Special Issue Article

JDMS

Journal of Defense Modeling and Simulation: Applications, Methodology, Technology
2019, Vol. 16(4) 285–295

DOI: 10.1177/1548512919831363
journals.sagepub.com/home/dms
SAGE

Effective Use of Public Perceptions in Assessments

Joseph Lindquist[1], David Peng[2], Shawnette Rochelle[3], Dusty Turner[1], and Matt Warshaw[2]

Abstract

Over nearly two decades of war, the Department of Defense has spent tens of millions of dollars to survey the public in war-torn areas such as Iraq, Afghanistan, Philippines, Syria, and the Horn of Africa. Most often these surveys are conducted by a private polling firm, which, in turn, conducts the field work. Since field work occurs in contested or enemy-dominated terrain, direct supervision of the field work is not possible. This lack of oversight is concerning to senior leaders, who may inform decisions and assessments using this polling data. This article reviews some best practices for the expeditionary operations research system analyst when faced with the task of planning, contracting, executing, analyzing, and reporting public perception polling data in conflict affected areas. Additionally, this work codifies some of the insights uncovered during a community of practice meeting in March 2018 to include the use of latent data sources such as social media. The purpose of this effort is to provide analysts the tools to increase senior leader confidence in the data that is oft relied upon to inform operations.

Keywords

Public perception, survey, best practices

1. Introduction

In early 2018, the Military Operations Research Society (MORS) convened a special meeting with a single purpose: to advance the professionalism of assessments conducted in support of contingency operations. One working group of more than 20 multiservice, multidisciplinary, and multinational professionals discussed and distilled ''best practices'' for soliciting public perceptions in conflict affected areas. Surveys have been used effectively for decades to inform military decisions regarding doctrine, organization, training, materiel, leader development and education, personnel, and facilities.[1]

The US Army Training and Doctrine Command Analysis Center (TRAC), recognizing the complexity of developing and executing quality surveys, developed a code of best practices for survey efforts. Their technical report provides a comprehensive procedure to plan, design, develop, administer, and analyze surveys tailored to their mission. Their report is a solid foundation and what follows in this article are a collection of salient findings as well as key reference materials for a practitioner responsible for understanding public sentiment in support of contingency operations. Sources beyond the military for providing explanations and examples of frameworks, as well as guidelines for how to implement survey research and analyze the results are included in this paper.

The US Government Accountability Office (GAO) has a ''methodology transfer paper'' dedicated to developing and using questionnaires tailored to their mission to assist ''congressional decision makers by furnishing analytical information on issues and options under consideration.''[2] The GAO paper provides a robust framework for a study

[1]United States Military Academy, West Point, NY, USA
[2]D3 Systems, Tyson's Corner, VA, USA
[3]Center for Army Analysis, Fort Belvoir, VA, USA

Corresponding author:
Joseph M. Lindquist, United States Military Academy Thayer Hall, Bldg. 601 646 Swift Road West Point, NY 10930, USA.
Email: joseph.lindquist@westpoint.edu

including survey research and sound advice on how to minimize sources of bias and implement quality assurance procedures.

The American Association for Public Opinion Research (AAPOR), one of the largest professional associations of survey practitioners, maintains a set of recommendations intended to produce the best survey possible.[3] It should be noted that these recommendations are nested within AAPOR's code of professional ethics and practices—highlighting the virtue of conducting, analyzing, and reporting sound opinion research. AAPOR suggests that quality surveys:

1. Have specific goals.
2. Consider alternative data beyond the survey instrument itself.
3. Select samples that well represent the population.
4. Use designs that balance costs with errors.
5. Employ appropriate question format and wording.
6. Pretest questionnaires and procedures.
7. Carefully train interviewers.
8. Check quality at each stage of implementation.
9. Minimize non-responses while treating human subjects ethically.
10. Use appropriate analytical and reporting techniques.
11. Develop and fulfill pledges of respondent confidentiality.
12. Disclose methods to allow for evaluation and replication.

To establish a common, broad framework that, in spirit, encompasses both the TRAC and GAO guidelines, we will reference these recommendations as (AAPOR # xx) when similar findings from the working group are discussed.

Guiding principles clearly exist for developing surveys to capture public perceptions in any application (public or private), but there are significant challenges that arise when using surveys to assess military operations. First, the complex strategic objectives of military combatant commands typically involve close partnerships with the host-nation and interagency organizations. The plans developed within these partnerships require nuanced executions that take into consideration the needs (and often the approval) of all stakeholders before work can go forward. This means the analyst may need to work on many fronts at one time to satisfy the needs of the military and other decision-makers in the operating environment. Sizable geographic territory and dynamic security conditions often force choices to be made between methodology and ground reality. Routinely, military forces rely on both quantitative metrics collected by tactical operators and public sentiment data. A careful review of both sources of information provides greater opportunity for insights to the command.

Public perception surveys taken from probability samples that are representative of the population of interest become powerful research tools. Because military commands cannot interview every member of the host nation population, surveys provide a cost-effective way, in a resource-constrained environment, for a command to gain insight into the sentiments of the local population. Analysts can mine the survey data for an understanding of perceptions across various dimensions (e.g. demographic group, geographic region). These analytical insights empower leaders to make data-informed policy and resource allocation decisions.

LaRivee notes that public perception studies have value in assessing counterinsurgency operations because questions of motivation, satisfaction, degrees of trust or fear, as well as intentions regarding future actions are difficult to measure by monitoring actions alone.[4] The ability to gain an understanding of public sentiment and behavior expands the ability of the command not only to assess past or on-going operations, but also to inform future actions. LaRivee contends that, despite sources of error such as bias resulting from respondents being afraid to answer honestly, or the inability of researchers to follow the most rigorous of academic standards in the operating environment, the surveys are still able to provide a valuable component of any assessment framework when used properly.

The logical question then follows—if there is value in having the data, then how do practitioners use surveys properly to assess military operations? Kulzy and Fricker contend that good surveys result from ''excellence in design (both questionnaire and sample), fielding, and analysis.''[5] In their discussion, they present a framework that clearly outlines *how* to implement many of the AAPOR recommendations and avoid common pitfalls often observed in the process. Further, their work provides practical advice in presenting results to decision makers.

The Center for Army Analysis maintains a Deploying Analyst Handbook that provides pragmatic advice to Operations Research / Systems Analysts deployed in support of deploying headquarters.[6] The handbook includes a brief chapter on public perception survey development and analysis designed to ''operationalize'' concepts found to be effective over more than a decade of persistent conflict.

Given the robust collection of resources available to analysts, tailored to virtually any environment, another logical question might be, ''Why create another *how-to* survey document?'' It should be noted that this document is not:

- a tutorial that will create a well designed survey.
- a technical document that develops the mathematics to design and analyze survey data.
- final best practice for all time (techniques change, data sources change, people change)

Instead, this document attempts to summarize enduring principles and nascent techniques to quickly inform an analyst about quality standards and best practices. The article offers a comprehensive perspective on how to effectively design, execute, and deliver results of a study gathering public perceptions. Principles are written to assist assessment professionals who may have an analytical background, but lack specific experience dealing with challenges associated with gathering public perceptions, especially in a conflict environment. Findings have been derived from MORS working group participants that included survey practitioners, operations research analysts, survey contractors, statisticians, social scientists, and decision makers who use perception data. The authors recommend options for analysts preparing to manage a perception survey with basic options depending on whether they have 2 hours, 2 days, or 2 weeks to get themselves moving forward.

The paper is organized to speak first about design aspects of surveys and effective interaction with fieldwork teams. We then discuss techniques to effectively present and deliver analysis to decision makers and conclude with recommendations for new practitioners to improve their skills with designing and supervising survey research.

2. Effective Survey Design

Survey design is complex and requires that analysts approach it with a clear desired outcome (AAPOR #1). Spending time with senior leaders to thoroughly understand how they envision using the information, and then designing the survey with enough specificity and flexibility to fulfill that vision, is key. A survey that will inform campaign assessments should be designed to ensure information collected in future years will both support analysis of trends as well as provide opportunities to address new issues that may arise.

2.1 Mixed method design

A single source of data or single method for collection is rarely able to satisfy all information needs. Using a variety of methods can attack a problem from multiple angles and offer deeper insights. A mixed methods approach leverages the value of quantitative data to provide an overview of *what* is happening and qualitative data to provide decision-makers with greater perspective regarding *why* something is happening. After soliciting and receiving input from senior leaders, analysts must then collaborate with stakeholders to design and implement research instruments that employ a mixed methods approach. Collecting and analyzing both types of data has the potential to yield results that serve as key benchmark indicators to assess progress against political, economic and/or military objectives, while also providing leaders key contextual insights.

Successful implementation of mixed methods requires analysts to consider which collection methods are best suited to the task. While quantitative data can be collected through several means (e.g. traditional surveys, open source publications, operators on the ground), qualitative data collection may be less structured in an effort to create space for greater nuance to emerge from respondents. The use of structured or semi-structured interviews or focus groups aims to allow participants to bring issues to the surface that leaders may or may not be aware of in the theater of operations. This additional context may then directly, or indirectly, inform senior leader decisions.

In summary, a successful mixed methods approach provides analysts a robust set of tools for translating data into actionable information for senior leaders (AAPOR #2,4). It is difficult to build a house if you have only one tool in your toolbox. Mixed methods approaches allow you to use different tools depending on the type of problem you are trying to address. Typically, military commands contract local market research firms to co-design and field research projects. An effective survey design collaboration with locally contracted market researchers should produce a survey that meets senior leaders intent while also mitigating bias. In order to achieve this, several factors should be addressed in sequence.

1. The method of collecting data must be identified since this will inform the development of the questions.
2. The wording of the questions matters. Open-ended questions should be designed to elicit a wide range of responses. Closed-ended questions, on the other hand, should be drafted to be interpreted similarly by a wide range of respondents. The local research firm can make valuable contributions in this domain to ensure the appropriate terms are used during question formulation to ensure that respondents are able to offer their opinions about issues and concepts based on a similar understanding of the situation. (AAPOR #5)
3. The questions must be sequenced to ensure bias is not introduced by information provided through the text of the previous questions.
4. The questionnaire must be translated to the host nation language and then back-translated to English to confirm the wording correctly conveys the desired sentiment in the host nation language(s). (AAPOR #8)
5. The survey must be pre-tested to validate its design. Issues that emerge during the pre-test can be addressed prior to the full fielding of the survey. Pre-testing is another way to ensure responsible

> stewardship of financial and personnel resources while providing oversight of how the research agency executes the contract. (AAPOR #6)

The preference with the quantitative component of survey data collection is to use probability based methods. This means that the design of the sample selection is random yet reproducible. Furthermore, the target population as a whole should have a known, non-zero probability of selection. This allows researchers to accurately estimate sampling error, conduct design-adjusted statistical analysis and compare outcomes with other probability-based samples. In practice, non-probability methods are often cheaper for field partners to implement, so our recommendation is to explicitly request probability-based designs when budget and timeline allow for it. If high quality sampling frames of the individuals comprising the target population exist (for example, list of program participants) the design can be as straightforward as a simple random sample. However, for large-scale face-to-face nation-wide studies in developing or conflict environments, multi-stage complex samples using area-based frames are more the norm.

After the survey has been fielded and the data has been delivered by the contracted firm, analysis of the results can begin. It is important to remember that probability samples do not always result in samples that perfectly match population benchmarks, such as the latest census figures. Analysts need to note that distortions can occur in surveys from a variety of sources, including differential sampling probabilities and non-response.[9] The firm fielding the survey will likely address the latter two sources during the survey design and subsequent weighting of the data. As a simplified example, a final data set may have had a split of 60% men and 40% women, when the census for the country suggests it should be 52% men and 48% women. This can be corrected by weighting the data. Analysts must confer with the research firm to ensure a comprehensive understanding of all items in the data file, and to re-confirm the survey methodology and any post-field modifications to the data such as weighting. This allows the analyst to offer conclusions and recommendations based on the data and any of its limitations. While we will not dive deep into the theory and technical elements of weighting, a brief overview of common stages to survey weighting will follow. We provide several citations that analysts may investigate more in depth.

2.2 On weighting

Generally, there are three stages to the construction of survey weights: design (also known as sampling), non-response, and calibration.[7]

Design weights adjust for the probability of selection during the sampling process. For instance, a common survey design practice is to over-sample minority group members. Oversampling facilitates the gathering of more precise information from minority groups that may otherwise be represented as more monolithic based on the proportion initially sampled at a national level, but should be adjusted for through design weights when looking at the data combined with other portions of the population sampled under different probabilities. In practice, especially in multi-stage complex samples that use area frames, sampling probabilities of individuals will vary due to household size, cluster sizes, and sample allocation by strata.

Non-response weights adjust for unit non-response. If respondents of the survey differ from those that choose not to participate, there may be bias in the estimates from the resulting sample. Generally, the lower the response rate of a survey, the higher the chance of non-response bias.

Calibration weights adjust the weighted data to match an external source, or benchmark. In practice, official census or within-country statistical organization estimates are used to adjust the sample. Calibration can be done through post-stratification or raking, depending on the granularity of the benchmark data and resulting sample composition by benchmark cells. For more information, we recommend reviewing Kolenikov's work on survey calibration.[10]

However, weights also introduce added design effect to the data and its resultant analysis. Thus, while weights are often necessary to reduce bias, they can further complicate complex survey design and the analysis of survey results. Principally, this results in an increase of the standard error of the sample statistics, causing a loss of precision. That loss of precision leads to greater variability. Failure to apply survey weights may result in biased results, so there is a tradeoff (AAPOR #9,10,12) Analysts considering the decision to weight or not weight should conduct the analysis both weighted and unweighted in order to observe the difference in the results and to account for the underlying drivers of the difference. Further, an analyst could refer to the comprehensive considerations provided by Lavallée and Beaumont.[7]

In practice, especially in international survey research, not all three stages are always calculated and combined into a single weight for end-consumers. This can be for a variety of reasons, including, but not limited to, inability to calculate the weights (or a only being able to calculate certain stages), lack of recent and reliable external benchmark statistics (for example, an outdated census), and purposely not providing weights depending on the design and resulting sample. However, we always recommend that you be involved in the weighting process or that you require detailed documentation on how the weights that accompany your data were calculated. Furthermore, if the

research firm that collected the data for you has a methodologist or sampling statistician, we recommend that you request to consult with him or her directly.

3. Practical Advice for Supervising Survey Efforts

Close collaboration with the fieldwork contractor from the beginning to the end of survey design is integral to achieving the desired outcomes. It is not enough to email a list of specifications to a research contractor and hope for the best. Analysts, to the extent possible, should be monitoring progress at each phase of the project from questionnaire design, translation, pre-testing, training, field implementation, field quality control, data cleaning, data quality control, and weighting. The analyst should endeavor to maintain a dialogue with the research contractor even post-delivery to ensure proper use of the data. This continued dialogue will grant analysts additional understanding of how the survey methodology was implemented in reality on the ground, and what it may mean for the results. It will also provide both the analyst and the contractor the chance to identify, and build from, lessons learned for potential future iterations of survey fielding.

3.1 Contractor interaction

Experiences by many working group members responsible for a large-scale survey revealed a common theme, namely that contractors responsible for executing survey efforts generally want to satisfy their ''customer''. The authors of this article consist of both sides of the contract, the ''customer'' and the contractor. For conciseness, we will refer to the ''customer'' as the contracting officer's representative (COR)—often an analyst—for the remainder of this paper. What follows in this section are some pragmatic suggestions for the COR to maximize quality of the end product.

When executing a large-scale survey, effective supervision of a contract requires the COR to:

- Be specific in the request for proposal and contracting phases concerning deliverables expected from the contractor. This will ensure that contractual obligations are fulfilled but will also give the contractor clear guidance on what they are bidding and expected to do on their end. This clarity will help keep the project on track and avoid schedule delays due to changes in requests mid project.
- Establish a schedule of deliverables for the instrument, sampling plan, translation, pre-testing, fieldwork, and delivery of data. This schedule should be complete with intermediate deadlines that take into account the cultural and religious holidays in the country of interest. The COR should also keep in mind seasonal weather that may affect field operations, such as a monsoon season or heavy snow in winter.
- Plan for interaction between the COR and contractor on a regular basis, and even after the contractor has delivered the data and other required deliverables, such as a methods report. It may be necessary for the COR or other analysts to interact with the contractor during the analysis phase to obtain clarifications about the data or methods.
- Require detailed methodology documentation that fully describes the design and implementation of the project.
- If possible, supervise training of fieldwork. These may be held at central locations where many (if not all) supervisors and some key interviewers attend. This effort should be discussed with the contractor to ensure COR presence does not bias the training/data collection in some unwanted way. Safety for the field team should also be a consideration if the COR's presence increases security risks.

3.2 Effective oversight—A contractor's perspective

The importance of effective communication between a COR and survey contractor cannot be overstated. This communication begins with the COR clearly defining the analytical goals of the study. In practice, these goals are likely to be numerous and potentially at odds with a well-crafted and executed survey. Working group members lament the numerous cases where a sponsor requested a survey designed to accomplish myriad tasks, resulting in a questionnaire that induced respondent fatigue (too long) and ultimately decreased the quality of the data. When goals are numerous, the COR should rank order the objectives. If the COR can clearly articulate the priority, there is a much higher chance the contractor can design a methodology to ensure that most important goal is addressed. If there are too many stakeholders/objectives, without clear guidance, a study design may be modified so many times that it may ultimately fail to address any key objectives adequately.

If the contractor specializes in scientific public opinion research, they are likely to have staff (survey analyst, survey methodologist, survey statistician) that will be able to work with the COR on various aspects of the overall design to the instrument, sampling methodology, and/or analysis. If they are a smaller firm and/or focus on domestic market research, they may be more limited, and it is likely that the COR will have to lead on numerous technical elements of the data collection process. If that is the case, here are some key things we recommend:

- Clearly identify the source data (sampling frame) from which the sampling is being conducted to select the target population and assess whether the quality is suitable for your needs. This is the foundation upon which the sampling methodology is being designed and if it is not adequate it may lead to incorrect results and findings.
  - Is the source of the sampling frame verifiable and from a reputable source?
  - Is the source up-to-date?
  - Does the source fully capture your target population of interest?
- Request that the contractor perform an honest assessment, using the sampling frame, of areas that are accessible possible for inclusion in the sampling. In conflict affected areas, security concerns may preclude even local field teams from accessing the population of interest. Ideally, contractors should state this in the proposal, but sometimes firms claim an ability to sample the entire country in an effort to win the contract, only later to exclude areas the COR is interested in. If coverage of a particular area is key to a COR's objectives, it is essential to know if that area is accessible. This process will also give you a clear definition of the population to be sampled and potentially help avoid or minimize common in-field decisions that are made such as substitutions for one sampling area over another.
- Request step-by-step information about how the sampling stages will be conducted. If possible, have the contractor show you the sampling process in person. Just because the desired technical terminology is used in the proposal or during communication does not mean that it is actually being performed correctly.
  - Although one can check if the sampling stages are being conducted properly remotely, it requires that the contractor share the entire sampling frame, which is unlikely if it distinguishes that firm's capabilities from its competitors. Access to a good sampling frame may be something the firm built or purchases, and they will not want to give it away. If they are unwilling, there are certain things we recommend you check or request to be provided when overseeing international face-to-face surveys. While this takes time, it will allow you to appropriately account for the design in the weighting and analysis phases and ensure the design is actually being properly implemented.
    * Double (or triple) check that the in-country vendor is using the same definition of stratification probability proportional to size (PPS) clusters sampling. These often get confused in practice. A good starting point, even if the text on the proposal sounds technically correct, is to ask, ''What is the first unit you are randomly selecting, and how?''
    * Ask for a table showing the proposed sample distribution by agreed upon strata. Check to see if certain stratum do not have any clusters allocated to them after rounding (we recommend at least two primary sampling units for variance estimation purposes), as that essentially excludes what tend to be smaller administrative units from having a chance of inclusion in practice.
    * Most surveys tend to be multi-stage cluster samples using area frames, such as administrative units officially recognized and estimated by country's central statistics office. Many designs use PPS cluster sampling of primary sampling units (PSU). In turn, in-country vendors may have workbooks set up to perform systematic PPS cluster sampling using something such as population as the measure of size. If possible, ask for this workbook or proof of sampling. If not, ask that the following be documented:
      - The measure of size—is it total population, 18+ population, or something else such as households?
      - Request the measure of size for all selected PSU.
      - Is the frame sorted, and if so how is it sorted prior to sampling?
      - The step size and random starting number (or random seed if using software)
      - The total number of PSU in each stratum
- Assess the quality control (QC) the contractor has proposed, and request more if you do not think it is sufficient to ensure high quality data. The goal is to ensure that the contractor is performing the level of quality control agreed upon. For example, are back checks to interviews being conducted as discussed?
  - If contracting multiple surveys of the same population, with the same contractor, there is an opportunity to do a post-field QC assessment of interviewer/supervisor performance compared with his/her peers. This, combined with other logic checks, can help you spot potential anomalies and investigate in preparation for the next study.
  - When possible, for face-to-face surveying, use computer-assisted personal interviewing (CAPI). The software, when programmed

correctly, will help with more complex instruments' filtering and will give you additional data and QC options.
- Set the training according to your standards, beyond what the contractor does by default. Attend if possible and appropriate.
- Require paradata and/or contact data in addition to the final completed interviews. This will help you investigate whether or not the contractor implemented the agreed upon methodology.

4. Latent Sources of Perception Data

Two of the main challenges analysts have with conducting assessments of operations and public opinion is having the data available and also the ease of accumulating necessary data to conduct assessments. Once a commander determines that he/she needs this type of information, there is often a lag period from when the need was defined to when the appropriate data to prepare the assessment is available. This lag can be frustrating because the commander may need to make decisions based on gut feelings and the gut feelings of their staff. Another option is to tour the public landscape and quiz the local population encountered. This option is time-consuming and could be severely biased because commander will only receive a few data points over a small number of locations. Depending on who the commander speaks to, the findings could be severely skewed by whoever happened to be available when the commander visited.

In practice, it may take weeks or months to properly design, contract, execute, and analyze a public perception survey. For a decision maker, this may be too lengthy. Analyzing what we will call ''latent data'' may be the only way to get a pulse of the public sentiment. This working group discussed two separate techniques to analyze public sentiment and the pros and cons of each method.

4.1 Harvesting sentiment from Twitter

The first method is analyzing data from Twitter. ''Scraping'' data from Twitter provides an almost immediate accumulation of thousands of opinions of the local populous. When scraping is targeted to the appropriate location over a specific time period with appropriate key terms, an analyst can zero in on the issues related to the population that he/she cares about. There are several ways to analyze the data to inform the commander of the populaces concerns. Querying the most used terms can shed light on the populations largest issues of concern for the commander to address and influence policy decisions.

An analyst can also take a more technical look at the information and apply sentiment analysis. While there are many different sentiment packages that glean and score sentiment from text, the general idea is the following:

- Break the text into single word tokens with a time stamp. Common sentiment-less words are removed (i.e. a, an, the, etc.) and each remaining word is scored according to a predetermined sentiment score determined by other researchers. Other ways of scoring the sentiment include bi-gram analyses when words are analyzed in pairs and also word cluster analysis where words are amplified/deamplified or negated depending on each word in its vicinity. Each method has its pros and cons.
- Once sentiment is scored, an analyst can either aggregate the analysis by tweet or over a certain time period (i.e. minute, hour, day). This type of aggregation can provide the commander a view of how the sentiment of the populace in a certain area changes over time about a certain topic.

This type of analysis is beneficial when you have no other means of evaluating public opinion, but it is not without drawbacks. Analyzing Twitter data biases your sampling to only those people who have access to Twitter, and who are willing to make statements on Twitter. Often, Twitter users have strong opinions and do not always reflect the population at large. Also, in an adversarial situation, Twitter bots can strongly influence analysis. Analysts should take steps to identify and remove these Tweets from the data. Using Twitter data also biases your sample to include only those who have access to a computer/cell phone, internet, and electricity, which may be a significantly limited group of people in a disaster or conflict environment.

4.2 Harvesting sentiment from news articles

Another mechanism for analyzing latent data is to apply similar sentiment analysis to news articles. Scraping data from news aggregation sources like Google News can quickly provide a way to accumulate articles, headlines, time stamps, and sources of information that media outlets publish. Analyzing news may prove worthwhile to a commander as an optic on public opinion and how particular sources aim to manipulate public opinion. Once the news is accumulated, an analyst can apply any number of sentiment analysis techniques discussed in the previous section about Twitter. A unique feature of analyzing news from a news aggregation source is that the analyst can attribute sentiment to different sources. The analyst can identify sources producing high and low volumes of news, as well as the most positive and negative sources. This can be very helpful to the commander when taking special measures to influence news organizations in their region of concern.

Just like with Twitter analyses, there are drawbacks to analyzing news. An analyst should take special measures

to ensure that news keyword searches are not biased by any particular users news preferences. Specifically, Google News will provide the user more information from sources it already knows the user is likely to read. To prevent this from happening, we recommend that the analyst be logged out of any particular account when accumulating data from a news aggregation source and using features that block cookies and other tracking that might start to influence what is presented to the analyst. For example, Google has a mode known as ''Incognito'' that may help to reduce this bias. In addition, unless the analyst pulls the news every day, the analysis will suffer from a recency bias. Where Twitter will provide all the tweets over a specific timeframe, news aggregation sources will provide the most recent articles and the most popular articles over a time span. This will bias the result in a specific way unless the analyst does a pull of news articles on a regular and frequent basis.

4.3 Presenting latent data

It is important for an analyst to properly caveat analysis on public perception that is derived from latent data. Until we can conduct more research to validate the accuracy of public perceptions gleaned through latent data, we must acknowledge it for what it is: the best assumption about public perception given the data available in the shortest amount of time possible.

5. Delivering Analysis

If poorly delivered, even the best analysis will fall short of its full potential. A survey that was well designed, with all of the considerations presented in this paper still needs to be distilled to the most salient findings. The analyst's task is to interpret the results in an effort to ''operationalize'' the findings for the decision maker. The analyst needs to consider how to boil all the data down to key findings for the commander, but then dive deeply into the broader data to find insights for others working on the problem set as well. A strong visual presentation is also essential. Kulzy and Fricker provide some ''do's'' and ''dont's'' that include advice to present univariate results, develop handouts, and anticipate logical follow-up analysis that is important to contextualize results.[5] Additional considerations follow.

5.1 "Morselizing"

In many nations, the US Department of State executes public perception surveys to capture longitudinal perceptions that inform policy. Often, this survey is conducted once per year, yet results are delivered in several technical reports, separated in time, that ''morselize'' the analysis into digestible portions. The authors contend that this technique maximizes the impact of the analysis by focusing on a particular facet of the data, allowing stakeholders to consider implications before moving on to another facet. To do this effectively, we suggest the following:

- As time and duties permit, analysts should develop smaller reports focused on a particular facet of the operational environment (geography, governance, development, violence perceptions, etc.) This technique allows for different staff to see the results in a context that informs their individual work. Analysts should consider how a smaller, more focused, report might augment the work of other staff, and thus help inform operations and plans. The analysis might prod follow-on analytical questions, relevant to filling a gap in understanding.
- Analysts should look for ways to weave survey results into the normal battle rhythm of the command they serve. Suppose planning is underway for an operation and perception data from the target area are available. An analyst should consider how the perception data could assist planners and units in understanding the operational environment. We recommend that analysts do not treat perception surveys just as stand alone analytical products to be shown once after fieldwork is complete, but also to weave the findings into daily briefings and other published documents.
- Analysts should consider creating interactive applications that have the ability to sort and summarize by key demographic data, location, question, etc. With the proliferation of interactive applications such as Shiny Apps, analysts should consider how to leverage these to make the most of the survey. This technique may allow for the widest use of the data, as it frees the analyst from guessing what might be of use to a particular staff section. Allowing the subject matter experts to query or filter the data they deem important can free up the analyst to focus on other tasks.

5.2 Other delivery considerations

Analysts should tailor results of perception data to the intended audience. General best practices discussed by the MORS working group include:

- Customize analysis to the decision maker and socialize results with key staff elements. Account for the style and preferences of the decision makers

the analysis is being prepared for and adjust accordingly.
- Anticipate follow-on questions of the audience to key findings. This might include breaking down responses by gender, ethnic group, geography, or other demographic considerations that may be relevant to different staff.
- When presenting results, provide quality control measures to the audience in layman's terms. While the analyst should be able to provide details on all facets of quality control measures employed, a full methodological brief should not be the focus of a results oriented briefing.
- Dial down the use of technical terms and focus on the insights found in the analysis. Analysts should be prepared to answer questions about any facet of how the study was conducted, quality control measures employed, and conclusions drawn, but be sure to start wherever your audience is starting from laymen to doctoral experts on methodology.
- Do not hide any of the products shortcomings. The MORS working group referred to these as the ''warts'' of the analysis. These ''warts'' can take the form of uncertainty due to sampling, interview protocol, measurement, coverage, or a host of other sources of error. Senior leaders understand that no source of data is perfect, so efforts to properly caveat and show transparency of limitations builds confidence in analytical products.
- Tell a cogent story. Do not simply summarize question responses. An unceasing barrage of summary statistics for each question of the survey instrument may seem like a thorough investigation of the data, but it is unlikely to inform decisions in a meaningful way. Filter down to the key findings. Be prepared to offer to return to any aspect of the survey to provide further information, but come to the presentation with a story to tell.
- Narratives are powerful. Quantitative findings that are anchored with qualitative anecdotes help to reinforce key points. Staff principles and decision makers may not remember a specific statistic but they may remember a well-crafted and salient narrative that summarizes the content of a statistic and puts a human face on the findings.
- Margins of error are highly influenced by the sample size. If the analysis filters down to a very small population, the sampling error increases. The margin or error may be low for the full sample, but once the analysis focuses on a very small aspect of sample, it may make the results less reliable. Analysts should guard against over-selling a finding if the sample size is too small; or at a minimum make sure the audience understands the risks.

6. Improving Survey Skill Development

Often analysts with little technical expertise in survey design and development are charged with gathering public perceptions or supervising a survey effort as a part of their duties. This dearth in training and experience can have significant consequences to decisions that may rely, in part, on those perceptions. We conclude with some educational recommendations for any analyst tasked with initiating or supervising surveys under various time restrictions.

6.1 Two hours to commit

If you have only 2 *hours* to commit to understanding a bit about Survey Research we recommend the following:

- View the video sequence:
  1. Sampling: Pew Research Sampling Design[a]
  2. Weighting: YouTube Weighting[b]
  3. Error Analysis:

YouTube Measurement Error[c]
YouTube Sampling Error[d]
YouTube Coverage Error[e]
YouTube Non-Response Error[f]

- Read: Chapter 5: Deployed Analyst Handbook

6.2 Two days to commit

If you have 2 *days* to commit to understanding ''Survey Research'' we recommend (in addition to items above) the following:

- Contact organizations that have experience in commissioning survey research for commands. One such organization is the Center for Army Analysis in Fort Belvoir, VA. For many years, this organization has served as a clearing house for connecting people to effectively improve operations. Their span of collaboration includes civilian and military contacts, international, intergovernmental, federally funded research and development centers, staff elements of the DoD, as well as key players in industry.
- Purchase and read a book on survey research. We recommend Fowler's work on survey research methods.[8] The text provides a well-organized and comprehensive view of survey methodology to get any analyst moving the right direction.

6.3 Two weeks to commit

If you have 2 *weeks* to commit to understanding ''Survey Research'' we recommend (in addition to items above) taking an open source course on survey design and implementation. Depending on the time available to commit to the coursework, we recommend a course similar to Corsera's Survey Data Collection and Analytics Specialization. Any coursework should include lessons on designing questionnaires, sampling, weighting, and dealing with missing data, as well as subsequent analysis of complex data. The goal of any course work should be to familiarize the user with major components of perception research.

7. Conclusion

The final important takeaway identified during the Working Group was: despite best efforts, no survey instrument will be perfect. Shortcomings will exist. However, if the collection of recommendations presented in this paper are considered, military commands will be in a position to field a survey that yields useful and actionable insights for senior leaders. Time spent eliciting clear objectives from leaders, and keeping leaders and other stakeholders informed of the survey process, will ensure the survey design meets the commanders intent. A survey that meets the commanders' intent has the greatest potential to deliver results that inform policy, operations, and the allocation of resources. The expense and risk of having boots on the ground in all places at all times to feed information to commanders is too high in most scenarios. Perceptions surveys offer an opportunity to cast a wide net to pull in information from the population.

As global conflicts continue to arise, and budget constraints remain, public perception surveys provide military commands with an effective instrument to inform the efficient allocation of constrained resources. A survey will never be perfect, but with technical knowledge, commitment, communication, and collaboration, the products that can be developed using the data collected can greatly inform and empower senior leaders in their strategic decision making.

Acknowledgments

The authors would like to thank the Military Operations Research Society for hosting the special session that facilitated the dialogue resulting in this paper. Additionally, we wish to thank military and civilian personnel serving in harm's way, protecting the very fabric that allows dissemination of thoughts, ideas, and efforts. Finally, we would like to acknowledge the scores of host-nation field workers often placed in significant risk to gather perception data that informs military operations.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

ORCID iD

Joseph Lindquist https://orcid.org/0000-0003-4172-9773

Notes

a. Web address: http://www.pewrfesearch.org/2017/05/12/video-explainer-understanding-random-sampling-for-public-opinion-surveys/
b. Web address: www.youtube.com/watch?v=KkqXbw43yxc
c. Web address: https://www.youtube.com/watch?v=zF37RvnNHnk
d. Web address: https://www.youtube.com/watch?v=XE7QDfdaQ68
e. Web address: https://www.youtube.com/watch?v=kaRQsW4nOcY
f. Web address: https://www.youtube.com/watch?v=3xY8QUklllo

References

1. Wenzel B and Nau K (2011) *Code of Best Practices for Survey Efforts*. White Sands Missile Range, NM.
2. Keenan B and Mauch M (1993) Developing and using ques tionnaires. *US Government Accountability Office* (PEMD-10.1.7). URL https://www.gao.gov/products/GAO/pemd-10.1.7.
3. American Association for Public Opinion Research (2018) Best practices for survey research. URL https://www.aapor.org/Standards-Ethics/Best-Practices.aspx.
4. LaRivee D (2011) Best practices guide for conducting assessments in counterinsurgencies. URL http://smallwarsjournal.com/jrnl/art/best-practices-guide-for-conducting-assessments-in-counterinsurgencies.
5. Kulzy W and Fricker R (2015) The survey process: With an emphasis on survey data analysis. *Phanlanx* 48(2): 32–37.
6. Bahm V, Collins G, Kollhoff R, Lindquist J, Stollenwerk M, Gaul M, Jesse E and Dunaway F, S T, Huddleston S, Hall H, Delara N, Haskell M, Shilling A and Anderson J (2016) *Deployed Analyst Handbook*. Fort Belvoir, VA. URL http://www.caa.army.mil/CAA-DAHB-30AUG2016_FINAL.pdf.
7. Lavallée P and Beaumont J (2015) Why we should put some weight on weights. survey insights: Methods from the field. *Survey Insights* URL https://surveyinsights.org/?p=6255.
8. Fowler F (2014) *Survey Research Methods*. Applied Social Research Methods. SAGE Publications. ISBN 9781412958417.

9. Groves R and Lyberg L (2010) Total survey error past, present and future. *Public Opinion Quaterly* URL https://doi.org/10.1093/poq/nfq065.
10. Kolenikov S (2016) Post-stratification or non-response adjustment? *Survey Practice* URL https://www.surveypractice.org/article/2809-post-stratification-or-non-response-adjustment.

Author biographies

Joseph Lindquist is an Associate Professor of Mathematics at the United States Military Academy where he leads the first year math program. He has 15 years of experience as an Operations Research Systems Analyst for the United States Army, providing analytical support while deployed to Operation Enduring Freedom (Afghanistan and the Southern Philippines) and leading dozens of analytical ''reach-back'' teams in support of ongoing operations. His research involves identifying and flexing OR techniques to support current operations and enjoys ''operationalizing'' mathematics for the students he teaches.

David Peng is the Chief Statistician and Vice President for Research at D3 Systems (Designs, Data, Decisions). In this role, Mr. Peng manages D3's Statistics Team and oversees sampling design and methodology, variance estimation, quality control, data manipulation, non-response analysis, and statistical data analysis. He has extensive experience in designing methodology for and analyzing data from difficult to reach environments across the world including conflict, post-conflict, and developing nations. Mr. Peng is an active member of the American Statistical Association and the American Association for Public Opinion Research (AAPOR).

Shawnette Rochelle retired from the U.S. Army as an Operations Research Systems Analyst in Fall 2018 at the rank of Lieutenant Colonel. In addition to her analytical assignments in the Army, she served as a fellow at Gallup through the US Army Training with Industry program. She is currently on sabbatical researching financial inclusion and the ways in which microfinance institutions, financial technology startups and impact investors contribute to financial inclusion.

Dusty Turner is an Operations Research Systems Analyst for the United States Army and an Assistant Professor at The United States Military Academy at West Point, NY. In this role, he leads a variety of courses including Advanced Introduction to Probability and Statistics and Sabermetrics. His research interests include sabermetrics, sentiment analysis, and pedagogy.

Matthew Warshaw is the Chief Operating Officer at D3 Systems (Designs, Data, Decisions). Mr. Warshaw has over 20 years of global social science research experience conducting public opinion surveys, media research, measures of performance and effectiveness studies, impact monitoring and evaluation projects, and consumer studies. Mr. Warshaw has extensive experience setting up research operations, logistics oversight, and foreign personnel management, with particular experience in conflict environments and developing countries. Mr. Warshaw serves as the managing director of D3's subsidiary, ACSOR-Surveys, in Kabul, Afghanistan and joined the board of Infinite Insight, D3's partner in Africa, in August 2011.