EXPECTED DATA

For this study we will collect three waves of survey data and one wave of qualitative interview data from an international military base site visit. The survey will include a survey of veterans within the United States (wave I) – former members of the military who are not currently on active duty. This survey will be distributed by a reputable survey sampling firm, who will recruit participants. The second survey will be conducted on active duty military members within the United States (wave II). The third survey will be conducted on active duty military members around the world (wave III). In Waves II & III, from each participating base, we will randomly select a subset of active duty members to participate in the survey. All individual survey responses will be de-identified. The datasets will include basic demographic information (e.g. race, age, gender, etc.), responses to questions about the participants' political and social attitudes and behavior, and data about individual's experiences with offending and victimization during their time in the military.

For these surveys, no information capable of uniquely identifying the respondent will be collected as part of data collection. Should local sample vendors include any unique identifiers (such as respondent IDs), these will be removed from any dataset. We plan to share individual-level respondent data, so that other researchers can conduct their own original analyses and verify our analyses. While individual demographic data could, in theory, be used to identify members from a particular base, we will not include information about which base individual respondents are from in the dataset, outside of an anonymized base indicator (base A, base B, etc.).

The qualitative interview data will be gathered from stakeholders during a site visit to a military base abroad. The primary purpose of these interviews will be to gather information that can contextualize the quantitative survey findings. Stakeholders will include military personnel stationed on the base as well as local officials who may interact with the base and/or servicemembers when crime or victimization occurs. An interview protocol will be developed for site visit interviews and responses will be recorded by hand. Handwritten notes will then be transcribed into an electronic format for analysis. Qualitative data management software, ATLAS.ti, will be used to facilitate content and thematic analyses. A summative report will be produced based on the interview data.

MANAGEMENT AND MAINTENANCE OF DATA

Data will be stored in a number of places, each requiring different security procedures. Survey data will be collected using Qualtrics, and raw response data will be retained on Qualtrics' secure servers. Although we will not collect any data capable of uniquely identifying a respondent, sample vendors usually require a unique identifier (such as a respondent ID provided by the vendor) to verify completion. These data will be kept on Qualtrics' secure server; there is no way to remove these data. No one except the Co-PIs will be given access to the data stored by Qualtrics. Data will also be downloaded, cleaned, formatted. These datasets will have no uniquely identifying information included – that is, any unique identifiers from the data stored in Qualtrics will be removed. Local copies of these datasets will be maintained on computers used by the Co-PIs and the research assistants.

Interview notes from the site visit will be recorded by hand and then transcribed into an electronic format. The PI and research assistants will conduct the transcription. The original notes will be stored in a locked filing cabinet and maintained for three years, at which time they will be destroyed and only the electronic versions retained. A unique identifier will be given to each interview and no personally identifying information pertaining to the interviewees will be recorded. The unique identifier will be used for the purpose of distinguishing data gathered from military personnel/stakeholders compared to data gathered from host country stakeholders. The transcribed notes will be uploaded into a qualitative data management software program, ATLAS.ti. Local copies of this dataset will be maintained on the computers used by the PI and the research assistants. All computers will be password protected and maintain up-to-date antiviral software to avoid security breaches related to malware.

FACTORS POSSIBLY IMPACTING ABILITY TO MANAGE DATA

The primary issues we face in managing and distributing our data are related to confidentiality. We are collecting public opinion and self-report offending and victimization data, which will include opinions and experiences that respondents would prefer not become publicly known. We have protocols in place, including limiting access to the Qualtrics data, and using password- and antiviral software-protected laptops to prevent any data breaches. Respondents will be informed of the (very minimal) risk of identification due to security breaches as part of the process of gaining informed consent.

MECHANISMS FOR DATA SHARING AND METADATA

We will use Harvard Dataverse to distribute data to the public. We will also use ScholarWorks from Boise State. Data will be embargoed until a paper using each dataset is accepted for publication; this is standard practice in our field. Harvard Dataverse stores a great deal of metadata, such as collection dates, associated publications, sample vendors, etc. We plan to take full advantage of these capabilities and enter any and all metadata we can, so that other researchers can have access to it. We will work with librarians at Boise State to ensure we have appropriate metadata available for each dataset. Additionally, we will develop a codebook for both datasets, with variable descriptions, coding rules, and any relevant notes that will be stored along with the dataset. The qualitative data (interview notes from site visit) will be disseminated in the form of a summative report. Given the small sample size and need to maintain confidentiality, the raw interviewer notes will not be made publicly available.

OTHER DETAILS

Period of data retention and distribution embargo. We will follow the rules of Boise State and store raw and formatted data for a minimum of three years. Qualtrics data will also be stored for three years, barring any policies maintained by Qualtrics to the contrary. All quantitative data will be made publicly available via Harvard Dataverse as soon as the paper using the datasets is accepted for publication, or within one year after the grant period ends, whichever is sooner.

Data formats and dissemination. All data will be stored either in secure cloud servers (including but not limited to those maintained by Qualtrics and Dropbox) or on password- and antiviral software-protected computers maintained by the project team. Uniquely identified data will only be accessible by the Co-PIs. Quantitative data will be made publicly available (after an embargo period) via Harvard Dataverse. Datasets published through Harvard Dataverse will be released under a modified Creative Commons license. Additionally, descriptive metadata records and permanent DOIs will be created for each dataset to ensure proper citation and permanent retrieval of the materials.

Data storage and preservation of access. All data, including unpublished data, will be stored on a local, networked drive managed by Boise State's Office of Information Technology. Drives are Access Control List (ACL) protected on high availability NAS and SAN devices hosted on secured networks accessed by users and applications via CIFS (SMB), NFS or iSCSI protocols as appropriate. OIT manages daily backups and utilizes dedicated disaster recovery/business continuity facilities off site. Should the PIs leave Boise State University during the study period, they will retain these data as intellectual property. The PIs will work with their institution to ensure continued access and storage of these data meet the University's data management requirements. Research data generated by the project will be openly shared via ScholarWorks, Boise State's institutional research repository, unless prohibited by the PIs' respective Institutional Review Boards or the original source of any aggregated state-level data. Managed by Boise State's Albertson's Library, ScholarWorks is a searchable Open Access hosting platform for data and research outputs. Datasets published through ScholarWorks will be released under a modified Creative Commons license. Additionally, descriptive metadata records and permanent DOIs will be created for each dataset to ensure proper citation and permanent retrieval of the materials.