Innovation in Targeted Violence and Terrorism Prevention: Phase 1 Summary Report

Joel Elson
Erin M. Kearns

Follow this and additional works at: https://digitalcommons.unomaha.edu/ncitereportsresearch
Please take our feedback survey at: https://unomaha.az1.qualtrics.com/jfe/form/SV_8cchtFmpDyGfBLE
Innovation in Targeted Violence and Terrorism Prevention

Developing and Testing an Intelligent Chatbot to Help Individuals Identify Threats and Improve Tips Reporting
RESEARCH TEAM

Members

Joel S. Elson (Principal Investigator), University of Nebraska at Omaha, NCITE
Erin M. Kearns (Co-Principal Investigator), University of Nebraska at Omaha, NCITE

Key Personnel

Isaac Moelter, University of Nebraska at Omaha, NCITE
John Pintar, University of Nebraska at Omaha, NCITE
Sam Schneider, University of Nebraska at Omaha, NCITE
Callie Vitro, University of Nebraska at Omaha, NCITE
Lauren Zimmerman, University of Nebraska at Omaha, NCITE

About NCITE. The National Counterterrorism Innovation, Technology, and Education (NCITE) Center was established in 2020 as the Department of Homeland Security (DHS) Center of Excellence for counterterrorism and terrorism prevention research. Sponsored by the DHS Science and Technology Directorate (S&T) Office of University Programs, NCITE leads an elite academic consortium of more than 50 researchers at partner institutions across the U.S. and Europe. Headquartered at the University of Nebraska at Omaha, NCITE is the principal U.S. academic partner for counterterrorism research, technology, and workforce development.

Acknowledgement. The research in this report was supported by the U.S. Department of Homeland Security’s Center for Prevention Programs and Partnerships under Grant Award Number DHS-21-TTP-132-00-01.

Disclaimer. Any opinions or conclusions contained herein are those of the authors and do not necessarily reflect those of the Department of Homeland Security, the DHS Science and Technology Directorate, or the University of Nebraska System.
EXECUTIVE SUMMARY

This report provides summary findings for data collection efforts in Phase I of the entitled FY21 TVTP grant. The primary aims in Phase I were to 1) better understand the many points where breakdowns can and do occur during the process of identifying and sharing information about potential terrorism and targeted violence (TTY) threats and to 2) develop a prototype basic chatbot to overcome some of these breakdown points that is then tested against a static webform, which is the industry standard. This report presents key findings from reporting data, threat assessment teams, and members of the public to support these objectives.

Report Key Findings

1. Archival Analyses
   a. Threat reports are often submitted by individuals acting in a professional capacity.
   b. People – even professionals – struggle with identifying incident type on reporting forms.
   c. Most reports are filed anonymously.

2. Focus Groups
   a. While team structure varies, law enforcement, education, mental health are core components of behavioral threat assessment and management teams (BTAMTs).
   b. Trust is built through open, transparent communication – both within BTAM team members and between BTAMs and communities.
   c. Buy-in begins with leadership, who provide support and a central connection point.
   d. Team members should be competent, know what is expected, and support each other.
   e. Rural areas often lack resources, including the need for better reporting systems and threat assessment tools.

3. National Survey
   a. People have poor awareness of SAR indicators or pre-incident TTY behaviors.
   b. Willingness to report is abstract: what people say v. what they have done is mismatched.
   c. Many are not sure how to report TTY and would turn to the internet for information.

4. Lab Study
   a. People struggle to identify incident type on reporting forms, which may deter them from continuing in the report.
   b. Reports received through a prototype basic chatbot are more concise than reports received through a static webform.
   c. Reports received through a prototype basic chatbot are at least as accurate as those received through a static webform, and some elements are more accurate.
   d. There is no difference in system-level trust or usability between the static webform and the prototype basic chatbot.

5. National Survey on Chatbot Experiences and Attitudes
   a. Most people interact with chatbots anywhere from multiple times a day to weekly.
   b. Younger people are more trusting of chatbots, view them favorably, and intend to use them.
## TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>RESEARCH TEAM</td>
<td>2</td>
</tr>
<tr>
<td>EXECUTIVE SUMMARY</td>
<td>3</td>
</tr>
<tr>
<td>PROJECT OVERVIEW</td>
<td>5</td>
</tr>
<tr>
<td>ARCHIVAL ANALYSIS</td>
<td>6</td>
</tr>
<tr>
<td>Understanding Incident Types and Anonymity</td>
<td>6</td>
</tr>
<tr>
<td>Understanding Reporter Identity and Anonymity</td>
<td>7</td>
</tr>
<tr>
<td>Understanding Trends over Time</td>
<td>7</td>
</tr>
<tr>
<td>Understanding Responder Usage</td>
<td>7</td>
</tr>
<tr>
<td>FOCUS GROUPS</td>
<td>10</td>
</tr>
<tr>
<td>NATIONAL SURVEY</td>
<td>13</td>
</tr>
<tr>
<td>Identifying Behavior as Concerning</td>
<td>13</td>
</tr>
<tr>
<td>Willingness to Report</td>
<td>14</td>
</tr>
<tr>
<td>Knowing How to Report</td>
<td>15</td>
</tr>
<tr>
<td>LAB STUDY</td>
<td>16</td>
</tr>
<tr>
<td>Static Reporting Webform – Interaction Experiences</td>
<td>16</td>
</tr>
<tr>
<td>Static Webform v. Prototype Basic Chatbot – Response Length and Sentiment</td>
<td>18</td>
</tr>
<tr>
<td>Static Webform v. Prototype Basic Chatbot – Willingness to Share Contact Information</td>
<td>18</td>
</tr>
<tr>
<td>Static Webform v. Prototype Basic Chatbot – Report Accuracy</td>
<td>18</td>
</tr>
<tr>
<td>Static Webform v. Prototype Basic Chatbot – System Trust and System Usability</td>
<td>19</td>
</tr>
<tr>
<td>BONUS SURVEY – NATIONAL CHATBOT ATTITUDES AND EXPERIENCES</td>
<td>20</td>
</tr>
<tr>
<td>Familiarity and Frequency with Use of Chatbot Technology</td>
<td>20</td>
</tr>
<tr>
<td>Chatbots: Trust in Them, Positive Attitudes toward Them, and Intentions to Use Them</td>
<td>21</td>
</tr>
<tr>
<td>REFERENCES</td>
<td>22</td>
</tr>
<tr>
<td>METHODOLOGICAL APPENDIX</td>
<td>23</td>
</tr>
<tr>
<td>Archival Analysis</td>
<td>23</td>
</tr>
<tr>
<td>Focus Groups</td>
<td>24</td>
</tr>
<tr>
<td>National Survey</td>
<td>24</td>
</tr>
<tr>
<td>Lab Study</td>
<td>26</td>
</tr>
<tr>
<td>Bonus Survey – National Chatbot Attitudes and Experiences</td>
<td>29</td>
</tr>
</tbody>
</table>
**PROJECT OVERVIEW**

**Problem Statement**
Effective terrorism and targeted violence (TTV) prevention and intervention relies upon members of the public (reporters) to provide tips about concerning behavior to authorities. There are multiple points during the process of identifying and sharing information about potential TTV threats where breakdowns can occur, hindering intervention efforts resulting in deadly consequences. Three critical questions need to be addressed, specifically: 1) how do members of the public identify what behaviors are concerning, 2) what factors impact their willingness to report those behaviors to non-law enforcement entities, and 3) how can familiarity with tips reporting options be increased. Most research in this area has focused on public willingness to report concerns to law enforcement. Thus, we know little about how this process functions, and more importantly why it fails.

**Innovation Grant Goals**
Our goal is to improve the tips reporting process and minimize breakdowns in this process. Our innovation project is creating and testing a chatbot for TTV tips reporting. In Phase I, our aim was to better understand real-world tips reporting processes and barriers to reporting using mixed-methods design that combines real-world and lab-based data. This report contains summary findings from all Phase I data collection efforts and testing of a prototype basic chatbot. In Phase II, we will use insights from Phase I to refine and test more advanced chatbot technologies.

**Figure 1. Innovation Project Process**
ARCHIVAL ANALYSIS

With the archival analysis, our main aim was to understand system usage patterns and trends in reporting behavior. We partnered with a local sheriff’s office and their county behavioral threat assessment and management (BTAM) team that uses a national reporting platform to handle reports from community members. To enter a report, this platform provides access to a static webform that is available on various school and county websites. To start a report, a person must first select a location and an incident type from a list of options. The form then contains relevant fields for the person to enter information. BTAM team members can then read and manage report. The dataset we received contained deidentified, actual reports from August 8, 2016 through January 31, 2022. We conducted various exploratory analyses. See the Appendix of this report for methodological details.

Understanding Incident Types and Anonymity

At the beginning of a report, the person reporting must select the type of incident they are reporting. We plotted these incident types in a bar chart, as shown in Figure 2. Four incident types were most commonly reported: Drugs/Alcohol/Tobacco, Threat to Harm Others, Other, Bullying/Harassment/Intimidation. Throughout this document, we refer to these as substance use, threat, other, and bullying reports respectively. The high prevalence of other reports may indicate that reporters are having difficulty labeling an appropriate incident type.

Figure 2. Total Reports by Incident Type

To investigate anonymity preferences, we calculated the percentage of reports that were filed anonymous overall and for each incident type. Overall, 57% of reports were filed anonymously. Substance use reports were the most common and had the second largest proportion of anonymous reports (91.3%). Threat reports were the second most common but had the lowest proportion of anonymous reports (18.8%). One main reason for this is that threat reports were often filed by someone acting in a professional capacity (e.g., school administrator or a law enforcement officer) where identifiability is not a concern when reporting.
Understanding Reporter Identity and Anonymity

Students filed the most reports, followed by parents and then community members. Students had the strongest preference for reporting anonymously, which they chose to do 84% of the time. Community members had the second highest preference for anonymity, with 66% choosing to remain anonymous. Finally, parents reported anonymously 55% of the time.

Table 1. Who Reports

<table>
<thead>
<tr>
<th>Reporter</th>
<th>Number of Reports Filed</th>
<th>% of Total Reports</th>
</tr>
</thead>
<tbody>
<tr>
<td>Students</td>
<td>349</td>
<td>34.5%</td>
</tr>
<tr>
<td>Parents/Guardians</td>
<td>206</td>
<td>20.4%</td>
</tr>
<tr>
<td>Community Members</td>
<td>162</td>
<td>16.0%</td>
</tr>
<tr>
<td>Unknown</td>
<td>137</td>
<td>13.6%</td>
</tr>
<tr>
<td>Law Enforcement Officer</td>
<td>78</td>
<td>7.7%</td>
</tr>
<tr>
<td>School Administrator</td>
<td>62</td>
<td>6.1%</td>
</tr>
<tr>
<td>Staff Member</td>
<td>17</td>
<td>1.7%</td>
</tr>
</tbody>
</table>

Understanding Trends over Time

From 2016 through 2019, there was a steady annual increase in reports. While the number of reports dip in 2020 (corresponding with the COVID-19 pandemic), the 2021 number is slightly greater than 2019. Monthly reporting trends correspond with the academic calendar. Reports are highest in September (the first full month of school), drop in January (Winter break and return from break), and decline most dramatically over the summer break. Weekly reporting trends show the highest reporting on Thursday and Fridays and the lowest reporting on weekends.

Figure 3. Reporting Trends Over Time

Understanding Responder Usage

Once a report is filed, there are specific activities that members of the BTAMT can take in response to that report. See Table 2 for a list of these activities and their descriptions.

Table 2. Activities within the Reporting System

<table>
<thead>
<tr>
<th>Activity</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Read</td>
<td>Read the report</td>
</tr>
<tr>
<td>Acknowledge</td>
<td>Must be done for a responder to do anything other than read the report</td>
</tr>
<tr>
<td>ActionTaken</td>
<td>Add a note to the report (i.e., notes on interviews, discussions, suspensions)</td>
</tr>
<tr>
<td>OwnerChange</td>
<td>Change the owner of the report</td>
</tr>
<tr>
<td>AssignedTeam</td>
<td>Assign an internal team to the report, for example, local police, mental health team, etc.</td>
</tr>
<tr>
<td>Action</td>
<td>Description</td>
</tr>
<tr>
<td>---------------------</td>
<td>-----------------------------------------------------------------------------</td>
</tr>
<tr>
<td>PriorityChange</td>
<td>Change the priority of the report (options: none, notes only, monitoring,</td>
</tr>
<tr>
<td></td>
<td>low, moderate, high). Reports automatically begin with none.</td>
</tr>
<tr>
<td>GrantAccess</td>
<td>Allow another specified responder to access the report</td>
</tr>
<tr>
<td>GrantAdmin</td>
<td>Give another specified responder administrator access to the report</td>
</tr>
<tr>
<td>InvolvedIndividual</td>
<td>Update a report’s information on the involved individual (subject of the</td>
</tr>
<tr>
<td>Updates</td>
<td>report)</td>
</tr>
<tr>
<td>RemoveAdmin</td>
<td>Retract a specified responder’s administrator access to the report</td>
</tr>
<tr>
<td>RemoveAccess</td>
<td>Retract a specified responder’s access the report</td>
</tr>
<tr>
<td>TypeChange</td>
<td>Change the incident type of the report (e.g., weapons)</td>
</tr>
</tbody>
</table>

When a report is submitted, relevant BTAM members receive a notification. These BTAM members need to read the report and acknowledge it before they can take any other actions in the system. On average, it took roughly 34 minutes after submission for a report to receive its first read from a team member, and 50 minutes for it to be acknowledged. However, actions beyond these took much longer as shown in Figure 4.

**Figure 4.** Mean Days After Report was Filed to First Activity

As shown in Figure 5, approximately 75% of reports had only 3 unique activity types – ReportRead, Acknowledge and ActionTaken – suggesting that responders primary use the system to read reports and record notes.
Figure 5. Percentage of Reports that Contain Each Activity

Activity Type

- Acknowledge
- ReportRead
- ActionTaken
- OwnerChange
- AssignedTeam
- PriorityChange
- GrantAccess
- NameChange
- GrantAdmin
- InvolvedIndividualUpdate
- RemoveAdmin
- RemoveAccess
- TypeChange
FOCUS GROUPS

Our main aim with the focus groups was to explore the strengths and weaknesses that threat assessment team members perceive in current tips reporting and identify potential improvements. We conducted focus groups with members of BTAM teams from August 2022 through February 2023. First, we conducted in-person focus groups with the BTAM team in Sarpy County, NE (n=22) that were broken out by focus area (i.e., law enforcement, schools, mental health services). Next, we conducted focus groups over Zoom with available members of whole BTAM teams across the state of Nebraska (n=20) and across the country (n=21) including a range of jurisdictions including ones in Alabama, Connecticut, Oklahoma, and Oregon.

Team Structure

- Different structures across threat assessment teams in our sample

- Law Enforcement:
  - Often the first group involved in prevention and continue to lead these efforts
  - Every team discussed the importance of law enforcement’s roll in prevention
  - Concerns that lower public trust in law enforcement may reduce reporting
  - Multiple roles: local FBI, municipal police, sheriff’s office, SROs, public safety
  - Challenges: not needed in every situation; difficult to access in smaller towns

- School Leadership:
  - School leadership’s role is critical – need their buyin for this to work
  - School leadership can share information and resources across disciplines and districts
  - Schools usually work closely with law enforcement entities
  - Multiple roles: Superintendents, local school leadership (e.g., deans, principals, some teachers), university staff (e.g., HR, student affairs, Title IX, counseling services, student conduct)
  - Challenges: strained resources in some locations

- Mental Health Experts:
  - In communities where mental health professionals are available, they are involved
  - Varied level of embeddedness within the team (part of system v. ad hoc involvement)
  - Partner with law enforcement and school leadership to assist with different cases
  - Many stressed the importance of better mental health awareness in these situations
  - Multiple roles: crisis counselors, school counselors, therapists, social workers
  - Challenges: Many communities have few – or no – mental health professionals

- Other Resources
  - County Prosecutors: may get involved in extreme cases
Team Balance
- Different structures across BTAMTs means different types of balance
  - Some collaborations are tight knit while others are spread over larger areas
- Achieving team balance takes time and varies by location – there is no one-size-fits-all model
- Regardless of structure, teams note that interdisciplinary expertise is essential
  - Balanced teams provide comprehensive understanding of best approaches
- Actors must be willing to cross boundaries and see tips from others’ perspectives
- Challenges: Many states have no standard for individual schools to have threat assessment teams so this is ad hoc and may not have the critical mission support necessary for success

Building Team Culture and Trust
- Team solidarity and cohesion is key to success – and trust is essential for this
- Trust is built over time through informal conversation and table-top exercises
- Politeness and bonding activities help break tension and build relationships
- Team must be friendly and connected – though no need to all be friends
- Collaborative decision making with the right people involved

Building Trust with the Community
- Mindfulness of community concerns – listen, engage, respond
- Work with partners on clear communication plans for when something does happen
- Encourage participation across community stakeholders
- Inform communities about what threat assessment is and why this is important
- Help train community members about what and how to report suspicious activity
- Local considerations (e.g., university campus, tribal nations)

Open Communication and Transparency within the Team
- Many groups note this is the most important aspect of a threat assessment team
- Need to have a multidisciplinary team with a common goal working together
- Cultural differences among team members – promote collaborative environment
- Team members should be competent, know what is expected, and support each other
- Team members must also understand their limitations and know how to be effective
- Team members should know everyone’s individual mission and how it fits into the larger picture
- It is critical to respect expertise and boundaries – and communicate openly
- Be contentious to prevent territorial disputes – nothing “belongs” to one person
- No decision should be made by a few people – every stakeholder should be involved
- No one subset of the team should have priority - who leads should depend on the situation
- Sense of a camaraderie and open communication go together
- Debrief following every incident, no matter how minor the incident seems
Open Communication and Transparency with the Community

- Have a simple, written plan that fits the community, so actions are transparent
- Communicate openly with other jurisdictions and ensure you’re communicate back
- Be transparent about issues and share learning experiences with other groups
- Articulate the rationale around decisions and communicate that openly

Internal Buy-In

- Buy-in begins with leadership, who provide support and a central connection point
- Everyone must be self-motivated, value the work, and have shared responsibility
- Buy-in is maintained through careful selection of new members with the same drive
- Training for everyone is critical - no one jumps into a situation without preparation

Having the Right Resources

- Whatever tool(s) they use should be supported by research
- Everyone needs to well-trained on the tool(s) they are using
- Critical to choose the right people to be involved and organically shift to changing needs
- Need to stay up to date on research and disseminate that out to their whole team
- Know the resources and limitations of your individual threat assessment team

Needs and Challenges

- Many rural areas lack adequate resources and technologies
- Need for a better reporting system or threat assessment tool
- Need better support system for people impacted by reports
- Some threat assessment teams experience politicking and territorial disputes
- Unclear when to bring up situations, especially without an actual threat assessment tool
- Difficulty communicating across groups if people don’t prioritize information sharing
- Difficult to get people invested as it takes time out of the workday and requires extra hours
- Frustration with lack of resources if situations come back with a red flag
- Shortage of intervention options in some areas, especially with less community involvement
- Suggestion: have a full-time coordinator to ensure that nothing slips through the cracks
NATIONAL SURVEY

The main aim here was to understand breakdowns in the pre-incident tips reporting process. Specifically, 1) how members of the public identify what behaviors are concerning, 2) what factors impact their willingness to report those behaviors to non-law enforcement entities, 3) how knowledgeable are people about how and where to report concerns, and 4) perceptions of the current terrorism and targeted violence threat landscape in the United States. We hired Qualtrics, an online survey research company, to provide a sample of U.S. adults (n=1,302) in October 2022. See the Appendix of this report for methodological details.

Identifying Behavior as Concerning

Survey participants were presented with 3 unique scenarios that contained 3 pieces of information about the potential suspect: gender, age, and behavior. Possible behaviors included 6 validated SAR indicators that are split into criminal and non-criminal acts and then further vary from low to high prevalence. We also included 2 non-SAR behaviors – one criminal and one non-criminal. See below for an example of what one of the scenarios looked like:

A male who is about late teens/early 20s is seeking information about the personnel, facilities, systems, and functions of a nearby company.

After reading each scenario, participants were asked to evaluate (a) how threatening and (b) how suspicious they thought the scenario was. On average, people viewed the scenarios as slightly more suspicious than threatening.

Participants were also asked whether they thought each scenario should be (a) reported to law enforcement and (b) reported to other non-law enforcement authorities. On average, people were moderately more likely to think scenarios should be reported to law enforcement compared to a non-law enforcement authority.

In Figure 6 below, the first figure estimates the effects of a suspect’s gender, age, and the behavior they’re engaging in on how threatening and how suspicious people think the scenario is. The second figure examine how a suspect’s gender, age, and their behavior impact what people view should be reported to both law enforcement and to non-law enforcement authorities.

People view a scenario as slightly less threatening when the suspect is female, otherwise neither gender nor age are related to perceptions of threat or suspicion. Though all the behaviors aside from yelling and selling drugs are validated SAR indicators and thus should be viewed as more threatening and suspicious and people should think these are more reportable to law enforcement and non-law enforcement authorities alike. Results show, however, that this is clearly not the case.
Whether people think something is threatening or suspicious and whether the think it should be reported to either law enforcement or non-law enforcement authorities is largely a function of whether that behavior is a criminal act. While this is a positive take-away from criminal behaviors, it suggests that more education is needed about validated SAR indicators.

**Willingness to Report**

We first asked people’s hypothetical willingness to report crime, which was generally high. People indicated greater willingness to report a crime than a person who broke the law with a moderate effect size.

We then asked about people’s actual experiences witnessing crime and reporting it to police, as shown in Table 3. Most people have not witnessed a crime and only a bit over half think they have been in a situation where they thought they should call the police. Of those who have been in a situation where they thought that they should call the police, 25% did not do so.

**Table 3.** Witnessing and Reporting Crime

<table>
<thead>
<tr>
<th></th>
<th>Yes</th>
<th>No</th>
<th>Unsure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Witnessed a crime or potential crime</td>
<td>37.9%</td>
<td>54.0%</td>
<td>8.1%</td>
</tr>
<tr>
<td>Been in a situation where you thought you should call the police</td>
<td>53.7%</td>
<td>41.2%</td>
<td>5.2%</td>
</tr>
<tr>
<td>Ever called the police to report a crime or a potential crime</td>
<td>45.9%</td>
<td>49.5%</td>
<td>4.6%</td>
</tr>
</tbody>
</table>
Knowing How to Report
We first asked participants if they had heard of the “See Something, Say Something” campaign. Despite existing for two decades, only 64.4% of US adults indicate familiarity with it while unsure 8.1% are unsure. We next asked people where they would turn for information about reporting and, as shown in Figure 7, the vast majority would turn to the internet for advice.

Figure 7. If you wanted info on how to report, who would you ask?
LAB STUDY

The main aim here was to compare behaviors and cognitive processes involved in reporting suspicious activity through a static webform v. a prototype basic chatbot. Specifically, we wanted to 1) identify the most challenging aspects of reporting through a static webform 2) examine reporting through a chatbot, and 3) compare reporting through a static webform v. a prototype basic chatbot.

We developed a mock static reporting webform with a visual interface and questions that reflect national reporting platforms, as shown in Figure 8. Ninety-nine adults in the Omaha metro area completed this study from November 2022 to February 2023.1 Participants were randomly assigned to report either using a static webform or the prototype basic chatbot. In both conditions, participants were hooked up to an eye tracker while reporting. See the Appendix of this report for methodological details.

Figure 8. Static Webform v. Prototype Basic Chatbot

Static Reporting Webform – Interaction Experiences

Heatmaps depict gaze time in color gradation where red indicates the most concentrated gaze behavior and green indicates the least. Figure 9 shows the static webform’s landing page with the areas of interest (AOIs) outlined on the left and the heatmap on the right.

Gaze duration was greatest for report type – which asked people to categorize the incident they were about to report. People paid more attention – ~2.5x more time per word – to the report type than the rest of the landing page. This indicates that the first question people are asked – the type of incident they are reporting – requires higher cognitive effort, which may be a deterrent to continuing with a report.

1 As discussed in the Appendix, 19 participants were dropped from analyses. This resulted in 80 participants in the analytic sample—43 in the web form condition and 37 in the chatbot condition.
People spent the most time in the Incident Report Questions section followed by the Individuals Involved section (see Figure 10 for heat maps of each). There was a dramatic drop-off in visual attention given to other sections, as shown in Table 4.

Table 4. Time (in seconds) Spent on Each Main Reporting Section

<table>
<thead>
<tr>
<th>Section</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Incident Report Questions</td>
<td>150.12</td>
</tr>
<tr>
<td>Individuals Involved</td>
<td>110.28</td>
</tr>
<tr>
<td>Location and Time</td>
<td>32.23</td>
</tr>
<tr>
<td>Report Types</td>
<td>28.76</td>
</tr>
<tr>
<td>Report Preferences</td>
<td>22.75</td>
</tr>
</tbody>
</table>

Figure 10. Heat Maps of Incident Reporting Questions and Individuals Involved
Static Webform v. Prototype Basic Chatbot – Response Length and Sentiment
We compared word count and sentiment for the open-ended responses for two questions: request for incident details and involved individual description. See Table 5 for results by condition.

Table 5. Means and Standard Deviations by Variable and Condition

<table>
<thead>
<tr>
<th></th>
<th>Request for Incident Details</th>
<th>Involved Individual Descriptions</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Word Polarity Subjectivity</td>
<td>Word Polarity Subjectivity</td>
</tr>
<tr>
<td>Static Webform</td>
<td>27.47 (0.17) -0.04 (0.18) 0.33 (0.18)</td>
<td>13.05 (11.79) -0.06 (0.08) 0.25 (0.21)</td>
</tr>
<tr>
<td>Prototype Basic Chatbot</td>
<td>17.06 (0.13) 0.01 (0.23) 0.24 (0.23)</td>
<td>7.89 (6.18) -0.03 (0.09) 0.18 (0.20)</td>
</tr>
</tbody>
</table>

Responses from participants in the prototype basic chatbot condition contained significantly fewer words than their static webform counterparts for both questions. There were no differences, however, in polarity or subjectivity.

Static Webform v. Prototype Basic Chatbot – Willingness to Share Contact Information
There was no difference in willingness to remain anonymous between platforms.

Static Webform v. Prototype Basic Chatbot – Report Accuracy
Participants all watched the same suspicious scenario video that we created, so we were able to code the accuracy of report. We coded across 8 dimensions of accuracy, as seen in Table 6.

Table 6. Dimensions of Accuracy and Coding Notes

<table>
<thead>
<tr>
<th>Dimensions</th>
<th>Coding Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>Presence of a vehicle</td>
<td></td>
</tr>
<tr>
<td>Presence of a weapon</td>
<td>Binary: 1=accurate, 0=inaccurate</td>
</tr>
<tr>
<td>How aware of incident</td>
<td></td>
</tr>
<tr>
<td>Previously reported</td>
<td></td>
</tr>
<tr>
<td>Incident Location</td>
<td>0=inaccurate to 4=included accurate, identifiable location</td>
</tr>
<tr>
<td>Individual Description</td>
<td>Score for Accurate minus Score for Inaccurate</td>
</tr>
<tr>
<td>Incident Description</td>
<td>Accurate: 0=no accurate description – 4=accurately describe 4 parts</td>
</tr>
<tr>
<td></td>
<td>Inaccurate: 0=no inaccurate description – 4=inaccurately describe 4 parts</td>
</tr>
<tr>
<td>Others Present</td>
<td>Score for Accurate minus Score for Inaccurate</td>
</tr>
<tr>
<td></td>
<td>Accurate: 0=no accurate description – 4=correctly name each group, approximate numbers, and what each is doing</td>
</tr>
<tr>
<td></td>
<td>Inaccurate: 0=no inaccurate description – 3=inaccurately describe 2 or more of the groups/people present</td>
</tr>
</tbody>
</table>
Compared to reports submitted to the static webform, reports submitted to the prototype basic chatbot were more accurate on:

- Presence of a vehicle (accurate: 89.19% v. 55.81%)
- Presence of a weapon (accurate: 97.30% v. 58.14%)
- Previously reported incident (accurate: 91.89% v. 69.77%)

There were no differences in report accuracy between the static webform and prototype basic chatbot for: how participant found out about the suspicious incident (accurate: 89.19% v. 88.37%), incident location, individual description, incident description, or others present.

**Static Webform v. Prototype Basic Chatbot – System Trust and System Usability**

Trust scores for both the static webform and the prototype basic chatbot were both above average. Similarly, system usability scores for both the static webform and the prototype basic chatbot were above average score, which indicates that both of our reporting platforms are more usable than the typical system. There was not a difference in system-level trust nor system usability between the platforms.
BONUS SURVEY – NATIONAL CHATBOT ATTITUDES AND EXPERIENCES

This survey was not part of our proposed project (as shown on Figure 1) though the insights provided here have helped our Phase I aim of understanding how people view and use chatbots. The main aim here was to understand both how members of the public use conversational agent technology in their daily lives and how they perceive of these experiences.

We hired Qualtrics, an online survey research company, to provide a sample of U.S. adults (n=1,003) in November and December 2022. See the Appendix of this report for methodological details.

Familiarity and Frequency with Use of Chatbot Technology

As Figure 11 shows, people interact with chatbots on a regular basis. Most people (79.3%) have had experience with chatbots – about half (49.1%) said that they interact with one on at least a weekly basis and 19% use them at least once a day. Unsurprisingly, younger people report more chatbot engagement.

Figure 11. Frequency of Interaction with Chatbots

As Table 7 shows, participants use voice assistant chatbots most (78%) and robot or physically embodied chatbot least (33.5%). Most participants (71.2%) also indicated that they were somewhat comfortable or very comfortable interacting with chatbots. Younger people are also more comfortable using chatbots.
Table 7. Frequency of Interactions with Various Forms of Chatbots

<table>
<thead>
<tr>
<th></th>
<th>Multiple times a day</th>
<th>Daily</th>
<th>Multiple times a week</th>
<th>Weekly</th>
<th>Multiple times a month</th>
<th>Once</th>
<th>Never</th>
</tr>
</thead>
<tbody>
<tr>
<td>Voice assistant</td>
<td>10.3%</td>
<td>13.0%</td>
<td>19.2%</td>
<td>10.8%</td>
<td>13.5%</td>
<td>11.2%</td>
<td>22.0%</td>
</tr>
<tr>
<td>Text-based</td>
<td>4.0%</td>
<td>6.6%</td>
<td>13.3%</td>
<td>9.4%</td>
<td>24.1%</td>
<td>20.3%</td>
<td>22.2%</td>
</tr>
<tr>
<td>Interactive voice response</td>
<td>4.5%</td>
<td>6.7%</td>
<td>13.5%</td>
<td>9.4%</td>
<td>28.0%</td>
<td>21.3%</td>
<td>16.6%</td>
</tr>
<tr>
<td>Robot or physically embodied</td>
<td>2.2%</td>
<td>4.5%</td>
<td>5.1%</td>
<td>5.8%</td>
<td>6.8%</td>
<td>9.1%</td>
<td>66.5%</td>
</tr>
<tr>
<td>Virtual human</td>
<td>3.2%</td>
<td>4.6%</td>
<td>5.0%</td>
<td>6.0%</td>
<td>8.5%</td>
<td>11.5%</td>
<td>61.3%</td>
</tr>
</tbody>
</table>

Note. Table percentages were calculated w/o NA values.

Chatbots: Trust in Them, Positive Attitudes toward Them, and Intentions to Use Them

Overall, people indicate that they trust chatbots, have generally positive attitudes toward them, and intend to use them in the future. Further, there is significant variation in each of these three variables as a function of age.
REFERENCES


METHODOLOGICAL APPENDIX

Archival Analysis

Understanding Incident Types and Anonymity
Questions: What is reported to the online reporting system? What are the anonymity behaviors when reporting different kinds of information? To address this knowledge gap, we investigated the user selected incident types and whether users selecting those types chose to remain anonymous.

Methodological Approach: At the start of a report, the person must select the type of incident they are reporting. We plotted these incident types in a bar chart. To investigate anonymity preferences, we calculated the percentage of anonymous reports for each user selected incident type.

Understanding Reporter Identity and Anonymity
Questions: Who is using the reporting platform and what are their anonymity behaviors? To answer these questions, we investigated the reporters’ answers to the field “I am a:”, and whether they chose to make an anonymous report or provide contact information.

Methodological Approach: Reporters were grouped into identities according to their response to the field “I am a:” We calculated the total number of reports for each identity and calculated the percentage of anonymous reports for each identity.

Understanding Trends over Time
Questions: Are there temporal trends in the number of reports filed? To answer this question, we investigated the number of reports filed at the yearly, monthly, and weekly levels.

Methodological Approach: Reports were aggregated by week, month, school year, and calendar year. Aggregate data were then used to create time series plots to identify important or interesting trends.

Understanding Responder Usage
Questions: How are responders (e.g., BTAM members) using the system? How long did it take after making a report for it to be read by a responder? To answer these questions, we investigated the record of system usage.

Methodological Approach: Prior to 10/28/2020 there was no reliable data on system usage. The following discussion is only considering the 286 reports collected between 10/28/2020 and 1/31/2022. We began by preprocessing the record of system usage to enable analysis. This involved identifying the unique activities that could take place in the system. After identifying the unique activities, we determined which activities occurred in which report, and how long after the report was made it took for the activity to occur. Bar plots were created to investigate the prevalence of activity types within the body of reports, and average time after a report was submitted for the activity to occur. Additionally, we created a clustered network graph to identify prevalent paths of activity stages.
Focus Groups

Aim: Our main aim with the focus groups was to explore the strengths and weaknesses that threat assessment team members perceive in current tips reporting and identify potential improvements.

Methodological Approach: We conducted a series of focus groups with members of BTAM teams from August 2022 through February 2023. First, we conducted in-person focus groups with the BTAM team in Sarpy County, NE (n=22) that were broken out by focus area (i.e., law enforcement, schools, mental health services). Next, we conducted focus groups over Zoom with available members of whole BTAM teams across the state of Nebraska (n=20) and across the country (n=21) including a range of jurisdictions including ones in Alabama, Connecticut, Oklahoma, and Oregon.

National Survey

We hired Qualtrics, an online survey research company, to provide a sample of U.S. adults. We set quotas for key demographics so that the sample reflects the general population on these factors as best as possible. Data were collected in October 2022 with a total of 1,302 completed surveys.

Identifying Behavior as Concerning

Table A1. Comparing Perceptions of Scenarios

<table>
<thead>
<tr>
<th></th>
<th>Mean (SD)</th>
<th>Difference in Means</th>
</tr>
</thead>
<tbody>
<tr>
<td>Suspicious</td>
<td>3.17 (0.87)</td>
<td>Significant difference (t(3903)=8.70, p&lt;0.01), small effect size (d=0.14)</td>
</tr>
<tr>
<td>Threatening</td>
<td>3.08 (0.89)</td>
<td>Significant difference (t(3902)=23.82, p&lt;0.01), moderate effect size (d=0.38)</td>
</tr>
<tr>
<td>Should be reported to law enforcement</td>
<td>3.09 (0.94)</td>
<td></td>
</tr>
<tr>
<td>Should be reported to other authorities</td>
<td>2.67 (1.03)</td>
<td></td>
</tr>
</tbody>
</table>

Willingness to Report

Questions: Most research on willingness to report crime generally and terrorism and targeted violence specifically has focused on views of law enforcement. To address this, we evaluate prior experiences witnessing and reporting crime. We also showed participant 3 unique situations and asked about views on reporting in each.

Methodological Approach: First, we asked participants about their prior experiences witnessing and reporting crime. We then presented participants with scenario that include 3 contextual factors about reporting: whether it can be anonymous, the reporting mechanism, and who they will report to. Anonymity is binary. There are 5 reporting mechanism options (a website, a text message, a phone call, an app on a cell phone, a chatbot). There are 5 entities to report to (a community leader, a religious leader, local law enforcement, federal law enforcement, an issue-specific tip-line for that kind of incident). There were 50 possible scenarios that participants could evaluate, for example:

Reports can be submitted anonymously through a website to a community leader.
Findings: We first asked people’s hypothetical willingness to report crime, which was generally high. People indicated greater willingness to report a crime ($M=3.35$, $SD=0.75$) than a person who broke the law ($M=3.01$, $SD=0.82$) with a moderate effect size ($t(3905)=30.25$, $p<0.01$, $d=0.48$).

The experimental component focused on how anonymity, reporting mechanism, and entities to report to impact various views on reporting. However, with incident-level context (i.e., the perpetrator and the behavior) removed, none of these reporting-level factors have a relationship with any of the outcomes below.

We also asked participants about other considerations and hesitancies that they may have related to reporting, as shown in Table A2.

<table>
<thead>
<tr>
<th>Table A2. Views on Reporting</th>
</tr>
</thead>
<tbody>
<tr>
<td>I would...</td>
</tr>
<tr>
<td>report in this situation.</td>
</tr>
<tr>
<td>worry that filing a report would waste people's time.</td>
</tr>
<tr>
<td>be hesitant to report because I wouldn't want to get involved.</td>
</tr>
<tr>
<td>not file a report in case what I saw wasn't really a big deal.</td>
</tr>
</tbody>
</table>

Knowing How to Report

Aim: Despite the prevalence of public awareness campaigns aimed at encouraging the public to identify and report terrorism and targeted violence threats, recent survey research indicates knowledge gaps in where and how to report.

Methodological Approach: We asked participants a short series of questions to assess their knowledge on national reporting campaigns and to ask what they would do if they wanted to report something and how they would get information about reporting.

Perceptions of Current Terrorism and Targeted Violence Threats in the US

Questions: What groups or ideologies do members of the public perceive to be the greatest terrorism and targeted violence threats currently facing the United States today? And, how do these perceptions vary as a function of other sociopolitical viewpoints? To address this, we asked participants to evaluate the level of threat currently posed by various ideologies in the US.

Methodological Approach: We asked the following question: “Currently in the USA,_________ violent extremists pose a high risk of committing terrorism and targeted violence.” Participants were presented with 5 ideologies – Racially or ethnically motivated, Anti-government or anti-authority, Animal rights or environmentalist, Abortion-related, and Islamist – and rated each on a 4-point scale.

Findings: Overall, U.S. adults view REMVE and AGAAVE violent extremists as posing the greatest threats to the country while Animal Rights and Environmental Rights extremists are perceived as posting the least threat out of these five ideologies.
Figure A1 shows the mean level of perceived threat for each extremist ideology broken out by participants’ political ideology. One of the most interesting findings about public perception of terrorism and targeted violence threats facing the country is now political ideology conditions these views. There is no difference in perceived threat from REMVE, AGAAVE, Animal/Environmental, or Abortion-Related extremists between people with very liberal and people with very conservative political views. They only differ in their perceptions of the threats posed by Islamist extremism.

**Figure A1. Mean Level of Perceived Threat by Political Ideology**

![Figure A1. Mean Level of Perceived Threat by Political Ideology](image)

**Lab Study**

Our static webform was built to mirror existing reporting webform, which are the current industry standard. Our prototype basic chatbot was built with the Microsoft Bot Framework SDK, open-source Bot Framework Composer IDE, and primarily coded in C# and JavaScript.

Of the 99 participants, 19 were dropped from analyses: 10 static webform users and 2 prototype basic chatbot users experienced a critical technology error that made their responses unusable, 5 static webform users reported something completely unrelated to the video presented, 1 static webform user filed multiple reports, and 1 prototype basic chatbot user stated that they did not watch the video. This resulted in 80 participants in the analytic sample—43 in the web form condition and 37 in the chatbot condition.

**Prototype Basic Chatbot – System Interaction Evaluation**

**Question:** What is the frequency and nature of errors when reporting through a basic chatbot prototype?
Methodological Approach: We reviewed recordings and database entries to identify the causes of chatbot- and user-error for the 48 participants who were assigned to report through the prototype basic chatbot. We classified errors into 3 non-mutually exclusive categories: user understanding error, chatbot understanding error, and internet connection error.

Findings: Of the 48 participants, 12 (25%) failed to understand one or more of the questions presented by the chatbot (user failure), 7 (14.6%) experienced chatbot-related technical issues (chatbot failure), and 5 (10.4%) experienced internet connection failure.

Static Webform v. Prototype Basic Chatbot – Response Length and Sentiment

Question: Are there differences in response length or sentiment between platforms?

Methodological Approach: Data come from the open-ended responses for two questions: request for incident details and involved individual description. First, we consider word count for each response. We then conducted sentiment analysis using two metrics: polarity, and subjectivity. Polarity measures the positive and negative emotion in text. Subjectivity measures the extent that text contains subjective opinion v. objective description. We used t-test to examine potential differences.

Findings: Responses from participants in the prototype basic chatbot condition contained significantly fewer words than their static webform counterparts for both the “request for incident details” question (t(68)=2.45, p<0.05) and the “involved individual description” (t(67)=2.81, p<0.01). There was no difference in polarity scores for either incident details (p=0.20) or individual descriptions (p=0.15). However, there was a near-significant difference in incident detail subjectivity whereby prototype basic chatbot reporters provided less subjective reports than static webform reporters (p=0.06). There was no difference in subjectivity for involved individual descriptions (p=0.21).

Static Webform v. Prototype Basic Chatbot – Willingness to Share Contact Information

Question: Are there differences willingness to share contact information between platforms?

Methodological Approach: People had the option to provide contact information or remain anonymous.

Findings: There was no difference in willingness to share contact information between the static webform condition (51.16%) and the prototype basic chatbot condition (48.65%), X²=(1, N=80)=0.05, p=0.82.

Static Webform v. Prototype Basic Chatbot – Report Accuracy

These errors were used to exclude user information from other analyses on a case-by-case basis.

We designed the prototype basic chatbot phrasing and presentation to mimic a standard reporting webform. We made this design choice to allow for the best comparison between the two reporting mediums, but it did require marked deviation from how a chatbot would typically present questions.
**Question:** Are there differences report accuracy across the platforms?

**Methodological Approach:** Since participants all watched the same suspicious scenario video that we created, we were able to code the accuracy of report across several dimensions. We created an initial codebook that was refined. All datapoints were double-coded and, to increase confidence, all coding discrepancies were discussed until a consensus for the final code was reached. We coded across 8 dimensions of accuracy.

**Findings:**
Compared to reports submitted to the static webform, reports submitted to the prototype basic chatbot were more likely to:

- Accurately state there was no vehicle present, $X^2=(1, N=80)=10.81$, $p=0.001$, (accurate: 89.19% v. 55.81%)
- Accurately state there was no weapon present, $X^2=(1, N=80)=16.84$, $p<0.001$, (accurate: 97.30% v. 58.14%)
- Accurate state that the incident had not previously been reported, $X^2=(1, N=80)=6.08$, $p=0.014$, (accurate: 91.89% v. 69.77%)

There were no differences in report accuracy between the static webform and prototype basic chatbot for: how participants found out about the suspicious incident, $X^2=(1, N=80)=0.01$, $p=0.91$, (accurate: 89.19% v. 88.37%); location information, $z=-0.36$, $p=0.72$; information about the suspect, $z=0.90$, $p=0.37$; information about others present, $z=0.45$, $p=0.66$; or information about the situation, $z=1.50$, $p=0.13$.

**Static Webform v. Prototype Basic Chatbot – System Trust and System Usability**

**Question:** Are there differences in system-level trust and system usability between platforms?

**Methodological Approach:** After reporting using either the prototype basic chatbot or the static webform, participants completed the Trust in a Specific Technology Scale (McKnight et al., 2011) and the System Usability Scale (Brooke, 1995). Scores on each scale were compared using t-tests.

**Findings:** There was no difference in system-level trust between the platforms ($t(78)=-0.97$, $p=0.34$). Further, the trust scores for both the static webform ($M=3.84$, $SD=0.75$) and the prototype basic chatbot ($M=3.67$, $SD=0.84$) were both above average the average score of 3.4.

Similarly, there was no difference in system usability between the platforms ($t(78)=0.03$, $p=0.97$). System usability scores for both the static webform ($M=76.63$, $SD=19.20$) and the prototype basic chatbot ($M=76.49$, $SD=20.37$) were both acceptable, which indicates that both of our reporting platforms are more usable than the typical system.
Bonus Survey – National Chatbot Attitudes and Experiences

We hired Qualtrics, an online survey research company, to provide a sample of U.S. adults. We set quotas for key demographics so that the sample reflects the general population on these factors as best as possible. Data were collected in November and December 2022 (n=1,003).

Familiarity and Frequency with Use of Chatbot Technology
Questions: How often do people interact with chatbot technology?

Methodological Approach: We first asked participants how often they interact with any chatbot and provided examples. Among the people who indicated that they ever interacted with any chatbot we then asked about their frequency of interaction with each subgroup of chatbots: voice assistant chatbots (e.g., Siri, Alexa, Google Assistant), text-based chatbots (e.g., a customer service chatbot), interactive voice response chatbots (e.g., an automated business phone system), robot or physically embodied chatbots (e.g., a robot store greeter), or virtual human chatbots (e.g., a virtual companion chatbot). Finally, we asked people to rate their comfort level interacting with chatbots.

Findings: Younger people reporting more chatbot engagement ($X^2(12, N=846)=31.95, p<0.001$). Age is similarly related to comfort whereby younger people are more comfortable using chatbots ($r(843) = -0.15, p<0.001$).

Chatbots: Trust in Them, Positive Attitudes toward Them, and Intentions to Use Them
Questions: How much do people trust chatbots? How positive are people’s attitudes toward chatbots? And how likely are people to use chatbots?

Methodological Approach: To measure trust in chatbots, we used a modified version of the Muir and Moray (1996) trust scale. To measure intention and attitude toward chatbots, we used modified items from the Davis and Venkatesh (1996) scale.

Findings: Overall, people indicate that they trust chatbots ($M=3.26, SD=0.54$), have generally positive attitudes toward them ($M=3.37, SD=1.05$), and intend to use them in the future ($M=3.55, SD=1.05$). Further, there is significant variation in each of these variables as a function of age.
### Table A3. Trust, Attitudes, and Intentions to Use by Participant Age

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Mean</th>
<th>SE</th>
<th>F</th>
<th>P-Val</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Trust in Chatbots</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18-44</td>
<td>414</td>
<td>3.35</td>
<td>0.03</td>
<td>13.13</td>
<td>&lt;.001***</td>
</tr>
<tr>
<td>45-64</td>
<td>271</td>
<td>3.18</td>
<td>0.03</td>
<td></td>
<td></td>
</tr>
<tr>
<td>65 &amp; up</td>
<td>162</td>
<td>3.15</td>
<td>0.04</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Attitudes Toward Chatbots</strong></td>
<td></td>
<td></td>
<td></td>
<td>46.37</td>
<td>&lt;.001***</td>
</tr>
<tr>
<td>18-44</td>
<td>414</td>
<td>3.71</td>
<td>0.05</td>
<td></td>
<td></td>
</tr>
<tr>
<td>45-64</td>
<td>271</td>
<td>3.09</td>
<td>0.07</td>
<td></td>
<td></td>
</tr>
<tr>
<td>65 &amp; up</td>
<td>162</td>
<td>2.99</td>
<td>0.08</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Intention to Use Chatbots</strong></td>
<td></td>
<td></td>
<td></td>
<td>19.75</td>
<td>&lt;.001***</td>
</tr>
<tr>
<td>18-44</td>
<td>414</td>
<td>3.78</td>
<td>0.05</td>
<td></td>
<td></td>
</tr>
<tr>
<td>45-64</td>
<td>271</td>
<td>3.37</td>
<td>0.07</td>
<td></td>
<td></td>
</tr>
<tr>
<td>65 &amp; up</td>
<td>162</td>
<td>3.28</td>
<td>0.09</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. *p < .05. **p < .01. ***p < .001