Crisis Intervention Team (CIT) metrics: a novel method of measuring police performance during encounters with people in crisis

Lois James[1]* and Stephen James[2]

[1]Washington State University College of Nursing, Spokane, WA, USA

[2]Washington State University Elson S. Floyd College of Medicine, Spokane, WA, USA

Abstract

Police officers frequently respond to mentally ill individuals in crisis states. Crisis Intervention Team (CIT) training is designed to help officers manage these challenging situations. Our objective was to develop a set of metrics for measuring police performance during encounters with people suffering from mental illness or in crisis, thus facilitating evaluation of CIT training. We employed a two-phase research design to develop the CIT metrics: first, we convened a diverse panel of experts in a "Reverse Concept Mapping" focus group to establish the measurable elements of police performance in CIT encounters, as well as the relative difficulty of those encounters. Second, we used Thurstone scaling to assign interval-level scores to each resulting difficulty and performance indicator by surveying hundreds of police officers and mental health professionals. The outcome of this process was a set of CIT metrics with 90 difficulty indicators and 112 performance indicators, that can be used for measuring police performance while controlling for situational difficulty. We also produced a "Quick Sheet Performance Metrics" comprised of the most important performance statements that can be applied in real time. The CIT metrics provide a way to empirically measure the effectiveness of CIT training, as well as informing CIT training development.

Introduction

Crisis Intervention Team (CIT) training is designed to help law enforcement officers effectively and appropriately manage situations involving persons in crisis [1]. Integrating crisis services and collaboratively managing crisis issues are high priorities in many communities. Law enforcement officers frequently respond to situations involving mental illness, developmental disability, or other crisis states, and need the skills to manage these encounters in ways that maximize safety and minimize tragic outcomes [2]. Since CIT's birth in Memphis, TN in 1988 in response to the shooting of an individual with mental illness, the training course has spread across the nation and is being used by many agencies to better equip their officers to deal with these challenging and often unpredictable situations [3].

Recognizing the importance of managing situations involving persons in crisis, many law enforcement departments have mandated all officers receive CIT training. Unfortunately, despite studies evaluating aspects of CIT training's effectiveness [4-7], little empirical validation has been conducted to investigate just *how* effective CIT training is at improving officers' ability to respond to persons in crisis. This is in part due to the difficulty of measuring training impact on officer *behavior* using outcome measures such as reduction in use of force, decreased arrest rates, or referral to mental health services (which may not allow officer discretion). We responded to this problem by developing a set of metrics that provides a way to empirically measure the effectiveness of CIT training by measuring police performance during encounters with people in crisis, while controlling for the difficulty of those encounters. The resulting metrics can also be used for informing CIT training development.

Method

Design

This process developed a set of metrics for 1) evaluating police performance in encounters with individuals in crisis, and 2) measuring the difficulty of those encounters. The "performance scale" measures officer behaviors in crisis encounters, and the "difficulty scale" improves the validity of the performance scale by measuring the relative difficulty of different crisis encounters. The difficulty scale also improves our ability to understand the social dynamics of crisis encounters in order to assure that performance standards and training are realistic. To develop these scales, we involved police and mental health professional experts in a two-phase study design. First, we used a reverse concept mapping focus group [8,9] to extract measurable variables from subject matter experts, (SMEs) tacit knowledge of what makes crisis encounters difficult and what constitutes good performance by the responding officer. Second, we used the Thurstone equal-appearing interval scaling approach [10,11] to assign values to variables by having hundreds of participants, with experience interacting with individuals in crisis, rate them on a Likert scale (e.g. 0=least impact on difficulty, 6=greatest impact on difficulty). Together, these methodologies allowed us to

extract specific and measureable variables from abstract concepts such as empathy.

Subjects

We convened a diverse panel of eighteen SMEs from law enforcement and mental health backgrounds to participate in the reverse concept mapping focus group. Nine of the SMEs were mental health professionals, and nine were law enforcement. All recruitment was done by way of personal request from the research team. Each SME represented a unique law enforcement or mental health discipline to ensure diverse experiences and knowledge.

The Thurstone scaling process received input from 499 professionals with experience in crisis encounters; police officers (n=387, 78%), mental health professionals (n=60, 12%), and other professionals who encounter individuals in crisis – such as correction officers and councilors (n=52, 10%) from different agencies across the United States. Of the respondents, 20% were female, 87% were white, and the average age was 45 (sd=9.96) years old.

Within the law enforcement respondents, 35% were at the rank of police officer/deputy sheriff, 29% were Corporals or Sergeants, 10% were detectives, and 20% were at the rank of lieutenant or higher. Seventy nine percent of respondents worked for municipal agencies, 14% for a county Sheriff's office, 4% were state police/patrol and 3% were federal law enforcement. Half (49%) of the respondents were assigned to patrol duties from 43 different states. Forty five percent of responding law enforcement professionals had received CIT training.

Within the mental health professionals (MHP) respondents 69% of the agencies provided mental health services, 11% services to individuals with developmental difficulties, 8% homeless services, 6% Veteran services, and 6% elder services. Forty percent of MHPs were case workers and 38% provided outreach services. Thirty nine percent of MHPs reported working alongside law enforcement on a daily or weekly basis while 54% stated they occasionally work with law enforcement.

We used a snowball recruiting process, whereby SMEs from the focus group recommended our survey to colleagues, police agencies and departments, and fraternal orders around the country. Requests for officers to participate in the survey were also sent to members of Force Science Institute, the National Tactical Officers Association, International Law Enforcement Training Association, Frontier Behavioral Health and other organizations via their respective websites. Access to the survey was provided through a link that was posted on websites and emailed to agency distribution lists.

Materials

During the reverse concept mapping focus group process, we used cards printed with the statements generated by the SMEs for the categorizing process (see procedures for details). During the Thurstone scaling process, we used "Survey Monkey," (www.surveymonkey.com) a widely used online survey instrument, to enable efficient scoring of the statements by the survey respondents. After responding was completed, Survey Monkey generated output that was uploaded into Excel. Survey Monkey also was used to gather demographic information from each Thurstone scaling respondent.

Procedures

The reverse mapping focus group took place over two full (8-hour) days. On the morning of the first day, both law enforcement and mental

health SMEs came to a consensus to develop a "goal statement" for the optimal outcome of a crisis encounter. The goal statement selected was as follows:

The goal of a law enforcement officer in an encounter with an individual in crisis is to select the most appropriate course of action; maximizing the safety of civilians, officers, and the person in crisis.

Following the generation of this goal statement, SMEs were provided with the "focus prompt":

An element of the encounter that makes this goal more difficult is…

They were given approximately twenty minutes to write down as many statements that relate to measurable variables affecting the difficulty of a crisis encounter as they could. They then one by one read their statements to the group. A member of the research team carefully recorded each statement. This frequently led to further discussion and clarifications in which other statements relating to the difficulty of a crisis encounter were suggested. During this phase the researchers were careful to facilitate, but not to direct or guide the focus group, in order to minimize biasing SMEs. Once participants were satisfied that all aspects of crisis encounter difficulty were addressed, related statements were integrated (with input from SMEs to ensure we did not remove any statements they believed stood alone) to create a final list of statements.

On the afternoon of the first day, SMEs were provided with a set of cards containing all the difficulty statements generated in the morning session; one statement per card. The SMEs were asked to group the statements into self-selected categories. They were not given any input on how many categories to select or what to name those categories; we wanted to elicit subjective information on how SMEs would categorize the difficulty statements.

On the second day of the focus group SMEs then repeated day-1 procedures but instead of nominating difficulty statements, they nominated performance statements, using the following focus prompt:

An element of a law enforcement officer's performance that affects the likelihood of achieving this goal is…

When the two-day focus group was complete, the research team entered all statements and suggested categories relating to both difficulty and performance into an excel spreadsheet. Categories that appeared the most frequently were selected to group difficulty and performance statements in online surveys for the Thurstone scaling phase of the research.

The Thurstone scaling procedures involved preparing online surveys in Survey Monkey that contained the categorized difficulty and performance statements. An email request to participate containing a link to the survey was then distributed to potential respondents. If the email recipient chose to respond to the survey, they clicked on the link, entered their responses, and clicked submit. No personal information was collected from respondents to ensure anonymity (as such this data collection was exempt from institutional review). If they did not chose to participate they were instructed to ignore the link. Response time to complete the survey was approximately 30-45 minutes.

Half of the survey respondents rated the difficulty statements first, and the other half rated the performance statements first in an attempt to counterbalance any survey fatigue issues. Each respondent was asked to rate difficulty statements on a seven-point Likert scale from 0 "no impact on difficulty" to 6 "maximum impact of difficulty".

Each respondent was asked to rate performance statements on a nine-point Likert scale from -4 "extreme negative effect on performance" to +4 "extreme positive effect on performance" (a 0 indicated no effect on performance). The surveys were open for a period of 4 weeks, and weekly reminders were emailed via distribution lists to prompt willing participants to respond if they had not already. When the surveys were closed, data was downloaded into Excel and the median value given to each statement by the survey respondents was then was assigned as that statement's value. This process gave us interval-level scales for measuring the concepts judged to be most relevant for understanding officer performance in crisis encounters and the relative difficulty of different crisis situations.

Results

During the focus group, the SMEs generated 90 statements relating to the difficulty of a crisis encounter, and 112 statements relating to officer performance in a crisis encounter. The results of the grouping process revealed the following most frequently suggested difficulty categories: "safety issues," "persons in crisis – symptoms," "persons in crisis – characteristics," "communication issues," "environmental issues," "other people present" and "resources." The most frequently suggested performance categories were: "policy and law," "training and wellness," "pre-plan," "assess," "tactics," "self-control," "interacting with the person in crisis," "resources," and "disengage and follow up." Table 1 provides examples of statements for each category.

During Thurstone scaling, the median value given to each statement by the raters then was assigned as that statement's value (see table 1 for scores assigned to example statements). The scores assigned to difficulty statements ranged from 0 (no impact on difficulty) to 6 (maximum impact on difficulty). Examples of statements that received the highest difficulty score are:

- "Inadequate staffing being available to respond to the crisis"

- "Not knowing where the person in crisis is located"

- "Having the opportunity for multiple persons to go into crisis (e.g. a bomb explosion, school shooting)"

The scores assigned to performance statements by the survey respondents ranged from -4 (extremely negative impact on performance) to 0 (no impact on performance) and up to +4 (extremely positive impact on performance). Examples of statements that received the highest performance score are:

- "Seeking accurate information about the situation before arrival"

- "Reading non-verbal cues of the person in crisis"

- "Adapting based on changing level of threat"

In addition to the full set of difficulty and performance metrics (see appendix) we also developed a "Quick Sheet Performance Metrics" comprised of the 18 performance statements that were most frequently rated as being of greatest importance. The reduced set of metrics are presented in Table 2, and can be applied in real time (for example to score performance during a role play scenario) to assign officers a performance score ranging from 0-18.

Discussion

Approximately one in five Americans over the age of 13 experiences mental illness in any given year. One in four homeless adults live with severe mental illness, and over 70% of youth in the juvenile justice system have at least one mental illness. More than 90%

Table 1. CIT metric categories, with an example statement and corresponding score for each

Category	Example Statement	Score
Difficulty		
Safety Issues	Hostages being present	6
Person in Crisis - Symptoms	The person in crisis appearing to be hallucinating (visual or auditory)	6
Person in Crisis - Characteristics	The person in crisis having a history of negative encounters with law enforcement	4
Communication Issues	The person in crisis being unable to communicate due to language barriers	6
Environmental Issues	The person in crisis having barricaded him or herself (e.g. behind a locked door)	6
Other People Present	Children being present	4
Resources	Not having timely access to mental health professionals	3
Performance		
Policy and Law	Understanding the civil commitment laws for mental illness	4
Training and Wellness	Seeking opportunities to participate in Crisis Intervention Team (CIT) training	4
Pre-Plan	Seeking access to previous information about the person in crisis (e.g. personal history)	3
Assess	Recognizing key indicators of mental illness	4
Tactics	Coordinating on scene roles between law enforcement and mental health professionals	3
Self-Control	Managing personal attitudes about mental illness	4
Interacting with the Person in Crisis	Being able to see things from the person in crisis's point of view	4
Resources	Giving the person in crisis information that can help him or her	3
Disengage and Follow Up	Accurately documenting the details of the encounter	3

Table 2. "Quick Sheet Performance Metrics" comprised of the 18 performance statements that were most frequently rated as being of greatest importance

Category	Performance Statement	Achieved (yes=1, no=0)
Assess	Reading non-verbal cues of the person in crisis	
	Picking up on the tone of voice of the person in crisis	
	Recognizing weapons of opportunity in the environment (e.g. rocks, lamp shades)	
	Continually assessing the level of threat of the person in crisis	
Tactics	Adapting based on changing level of threat	
	Obtaining critical information from family members about the person in crisis	
Self-control	Controlling his or her own tone of voice	
	Appearing confident in him or herself during the encounter	
	Not taking things personally (e.g. verbal abuse)	
	Controlling his or her emotions during the encounter	
	Having the desire to succeed throughout the encounter	
Interacting with the person in crisis	Having the ability to de-escalate a situation (calm the person in crisis down)	
	Being aware of the words he or she uses and how they influence the person in crisis	
	Actively listening to the person in crisis during the encounter	
	Not patronizing or insulting the person in crisis	
	Providing clear instructions to the person in crisis	
	Demonstrating concern for the person in crisis's safety	
	Demonstrating patience with the person in crisis	
	Total Score:	

of children who commit suicide have a mental illness. Furthermore, it's estimated that serious mental illness costs America $193.2 billion a year in lost earnings. Yet only 41% of individuals with mental illness received mental health services in 2014, and that number is about half for African Americans and Hispanic Americans [12]. In the absence of

these services, individuals in crisis frequently come into contact with the criminal justice system. Due to the extreme behaviors sometimes exhibited by individuals with mental illness, their contacts with law enforcement can end in tragedy. Accordingly, it is critically important to train officers on effective techniques of interacting with persons in crisis, and effectively evaluate how well training prepares them for these encounters.

In response to this broad societal problem, we developed a set of interval-level metrics for measuring the performance of police officers responding to calls for service involving persons with mental illness or in crisis. Prior to the development of these metrics, the ability to evaluate performance in these types of social encounters was limited to probabilistic outcome measures (such as arrest rates, uses of force, or mental health referrals). Consequently, this reduced our ability to understand the dynamics of crisis encounters, or to evaluate CIT training effectiveness based on changes to officer behavior—which is the ultimate goal of any police training. We expect the CIT metrics to improve the ability to evaluate what kind of policies, practices, and training work best. In turn, this could improve the ability of law enforcement policy makers to make informed decisions regarding future implementation of CIT training. Validation studies are required to determine whether these metrics are in fact more accurate and precise than other methods (for example subjective expert review).

Since the development of the CIT metrics, they have been used to inform an "Enhanced" CIT (E-CIT) training class in a mid-sized urban police department with learning objectives extracted directly from the performance categories. Namely: 1) Knowing relevant policies and laws; 2) Practicing health and wellness; 3) Being able to put an effective plan in place before the encounter; 4) Being able to assess the situation on arrival; 5) Effectively using tactics during the encounter; 6) Being able to use self-control during the encounter; 7) Being able to appropriately interact with the person in crisis; 8) Being able to use resources effectively; and, 9) Being able to disengage and put a follow-up plan in place. Two cohorts totaling 35 officers have been trained on this model, and anecdotal evidence suggests that both officers and community members respond well to the training. Plans are underway to evaluate officer performance before and after E-CIT training by using the CIT metrics to score incident reports and (where feasible) body camera footage.

Law enforcement officers face huge challenges when it comes to effectively and appropriately interacting with people in crisis, who may be exhibiting extreme and dangerous behaviors. They are required to alleviate harm to the person in crisis, promote decriminalization of individuals with mental illness, reduce the stigma associated with mental disorders, and use a team approach when responding to crises, all while maintaining personal safety and reducing risk of injury to fellow officers and members of the public [13]. Despite it being well acknowledged that CIT training is critical for providing officers with the skills to manage these unpredictable and complex situations, rigorous evaluations of CIT training have yet to determine whether the knowledge gained during training translates into actual behavior change in the field. It is our hope that the CIT metrics developed here will facilitate such evaluations, and meaningfully impact CIT training nationwide.

Authorship and contribution

Both authors were responsible for the conception and design of the study described in this manuscript. As co-PIs on the research, they were jointly responsible for data collection, analysis, and dissemination.

Both authors had an active role in drafting the manuscript, and both have reviewed it thoroughly.

Competing interest

The authors have no financial or non-financial conflicts of interest to declare

References

1. Sanow Ed (2006) Crisis Intervention Team. Law & Order 54: 26-35.

2. Augustin D, Fagan T (2011) Roles for mental health professionals in critical law enforcement incidents: An overview. *Psychological Services* 8: 166-177.

3. Teller J, Munetz M, Gill, Ritter C (2006) Crisis intervention team training for police officers responding to mental health disturbance calls. *Psychiatr Serv* 57: 232-237.

4. Steadman HJ, Deane MW, Borum R, Morrissey JP (2000) Comparing outcomes of major models of police responses to mental health emergencies. *Psychiatr Serv* 51: 645–649.

5. Compton MT, Bahora M, Watson AC, Oliva JR (2008) A comprehensive review of extant research on Crisis Intervention Team (CIT) programs. *J Am Acad Psychiatry Law* 36: 47-55. [Crossref]

6. Young AT, Brumley N (2009) On-scene mental health services: A case study of the Lubbock Police Department's Victims Services Crisis Team. FBI Law Enforcement Bulletin.

7. Compton MT, Bakeman R, Broussard B, Hankerson-Dyson D, Husbands L (2014) The police-based crisis intervention team (CIT) model: II. Effects on level of force and resolution, referral, and arrest. *Psychiatr Serv* 65: 523-529. [Crossref]

8. Vila B, Stephen J, James L (2016) Novel Process for Developing Metrics That Measure What Police Do. Interservice/Industry Training Simulation and Education Conference, I/ITSEC Proceedings.

9. Vila B, James S, James L (2017) How police officers perform in encounters with the public: Measuring what matters at the individual level. Policing: an International Journal of Police Strategies & Management. Forthcoming.

10. Thurstone L, Chave E (1929) The Measurement of Attitude. Chicago: University of Chicago Press.

11. Trochim WM (2001) The Research Methods Knowledge Base, 2nd Edition. Cincinnati: Atomic Dog.

12. National Alliance on Mental Illness. https://www.nami.org/

13. Jines J (2013) Crisis Intervention Teams: Responding to Mental Illness Crisis Calls. FBI Law Enforcement Bulletin.