy concerns. Two participants did not complete the satisfaction survey. Full details of participant allocation and attrition can be found in the CONSORT diagram in Fig. 2.

Power In addition to difference testing, effect sizes were calculated and are reported as partial η^2.

Benefits of DEFUSE on performance competence

A univariate treatment-by-repeated-measures ANOVA was conducted on the summed scores obtained from de-escalation simulations administered at pre- and post-testing. A significant interaction, $F(1, 10) = 5.358$, $p = 0.043$, partial $\eta^2 = 0.349$, indicated that DEFUSE produced improvement in performance competence of de-escalation skills.

Benefits of DEFUSE on CIT measures

A treatment-by-repeated-measures MANOVA was conducted on the pre- and post-test scores of the CIT battery to control for familywise error. The interaction was significant (Wilk's $\lambda = 0.290$, $F(4, 19) = 11.645$, $p = 0.000$, partial $\eta^2 = 0.710$) indicating beneficial

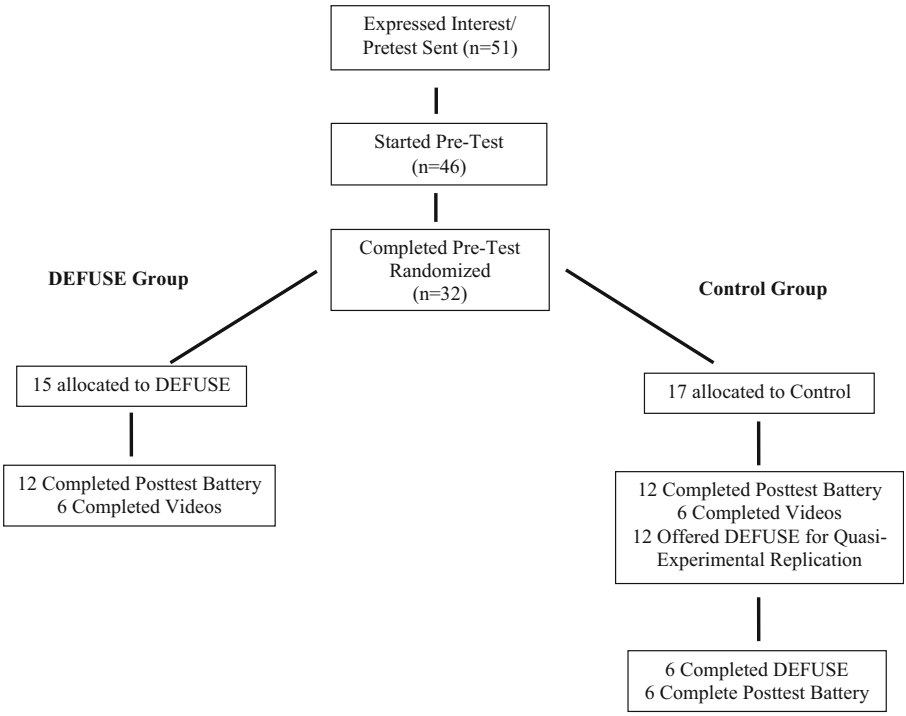

Fig. 2 CONSORT diagram. This figure shows treatment and control condition allocation

effects favoring DEFUSE. The interactions on follow-up repeated measure ANOVAs for each dependent variable were also significant, indicating that beneficial effects favoring DEFUSE were seen on each of the individual measures: *Empathy*, $F(1, 22) = 12.025$, $p = 0.002$, partial $\eta^2 = 0.353$; *Stigma*, $F(1, 22) = 10.464$, $p = 0.004$, partial $\eta^2 = 0.322$; *Self-efficacy*, $F(1, 22) = 29.295$, $p = 0.000$, partial $\eta^2 = 0.571$; *Self-reported behavior outcomes*, $F(1, 22) = 11.871$, $p = 0.002$, partial $\eta^2 = 0.350$. Neither the MANOVA treatment nor repeated measures main effects are relevant to the hypotheses. In sum, as a result of training, DEFUSE participants showed greater empathy toward individuals with mental illness, lower levels of stigma toward mental illness, improved self-efficacy for handling mental health crises, and better recognition of favorable behavioral outcomes.

Benefits of DEFUSE on supplemental DEFUSE measures

An additional follow-up to the MANOVA just reported indicated a significant univariate effect on the *Knowledge* measure, $F(1, 22) = 11.851$, $p = 0.002$, partial $\eta^2 = 0.350$, meaning that participants showed increased understanding of mental illness and de-escalation following completion of DEFUSE training.

The *Satisfaction Survey* data from all participants in the experimental and delayed treatment conditions who had received DEFUSE training were collapsed ($n = 22$). In general, participants were very happy with DEFUSE. Mean responses on each of the 4-

point Likert items ranged from 3.45 to 3.7. Participants were happiest with the breadth of the curriculum ($\bar{x} = 3.7$). Scores on the other rated dimensions were only slightly lower: maintaining interest ($\bar{x} = 3.45$), ease of navigation ($\bar{x} = 3.5$), the acronym DE-FUSE ($\bar{x} = 3.5$), and overall experience ($\bar{x} = 3.5$). Features most liked about the program were its structure and content. Suggestions for improvement most often called for additional video simulations for further skill rehearsal.

Interaction details

The interactions emerging from the aforementioned treatment-by-repeated-measure ANOVA and MANOVA analyses are necessary and sufficient to confirm the hypotheses for differential benefits favoring DEFUSE in comparison to controls. We have added pairwise *t* tests in Table 3 in response to an editorial request.

Participants given the DEFUSE training showed significant pre-/post-improvement on five of the six outcome measures and marginal gains on the sixth. Moreover, DEFUSE participants were superior to controls on four of the six measures at post-test. Finally, control participants showed pre-/post-gains on performance competence and deteriorated on empathy. This outcome pattern is consistent with the previously reported ANOVA and MANOVA interactions indicating differential benefits for DEFUSE participants.

Quasi-experimental replication analyses

A treatment-by-repeated-measures MANOVA was also conducted to control for familywise error in comparing the control group's post-test data with scores obtained following their completion of DEFUSE. It was not significant (Wilk's $\lambda = 0.026$, $F(5, 1) = 7.418$, $p = 0.271$, partial $\eta^2 = 0.974$); however, given the large effect size, we opted to review the univariate follow-ups as the small sample size would greatly limit the available power in finding a significant result. Univariate follow-ups were significant on four of the five measures: *Empathy*, $F(1, 5) = 20.298$, $p = 0.006$, partial $\eta^2 = 0.802$; *Self-efficacy*, $F(1, 5) = 11.468$, $p = 0.020$, partial $\eta^2 = 0.696$; *Reported behavior outcomes*, $F(1, 5) = 6.231$, $p = 0.055$, partial $\eta^2 = 0.555$; *Knowledge*, $F(1, 5) = 53.57$, $p = 0.001$, partial $\eta^2 = 0.915$. The effect on the *Stigma* measure was not significant: $F(1, 5) = 2.609$, $p = 0.167$, partial $\eta^2 = 0.343$. In sum, these replication analyses further indicate that DEFUSE can be an effective intervention for increasing empathy toward individuals with mental illness, improving self-efficacy in effectively handling mental health crises, raising awareness of behavioral outcomes to benefit such interactions, and favorably impacting participant understanding of mental illness and de-escalation (Table 4).

Discussion

Despite numerous options being available to train law enforcement worldwide on how to better recognize and intervene in mental health crises, current

research is lacking and few randomized controlled trials have been conducted. Training designs vary. Delivery of content is inconsistent. Access is limited, and it is unclear how to best distribute such training to law enforcement audiences. DEFUSE was developed and evaluated to offer a resolution to these concerns. The curriculum derives from a stress inoculation framework, requires only 2 hours of interactive online training, and is consistently scalable worldwide at negligible cost.

We randomly assigned participants to either the DEFUSE or delayed treatment control condition, later providing control participants with DEFUSE training. Those initially receiving DEFUSE showed statistically significant improvements in performance competence assessed from their behavioral responses to video-simulated mental health crises. For comparative purposes, we also evaluated DEFUSE with a full array of measures from CIT research, plus additional measures tapping specialized knowledge and satisfaction with the DEFUSE training. Multivariate and univariate analyses indicated that DEFUSE produced significant beneficial effects on all CIT measures (empathy, stigma, self-efficacy, and behavioral self-report) as well as the knowledge and satisfaction outcomes targeted by DEFUSE; despite limited power during quasi-experimental replication, control participants obtained comparable benefits after DEFUSE training on four of the five measures.

DEFUSE thus appears to be a promising tool for teaching law enforcement how to intervene in mental health crises. It is important to note that in addition to achieving statistical significance, DEFUSE produced strong effect sizes on each of the CIT measures. Conversely, research conducted on CIT has found moderate effects, at best, on the same measures (Compton et al. 2014a). DEFUSE is far less costly, consistently delivered, less time intensive, and easily scalable to officers worldwide. Using the same numbers as El-Mallakh et al. (2014), DEFUSE could be used to train a recruit for less than 10% of the personnel cost of CIT; officers only need access to a computer with Internet capabilities.

Despite the limitations imposed by a small *n*, the DEFUSE training had a significant impact on a wide array of measures. Nevertheless, for these results to become more meaningful and convincing, replications with larger sample sizes are necessary. Additionally, participation was not limited solely to law enforcement making generalization to the broader law enforcement population more challenging. Recruitment of law enforcement officers was complicated by department structures and individual concern about privacy related to being recorded through webcam technology. Additional information about noncompletion of the study was not solicited and participants self-selected for participation. As a result, selection bias is likely inherent. Interestingly, Compton et al.(2011) found that officers voluntarily completing CIT training were neither more psychologically minded nor more empathic than non-trained officers or officers mandated to complete the training. Voluntary officers, however, were found to have greater contact with mental health professionals.

Although the performance competence measure unique to this study allowed for the assessment of participant behavior during simulated interactions, no follow-up occurred to assess participant behavior outside of the classroom setting or to re-evaluate these outcomes over time. Despite this, two of the

law enforcement participants sent unsolicited emails months after their completion of the course noting a call in which they used the DEFUSE skills and the benefit that came from this.

Given these limitations, future studies will seek to obtain collaboration with law enforcement agencies for recruitment. This will provide a larger sample of the population of interest and limit selection bias allowing for greater generalizability of study findings. Longitudinal data and evaluation of DEFUSE's impact on level of force and resolution, referral, and arrest could be gathered more easily through this structure as well. In regard to DEFUSE, future versions will evolve to incorporate additional content recommendations received from participants. We also expect to develop an array of video simulations to further enhance the learners' experience. Although the B-PAD group was generous in making the incorporation of the performance competence measure possible, cost per participant is an impediment to extensive research and development activity. Moreover, creation of videos would permit closer alignment with instruction materials and a more sensitive outcome battery. Similarly, the creation of videos would allow us to embed additional real-life simulations within the course itself, rather than just at pre- and post-assessment points; doing so would offer more opportunity for practice, as requested by participant feedback, while enhancing the learning experience.

The DEFUSE model could also be easily extended and applied to specialized clinical issues including trauma, addiction, and developmental disability, or customized for different occupations within law enforcement such as dispatchers, correctional guards, and parole officers. Certainly, DEFUSE has strong potential for enhancing law enforcement's capacity to serve and protect.

Funding information This research was supported in part by a grant from the American Psychological-Law Society Grant-in-Aid program and by a grant from the Graduate and Professional Student Association at Arizona State University. B-PAD also provided assessment resources at a discounted rate.

Appendix

Table 1 Basic sociodemographic characteristics of the study sample ($n = 24$)

	Control	Experimental
Age, Years	Range, 26–60 (M age = 39.92)	Range, 23–56 (M age = 34.92)
Educational attainment	4 had less than a bachelor's degree 8 had a bachelor's degree or higher	3 had less than a bachelor's degree 9 had a bachelor's degree or higher
Gender	9 male 3 female	8 male 4 female
Race	10 Caucasian	12 Caucasian
Law enforcement experience	2 in law enforcement	4 in law enforcement
Familiarity with mental illness	10 know someone with mental illness 8 know someone in treatment	9 know someone with mental illness 7 know someone in treatment

Table 2 Study 1: means, standard deviations, and confidence intervals for DEFUSE and control conditions by testing occasion

	Pre-test		Confidence Intervals		Post-test		Confidence Intervals	
	Control *M* (SD)	Experimental *M* (SD)	Control Lower bound/ Upper bound	Experimental Lower bound/ Upper bound	Control *M* (SD)	Experimental *M* (SD)	Control Lower bound/ Upper bound	Experimental Lower bound/ Upper bound
Performance competence	7.00 (3.85)	7.08 (5.95)	2.96/11.04	0.84/13.33	12.08 (6.12)	19.83 (11.51)	3.70/20.47	11.45/28.22
Empathy	92.00 (14.75)	87.42 (19.85)	81.53/102.47	76.95/97.89	84.25 (17.29)	92.33 (15.72)	74.36/94.14	82.44/102.22
Stigma	26.42 (4.89)	27.58 (2.68)	24.06/28.78	25.22/29.94	25.50 (4.62)	30.67 (2.67)	23.24/27.76	28.41/32.93
Self-efficacy	41.25 (8.45)	40.83 (7.81)	36.38/46.12	35.96/45.71	41.58 (7.03)	51.83 (7.11)	37.35/45.81	47.60/56.06
Reported behavior	53.75 (4.69)	54.67 (2.71)	51.46/56.04	52.37/56.96	52.00 (4.51)	57.33 (3.14)	49.67/54.33	55.01/59.66
Knowledge	13.75 (1.66)	15.08 (1.51)	12.80/14.70	14.14/16.03	14.08 (1.68)	17.33 (0.888)	13.28/14.89	16.53/18.14

Table 3 Study 1: dependent *t* tests for the DEFUSE group and control group and independent samples *t* tests for post-test

	Pre-/post-comparisons for DEFUSE condition			Post-test comparisons for the DEFUSE and control conditions			Pre-/post-comparisons for the control condition		
	t	*df*	*p*	*t*	*df*	*p*	*t*	*df*	*p*
Performance competence	−4.509*	5	0.006	1.456	10	0.176	−2.948*	5	0.032
Empathy	−2.029	11	0.067	1.198	22	0.243	2.724*	11	0.020
Stigma	−4.093*	11	0.002	3.352*	22	0.003	0.935	11	0.370
Self-efficacy	−6.606**	11	0.000	3.553*	22	0.002	−0.316	11	0.758
Reported behavior	−3.646*	11	0.005	3.360*	22	0.003	1.686	11	0.120
Knowledge	−6.051**	11	0.000	5.935**	22	0.000	−0.804	11	0.438

*$p \leq 0.05$; **$p \leq 0.001$

Table 4 Study 2: means, standard deviations, and confidence intervals for quasi-experimental follow up of the control group

	Pre-test *M* (SD)	Confidence intervals Lower bound/upper bound	Post-test *M* (SD)	Confidence intervals Lower bound/upper bound
Empathy	77.00 (15.95)	60.26/93.74	90.83 (22.08)	67.67/114.00
Stigma	24.50 (5.68)	18.54/30.46	26.50 (7.94)	18.16/34.84
Self-efficacy	44.17 (6.97)	36.85/51.48	54.50 (7.58)	46.54/62.46
Reported behavior	52.00 (5.22)	46.53/57.47	56.50 (3.01)	53.33/59.67
Knowledge	14.67 (1.03)	13.58/15.75	17.17 (0.75)	16.38/17.96

References

Bahora, M., Hanafi, S., Chien, V. H., & Compton, M. T. (2008). Preliminary evidence of effects of Crisis Intervention Team training on self-efficacy and social distance. *Administration and Policy in Mental Health, 35*, 159–167.

Bogardus, E. S. (1925). Measuring social distances. *Journal of Applied Sociology, 9*, 299–308.

Booth, A., Scantlebury, A., Hughes-Morley, A., Mitchell, N., Wright, K., Scott, W., & McDaid, C. (2017). Mental health training programmes for non-mental health trained professionals coming into contact with people with mental ill health: a systematic review of effectiveness. *BMC Psychiatry, 17*(1), 196.

Borum, R., Dean, M. W., Steadman, H. J., & Morrissey, J. (1998). Police perspectives on responding to mentally ill people in crisis: perceptions of program effectiveness. *Behavioral Sciences & the Law, 16*, 393–405.

Broussard, B., Krishan, S., Hankerson-Dyson, D., Husbands, L., Stewart-Hutto, T., & Compton, M. T. (2011). Development and initial reliability and validity of four self-report measures used in research on interactions between police officers and individuals with mental illness. *Psychiatric Research, 189*, 458–462.

College of Policing. (2015). College of policing analysis: estimating demand on police service. Retrieved January 30, 2019 from https://www.college.police.uk/News/College-news/Documents/Demand%20Report%2023_1_15_noBleed.pdf.

Coleman, T.G., & Cotton, D. (2010). Police interactions with persons with mental illness: police learning in the environment of contemporary policing. Retrieved January 30, 2019 from https://www.mentalhealthcommission.ca/sites/default/files/Law_Police_Interactions_Mental_Illness_Report_ENG_0_1.pdf.

Coleman, T. & Cotton D. (2014). TEMPO: police interactions. A report toward improving interactions between police and people living with mental health problems. Retrieved January 30, 2019 from https://www.mentalhealthcommission.ca/sites/default/files/2016-05/TEMPO%20Police%20Interactions%20082014.pdf.

Compton, M. T., Bahora, M., Watson, A. C., & Oliva, J. R. (2008). A comprehensive review of extant research on Crisis Intervention Team (CIT) programs. *The Journal of the American Academy of Psychiatry and the Law, 36*(1), 47–55.

Compton, M. T., Bakeman, R., Broussard, B., Hankerson-Dyson, D., Husbands, L., & Krishan, S…Watson, A.C. (2014a). The police-based Crisis Intervention Team (CIT) model: I. Effects on officers' knowledge, attitudes, and skills. *Psychiatric Services, 65*(4), 517–522.

Compton, M. T., Bakeman, R., Broussard, B., Hankerson-Dyson, D., Husbands, L., & Krishan, S…Watson, A.C. (2014b). The police-based Crisis Intervention Team (CIT) model: II. Effects on level of force and resolution, referral, and arrest. *Psychiatric Services, 65*(4), 523–529.

Compton, M. T., Broussard, B., Hankerson-Dyson, D., Krishan, S., & Stewart-Hutto, T. (2011). Do empathy and psychological mindedness affect police officers' decision to enter Crisis Intervention Team training? *Psychiatric Services, 62*(6), 632–638.

Compton, M. T., Broussard, B., Hankerson-Dyson, D., Krishan, S., Stewart, T., Oliva, J. R., & Watson, A. C. (2010). System- and policy-level challenges to full implementation of the Crisis Intervention Team (CIT) model. *Journal of Police Crisis Negotiations, 10*, 72–85.

Compton, M. T., Esterberg, M. L., McGee, R., Kotwicki, R. J., & Oliva, J. R. (2006). Crisis Intervention Team training: changes in knowledge, attitudes, and stigma related to schizophrenia. *Psychiatric Services, 57*, 1199–1202.

Corrigan, P. W., Morris, S. B., Michaels, P. J., Rafacz, J. D., & Rusch, N. (2012). Challenging the public stigma of mental illness: a meta-analysis of outcome studies. *Psychiatric Services, 60*(10), 963–973.

Deane, M. W., Steadman, H. J., Borum, R., Veysey, B. M., & Morrissey, J. P. (1999). Emerging partnerships between mental health and law enforcement. *Psychiatric Services, 50*(1), 99–101.

Dupont, R., Cochran, S., & Pillsbury, S. (2007). Crisis Intervention Team core elements. CIT Center. Retrieved from http://www.cit.memphis.edu/information_files/CoreElements.pdf

El-Mallakh, P. L., Kiran, K., & El-Mallakh, R. S. (2014). Costs and savings associated with implementation of a police Crisis Intervention Team. *Southern Medical Journal, 107*(6), 391–395.

Fuller, D.A., Lamb, H.R., Biasotti, M, & Snook, J. (2015). Overlooked in the undercounted: the role of mental illness in fatal law enforcement encounters. Retrieved January 30, 2019 from https://www.treatmentadvocacycenter.org/storage/documents/overlooked-in-the-undercounted.pdf.

Gillig, P. M., Dumaine, M., Stammer, J. W., Hillard, J. R., & Grubb, P. (1990). What do police officers really want from the mental healthy system? *Hospital and Community Psychiatry, 41*(6), 663–665.

Herrington, V., & Pope, R. (2014). The impact of police training in mental health: an example from Australia. *Policing and Society, 24*(5), 501–522.

Kalisch, B. J. (1971). An experiment in the development of empathy in nursing students. *Nursing Research, 20*(3), 202–211.

Levy, S. R., Freitas, A. L., & Salovey, P. (2002). Construing action abstractly and blurring social distinctions: implications for perceiving homogeneity among, but also empathizing with and helping, others. *Journal of Personality and Social Psychology, 5*, 1224–1238.

Meichenbaum, D. H., & Deffenbacher, J. L. (1988). Stress inoculation training. *The Counseling Psychologist, 16*(1), 69–90.

Peck, Leonard W. (2003). Law enforcement interactions with persons with mental illness. Dissertation, Sam Houston State University.

Rafacz, J.D. (2012) Anti-stigma programs: stigma in campus police officers. Dissertation, Illinois Institute of Technology.

Reaves, B. A. (2009). *State and local law enforcement training academies, 2006 (NCJ 222987)*. Washington, DC: U.S. Department of Justice.

Reaves, B. A. (2016). *State and local law enforcement training academies, 2013 (NCJ 249784)*. Washington, DC: U.S. Department of Justice.

Sheehy, R., & Horan, J. J. (2004). Effects of stress inoculation training for 1st-year law students. *International Journal of Stress Management, 11*(1), 41–55.

Scantlebury, A., Fairhurst, C., Booth, A., McDaid, C., Moran, N., Parker, A., et al. (2017). Effectiveness of a training program for police officers who come into contact with people with mental health problems: a pragmatic randomised controlled trial. *PLoS One, 12*(9), e0184377.

Teagardin, J., Dixon, D. R., Smith, M. N., & Granpeesheh, D. (2012). Randomized trial of law enforcement training on autism spectrum disorders. *Research in Autism Spectrum Disorders, 6*(3), 1113–1118.

Teller, J. L., Munetz, M. R., Gil, K. M., & Ritter, C. (2006). Crisis Intervention Team training for police officers responding to mental disturbance calls. *Psychiatric Services, 57*, 232–237.

Tinney, M., & Rosenbaum, N. (2015). Police perceptions of the percentage of contacts in Albuquerque, NM that involve people living with mental illness. Unpublished manuscript, City of Albuquerque, Albuquerque, New Mexico.

West, D. J., Horan, J. J., & Games, P. A. (1984). Component analysis of occupational stress inoculation applied to registered nurses in an acute care hospital setting. *Journal of Counseling Psychology, 31*(2), 209–218.

Robyn L. Hacker is a psychologist with the Center for Dependency, Addiction, and Rehabilitation (CeDAR) at the University of Colorado. Here she coordinates the Professionals Program, supporting law enforcement officers in addiction and co-occurring mental health treatment. Dr. Hacker also has extensive experience working clinically with criminal offenders, most recently completing forensic addiction fellowships at Yale University School of Medicine. Her research has focused on the development and evaluation of evidence-based clinical intervention.

John J. Horan recently retired from long careers of teaching and research at Arizona State and Penn State Universities. In the past 20 years, he concentrated on the design and evaluation of Internet-delivered treatments. He continues to consult on evidence-based research in clinical psychology and behavioral health.

Affiliations

Robyn L. Hacker[1] · **John J. Horan**[2]

✉ Robyn L. Hacker
robyn.hacker@uchealth.org

[1] CeDAR, University of Colorado, Aurora, CO, USA

[2] Arizona State University, Tempe, AZ, USA