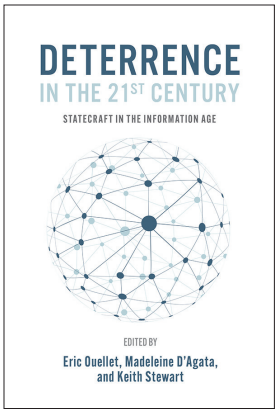
UNIVERSITY OF CALGARY
Press



DETERRENCE IN THE 21ST CENTURY: STATECRAFT IN THE INFORMATION AGE

Edited by Eric Ouellet, Madeleine D'Agata, and Keith Stewart

ISBN 978-1-77385-404-5

THIS BOOK IS AN OPEN ACCESS E-BOOK. It is an electronic version of a book that can be purchased in physical form through any bookseller or on-line retailer, or from our distributors. Please support this open access publication by requesting that your university purchase a print copy of this book, or by purchasing a copy yourself. If you have any questions, please contact us at ucpress@ucalgary.ca

Cover Art: The artwork on the cover of this book is not open access and falls under traditional copyright provisions; it cannot be reproduced in any way without written permission of the artists and their agents. The cover can be displayed as a complete cover image for the purposes of publicizing this work, but the artwork cannot be extracted from the context of the cover of this specific work without breaching the artist's copyright.

UNIVERSITY OF CALGARY
Press
press.ucalgary.ca

Acknowledgement: *We acknowledge the wording around open access used by Australian publisher,* **re.press**, *and thank them for giving us permission to adapt their wording to our policy* http://www.re-press.org

12

Assessing Influence in Target Audiences that Won't Say or Don't Know How Much They Have Been Influenced

Ronald D. Porter, Minqian Shen, Leandre R. Fabrigar, and Anthony Seaboyer

Introduction: The Challenge of Measuring Influence

Defending against and conducting influence operations has always been an important challenge facing the Canadian Armed Forces (CAF) and other national security organizations responsible for protecting Canadian citizens. For example, the CAF has long recognized the value of having a capability to influence the attitudes and behaviours of enemy forces in support of its military operations conducted abroad. Such a recognition by the CAF has resulted in the training of military personnel specifically tasked with conducting influence operations (i.e., psychological operations, or "PSYOPS," personnel).

However, with the increasing centrality of the Internet in every facet of citizens' lives and the prominence of social media platforms as a means of communication, the potential "battlefield" for social influence operations has expanded far beyond what might have been imagined by national security organizations even twenty-five years ago. In the online information environment of contemporary liberal democracies, both state and non-state adversaries are routinely targeting audiences with persuasive appeals designed to shape their attitudes and behaviours (e.g., see Kim et al., 2018)—though the degree of persuasiveness varies significantly depending on the adversary, the level of effort, and other mitigating factors (e.g., corruption in the

implementation of influence operations). For instance, the intelligence services of adversaries might be expected to conduct social influence operations in an effort to undermine support for a nation's leaders, policies, and institutions. Likewise, armed non-state groups and other radical organizations conduct influence campaigns in an effort to recruit new members or incite lone individuals to undertake violent or destructive actions. In response to such efforts, government organizations in some liberal democracies sometimes attempt to counteract the persuasive efforts of adversaries with their own influence operations.

At least since 2016, governments have understood the power non-kinetic influence campaigns can have compared to the more traditional measures of security organizations. The election of Donald Trump was, at the very least, supported by massive adversarial influence campaigns that were launched through micro-targeted, hyper-personalized influence campaigns (Lewis & Hilder, 2018). Many—but not all—subject-matter experts claim that the Russian influence campaign was effective enough to sway the election by 2–3 per cent, a margin that may well have been crucial to the outcome. Long before 2016, adversaries focused the larger part of their operations against the West in the non-kinetic environment. In a 2013 article, Russian general Valery Gerasimov famously described his perception that the way war is conducted has fundamentally changed and that non-kinetic means exceed kinetic means in a ratio of 4:1 (Gerasimov, 2016). China, other actors such as Iran and North Korea, as well as armed non-state actors, have certainly implemented similar strategies. Additionally, digitalization and the increasing use of social media are making influence operations more effective, easier, less risky for the actor, cheaper, and more efficient (Seaboyer, 2016, 2018; Singer, 2018). Finally, the need to understand which adversarial influence operations are actually effective derives from the fact that our information space in democratic societies is much easier to target than the information space of our adversaries—in which the Internet is heavily censored and (and at least somewhat) contained by firewalls and other measures to reduce foreign influence. Therefore, in order to defend our open democratic societies, it is essential to understand which adversarial influence campaigns are effective so that defence resources can be directed to where they are likely to be the most effective.

For these and other reasons, Western governments are increasingly seeing the importance of understanding which influence campaigns are effective,

and are therefore focussing efforts on increasing their abilities to measure the impact of influence campaigns.

Regardless of whether influence operations are being conducted on the traditional battlefield or in an online environment, key to evaluating the impact of an adversary's attempts at influence, as well as the efficacy of one's own efforts at influence, is the ability to measure attitudes in the target audiences of interest—as a first step to identifying the effectiveness of campaigns. More specifically, the impact of influence can only begin to be empirically evaluated when we are able to measure a target audience's attitudes both before and after exposure to that attempt. Alternatively, we must be able to measure attitudes in a subgroup of the target audience that has been exposed to an influence attempt, and then compare those attitudes to the attitudes of a comparable subgroup of the target audience that has not been exposed to the influence attempt. In either case, in the absence of an effective method for measuring attitudes, it is impossible to know which of an adversary's messages is proving especially effective, and thus to prioritize counteracting it. Likewise, it is difficult to know which of one's own influence operations are successful, and then accord them further resources.[1]

Unfortunately, the target audiences of adversaries' social influence operations are often not amenable to traditional methods of assessing public opinion (e.g., telephone or online surveys). For instance, members of radicalized audiences that are likely to be targets for recruitment by armed non-state actors might be expected to be unwilling to participate in a telephone survey on their views of political violence, and if they did participate, they might not be expected to give honest answers. Likewise, the target audiences of Canadian national security organizations' influence operations abroad are also unlikely to be audiences whose attitudes can be assessed using traditional approaches. For example, the soldiers of an adversary targeted by the CAF with PSYOPS leaflets urging surrender are unlikely to be in position to complete a survey indicating how seriously they are contemplating surrender. Thus, in many (perhaps most) cases in which national security organizations such as the CAF might wish to evaluate the efficacy of their own influence operations or those of their adversaries, the ability to measure the attitudes of target audiences is a major challenge.

The importance of measuring attitudes in such contexts, as well of the practical challenges of accomplishing this objective, have long been recognized by national security organizations such as the CAF. For example, in a

comprehensive review of military PSYOPS training manuals from the United States, United Kingdom, Canada, and NATO, Fabrigar and Porter (2008) noted that such materials routinely acknowledged the importance of assessing the impact of social influence attempts and the need to develop non-traditional measures for doing so. However, their review also noted the absence of concrete standardized procedures for constructing such measures in these training materials.

Chapter Overview and Objectives

The central goal of the present chapter is to discuss some of the challenges of assessing attitudes in the sort of environments and among the target audiences for which social influence must be evaluated by the CAF and other national security organizations. As it turns out, some of these challenges parallel those faced by social scientists in other contexts. In an effort to overcome these challenges, social scientists have developed a number of indirect measures of attitudes (e.g., see Gawronski & De Houwer, 2014; Kidder & Campbell, 1970; Petty et al., 2009; Webb et al., 1966). We begin by reviewing the reasons why social scientists have sometimes used indirect attitude measures before providing an overview of traditional indirect measures and more contemporary indirect measures of attitudes that have been proposed to overcome these problems. In discussing these traditional and contemporary approaches, we describe the procedural features of these measures, discuss their strengths and weaknesses, and evaluate their potential utility for use by the CAF and other national security organizations. In the next section, we propose potential adaptations to existing indirect measurement approaches that might enhance their utility for national security applications. We also discuss more novel procedural innovations that build on the principles of prior indirect measures that could potentially lead to other indirect measures with practical utility for national security contexts. In the final section, we present a set of key unresolved issues that must be addressed in order to develop an enhanced capability to assess the impact of social influence operations in national security settings.

Traditional and Contemporary Indirect Measures of Attitudes: The Origins of Indirect Attitude Measurement

Beginning in the 1920s, researchers in psychology and related disciplines began to develop formal procedures for assessing people's attitudes (e.g., see Guttman, 1944; Likert, 1932; Osgood et al., 1957; Thurstone, 1928). These various procedures all involved what have been traditionally called "direct measures" of attitudes and are now more commonly termed "explicit measures" of attitudes. Essentially, direct measures assess people's attitudes in overt ways by specifically prompting people to report their likes and dislikes (e.g., "Do you favour or oppose the death penalty for serious crimes?"). Such direct measures, when carefully constructed, have substantial utility, and they continue to be the most common form of attitude measures used in both research and application. However, even at a fairly early phase in the history of the research literature on attitude measurement, social scientists recognized that direct measures were not without their limitations (e.g., see Hammond, 1948; Proshansky, 1943). Concerns regarding direct measures arose from two potential problems.

First, because direct measures are so overt, the intent of what they are designed to assess is readily apparent. For many issues (e.g., "To what extent do you have a negative versus positive opinion of Crest toothpaste?," "To what extent do you dislike versus like spaghetti?"), this property of direct measures is unlikely to be a problem as people might be entirely comfortable reporting their attitudes. In other cases, issues might be more sensitive, but placing people in a sufficiently comfortable context (e.g., in a situation where their answers are anonymous) might be sufficient for people to respond accurately. However, in other cases, the issues might be so sensitive, or mistrust on the part of respondents might be so pronounced, that people are unlikely to respond honestly even when their responses are anonymous. In these cases, people might be expected to refuse to answer questions, or, if they do answer, to provide answers they believe the questioner wishes to hear rather than their true views (i.e., to engage socially desirable responding; see Paulhus, 1991).

A second potential limitation with direct measures that was recognized early on in the attitude measurement literature, and which has been even more prominently featured in contemporary discussions of attitude measurement, is that direct measures are to some degree based on the assumption that people can accurately access their own attitudes. That is, in order for a

person to directly report their attitudes, they must know what their attitude is. However, what if people have positive or negative reactions to something of which they are not consciously aware? Or alternatively, what if people have instant positive or negative "gut" reactions of which they are consciously aware, but whose accuracy they might doubt upon careful reflection? Despite their more considered doubts regarding these instant reactions, might these people's responses influence them when they are not actively monitoring these reactions? One might expect that direct measures of attitudes would do a poor job capturing such unconscious and/or spontaneously activated positive or negative reactions.

To overcome these potential problems, a number of "indirect measures" of attitudes (now more commonly termed "implicit measures" of attitudes) have been suggested.[2] Indirect measures involve a procedure for assessing attitudes that does not require overtly asking people their likes and dislikes. Rather, attitudes are inferred on the basis of some behavioural response or set of behavioural responses presumed to be related to the attitude of interest, or on the basis of how people perform some judgmental task presumed to be related to the target attitude of interest. Initial interest in indirect measures began in the 1940s and continued to grow through the 1950s and '60s (see Kidder & Campbell, 1970; Webb et al., 1966). While interest in indirect measures never entirely disappeared, it waned somewhat over the next thirty years, and then exploded in the early 2000s under the rubric of "implicit measures" (Porter, 2010). This interest has continued for the past twenty years, and the study of implicit measures remains a major topic of inquiry in contemporary social psychology and related disciplines (Gawronski & De Houwer, 2014; Petty et al., 2009).

Traditional Indirect Measures

Early attempts to indirectly measure attitudes were based on projective approaches (e.g., the thematic appreciation test; see Proshansky, 1943), but indirect measures soon evolved into more structured judgmental tasks (e.g., error choice; see Hammond, 1948) or behavioural observation procedures (e.g., lost letter; see Milgram et al., 1965).[3] Here, we discuss some of the better-known traditional indirect measures to illustrate the logic underlying these procedures and comment on their strengths and limitations.

BEHAVIOURAL OBSERVATIONS

One general approach to indirectly measuring attitudes is through the examination of a person's demonstrable behaviour. The underlying premise of this approach is that, if someone has a favourable or unfavourable attitude toward an attitude object, then it would presumably be reflected in their behaviour toward that attitude object. One of the best-known early examples of this approach is the lost letter technique (LLT) (Milgram et al., 1965). In this technique, a specific attitude object is identified (e.g., legalized abortion). A large number of pre-addressed and stamped envelopes are then randomly left in a variety of public locations. Half of the envelopes are addressed to an organization (fictitious but plausibly real) that someone could clearly identify as being positive toward the specific attitude object (e.g., "The Citizen Pro-Choice Coalition") and the other half addressed to an organization that could be clearly identified as negative toward the attitude object (e.g., "The Pro-Life Citizen Alliance"). The researcher then tracks how many letters are delivered to each addressee. The underlying assumption of this technique is that when a letter is found, people assume it has been accidentally dropped and are more likely to place it in a mailbox if it is addressed to an organization that is consistent with their own attitude, thereby providing a rough estimate of the popularity of each position. In this way, the people are not affected by social desirability because no one, other than themselves, are aware of their actions. Research has indicated that the LLT provides a reasonable overall estimate of the popularity of a given attitudinal position in a group of people (i.e., the group of people represented by the physical local in which the letters were initially distributed; Milgram et al., 1965), and can even function adequately in settings where people might fear for their physical safety were they to openly express their opinions (Kremer et al., 1986). However, one limitation of the approach is that although it can be used to infer the general distribution of two opposing views in a group of people, it does not provide individual-level information regarding the opinions of specific people (i.e., one has no way of deducing who specifically returned letters and thus what their opinions might be).

More recently, the LLT has been adapted to work in a more current technologically oriented environment focusing on emails rather than letters (Stern & Faber, 1997; Vaughan-Johnston et al., 2021). As a result, the name has been changed to the lost email technique (LET). The underlying premise of

this technique is similar to the LLT, except emails are sent "in error," with the rates of return assessed (Stern & Faber, 1997). In the LLT the participant has two options (mail or ignore the letter), whereas in the LET the recipient of the email can ignore or delete (interfere with the communication), send the message to the intended recipient, or return the email to the originator (letting them know that they made an error). The discrepancies in return rates in the LET can then be interpreted as either approval or disapproval of the contents (i.e., message) of the email (Bushman & Bonacci, 2004; Stern & Faber, 1997). In the case of the LET, one would often be able to infer the identity of individuals who received emails and whether they returned/forwarded the email or ignored/deleted them. Thus, one could infer individual-level attitudes, although such inferences would provide only a crude dichotomous assessment of attitudes (i.e., whether people are positive or negative in their evaluations, but not the extremity of those evaluations) and would likely reflect a substantial amount of error (e.g., some people might inadvertently miss the email or be very busy at the time the email arrives or regard the email as spam).

There are a number of other behavioural observation methods that have demonstrated validity in applied settings (Webb et al., 1966). Behavioural observation is the systematic recording of behaviour (usually surreptitiously) by an observer. The underlying premise of this approach is that, if someone has a favourable or unfavourable attitude toward an attitude object, then it would presumably be reflected in their behaviours toward that attitude object. Additionally, because evidence of people's attitudes is gathered from unobtrusive observation, attitudes can be assessed without affecting the behaviour of the people whose attitudes are being assessed (Webb et al., 1966). A number of general categories of behaviour have been suggested as reflective of attitudes. For example, Webb et al. (1966) noted that the physical distance people place between themselves in environments in which they can control their physical location can be used to infer interpersonal attitudes. Likewise, the tone of a person's voice when discussing a particular attitudinal position or when interacting with another person can be reflective of their attitudes toward that attitudinal position or that person. Obviously, any single behaviour will be determined by multiple factors and as such provides a very imperfect measure of attitudes. However, if a variety of behavioural responses can be aggregated, this aggregate score is likely to provide a more accurate assessment of attitudes.

Such observational behaviour approaches have often been advocated for use in military settings such as the assessment of PSYOPS activities (Goldstein & Findley, 1996). In theatre, for example, this could be the number of opposition soldiers that surrender following an information operation, or the number of posters torn down advocating a particular group or stand on a policy. However, discussions of behavioural observation measures have generally been highly specific and illustrative rather than leading to the development of standardized behavioural assessment procedures that might be applied broadly such as the LLT.

As with the LLT, in many cases it will not be possible to track the identity of specific people who have performed the target behaviours (e.g., the specific people who tore down posters). Thus, such observational measures will generally not permit the collection of individual-level information regarding people's attitudes as much as group-level information regarding the popularity of a particular position within a specified region or target group.

JUDGMENTAL BIAS APPROACHES

An early indirect approach to attitude measurement involved the use of a modified self-report measure called structured objective questionnaires. In this method, respondents are given what they believed to be an objective information test that assesses their knowledge on a particular subject; however, some of the questions are not objective and have no correct response. Rather, these questions have responses intentionally weighted for or against an attitude object and randomly dispersed within the information test (Coffin, 1941; Hammond, 1948; Kubany, 1953; Newcomb, 1940, 1946; Smith, 1947; Weschler, 1950a, 1950b). The underlying premise of this approach is that there is a relationship between a person's attitudes and how they interpret information presented as fact. That is, this method assumes that when people are presented with a question for which they do not know the correct response, their guessing reflects the respondents' attitudes (Coffin, 1941; Hammond, 1948; Newcomb, 1946).

Probably the best exemplar of this general approach is the error choice (EC) technique (Hammond, 1948). The EC technique involves presenting a set of objective knowledge questions that are in principle knowable but unlikely to be known and whose response options imply something either positive or negative about the attitude object. This procedure rests on two basic premises. First, when people are faced with a knowledge-based question for which they

do not know the answer, their guess will not be random; and one factor that they might rely upon in such guessing is their attitude. For example, when faced with a question where there are two factual possible answers, they will tend to pick the answer that best fits with their attitude. Thus, across a series of objective knowledge questions that are in principle knowable, but to which respondents are very unlikely to know the true answers, one might expect to find a systematic guess pattern that is consistent with people's attitudes. The second premise of the measure is that, because each of the items is presented as a factual question, people will not be aware that their attitude is being assessed. Early research suggested that the EC technique had promise, but its performance was never fully evaluated in subsequent research. More recent examinations of EC have provided further encouraging evidence (see Porter, 2010). Specifically, answers to EC questions do appear to reflect a single systematic response pattern that is comparatively reliable and at least in part represents the respondent's attitude. These studies also suggest that (as intended) this response pattern to the EC questions is highly resistant to socially desirable responding. Importantly, completion of measures allows for the collection of individual-level information about peoples' attitudes, just as completion of direct measures provides such information.

Contemporary Indirect Measures

Beginning in the late 1990s, interest in indirect measures of attitudes underwent a renaissance with the emergence of a new generation of indirect measures (e.g., Fazio et al., 1995; Greenwald et al., 1998), now more commonly referred to as implicit measures. These new methods built on methodological procedures used and phenomena documented in the research literatures within cognitive psychology and social cognition. These new implicit measures required the use of computers, which allowed for very precise timing in the presentation of stimuli and high-resolution recording of reaction times in responding to stimuli. Although different implicit measures vary in their specifics, all of these procedures involve presenting people with stimuli related to the topic of interest (i.e., the attitude object), usually in the form of words and/or images, and then asking people to perform some sort of judgmental task related to the stimuli. Some aspect of how these judgments are performed (e.g., the speed with which judgments are made) is assessed. This task performance criterion is, on the basis of some theoretical logic, presumed to be influenced by the attitude of interest. Importantly, these measures are all indirect in that

they never specifically ask people to report their attitudes. A number of such measures have been proposed (see Gawronski & De Houwer, 2014; Petty et al., 2009). For purposes of illustration, we will just briefly discuss three of the better-known of these contemporary indirect measures.

Implicit Association Task

The implicit association task (or IAT; see Greenwald et al., 1998) is a measure that, in its original form, assesses attitudes toward two competing persons, objects, or concepts. The technique has most famously been used to assess prejudice toward social groups (e.g., racial groups), but can be adapted to assess attitudes toward virtually anything. Participants complete rapid judgment tasks in which they are instructed to sort words (or images) into one of two categories as quickly as possible using one of two designated computer keys to indicate the group to which the word belongs.

For example, an IAT designed to measure attitudes toward Canada versus the United States would first present respondents with words either associated with Canada (maple leaf, Ottawa) or America (Washington, DC, bald eagle). Respondents indicate for each word presented whether it is a word related to Canada or America by pressing one of the two designated keys. They are then presented with a new list of words (e.g., death, love, vomit, peace) with a second categorization task of indicating whether the words are positive or negative, once again using the two designated response keys.

In the critical later phases, these two categorization tasks are combined so that words are randomly presented from either list (Canada/America and positive/negative), but only two response keys are used, which mean the keys must be shared for both categorization tasks. For example, if the classifying categories are Canada/America and positive/negative, one of the two keys might be designated for words that are related to Canada or positive, and the other key for words that are related to America or bad. In a later phase, this sorting task is repeated for the reverse combination of shared keys (i.e., if the first round used Canada/positive and America/negative, the next phase would use America/positive and Canada/negative as the shared response keys). The time it takes for participants to sort each word after presentation is recorded.

The theory behind the IAT is that strong congruent associations between concepts should lead to fast responses when they share a response key, and that strong incongruent associations between concepts should lead to slow responses when they share a response key. In other words, if people have very

positive attitudes toward Canada, they should be relatively fast at performing the task when Canada/positive share the same response key compared to when Canada/negative share the same response key. Likewise, very positive attitudes toward America should produce a response pattern in which people are much faster when America/positive share the same response key than when America/bad share the same response key. Thus, the difference in time it takes for people to perform the task when Canada/positive and America/negative share keys compared to when America/positive and Canada/negative share response keys provides a measure of whether people's attitudes toward Canada are more positive versus negative than their attitudes toward America. Revised versions of the IAT have been developed that can be used to assess attitudes toward a single group, concept, or person (Karpinski & Steinman, 2006).

The strengths of the IAT mainly revolve around its implicit nature; by assessing implicit evaluations through quick reaction time–based tasks, respondents do not have time to consider whether their responses are socially appropriate. Similarly, word sorting does not have very intuitive connections to attitude assessment, and thus respondents will be less likely to ascertain the intent of the measure, further shielding them from socially desirable modified responses. Another strength of the IAT is its versatility; it can be formatted to measure associations between any classification/concept (e.g., black/white, fat/thin, America/Iraq) and virtually any attribute (good/bad, strong/weak). Thus, the core procedure of the IAT can be adapted to study a wide range of judgments.

Of course, the practical weaknesses of the IAT include its resource demand and the vulnerability of its accuracy to outside interference. The IAT is a computer task that requires limited distractions for an extended period of time (often fifteen to twenty minutes) in order to gather high-resolution data based on reaction times. Thus, participants in uncontrolled settings might be unwilling or unable to complete the IAT appropriately, although reasonably good data can be collected in online settings if respondents are sufficiently motivated and have a location where they can perform the task that is not too distracting (e.g., Xu et al., 2014).

Evaluative Priming

Evaluative priming (or EP, also sometimes referred to as affective priming; see Fazio et al., 1995) involves presenting target words (or images) representing

the topic of interest for which one wants to measure attitudes along with words (or images) representing positive or negative evaluation. The words representing the topic of interest serve as the "primes," and the words representing positive or negative evaluation serve as the targets of judgment. In this task, respondents are told that they will first be presented with an orientation word to help focus their gaze on the appropriate location on the computer screen (the prime) and that this word will appear only briefly, rapidly followed by the target word. They must then judge as quickly as possible if the target word is either positive or negative. For each judgment, the speed with which the target word is judged is recorded by the computer.

For example, if EP was being used to measure attitudes toward Canada, the prime words used for each trial would be words strongly related to Canada (e.g., maple leaf, Ottawa). The target words would be words almost universally seen as positive or negative (e.g., love, vomit). The EP procedure is based on a well-documented phenomenon that when evaluative responses are evoked, they will tend to facilitate the ease with which people can make judgments about things congruent with that evaluation and will interfere with judgments about things incongruent with the evaluation. Thus, if people have very positive attitudes toward Canada, the Canada-related prime words should evoke positive evaluative responses in people, which will in turn make them very fast at categorizing positive target words (e.g., love) and very slow at categorizing negative target words (e.g., vomit). People with negative attitudes toward Canada should have negative evaluative responses evoked by the Canada-related prime words, thus showing a reverse pattern (i.e., fast at judging negative words and slow at judging positive words). The difference in the average speed of judging positive target words versus negative target words that are preceded by Canada-related prime words provides the measure of people's attitudes.

Evaluative priming shares some of the same practical strengths and weaknesses of the IAT. The task itself largely bypasses any effortful modification of responses due to the primed words being presented very briefly and the need to categorize target words very quickly. Importantly, people are never asked to make any judgments of the word primes themselves (which are the words actually related to the topic of interest), and thus the intent of the task is not readily apparent. However, like the IAT, it requires a reasonably large number of judgment trials to be valid, and thus requires some extended time and effort on the part of respondents. Likewise, the high-resolution

concerning reaction times required for the measure are vulnerable to outside distractions.

AFFECT-MISATTRIBUTION PROCEDURE

Similar to evaluative priming, the affect-misattribution procedure (or AMP; see Payne et al., 2005) uses words (or images) related to the topic of interest as "primes" in a judgmental task. However, the specific targets of judgment in the task are somewhat different in that they are stimuli that would not be expected to evoke a negative or positive evaluation (e.g., an abstract shape, symbol, or ideograph). Participants are then asked to judge target neutral stimuli as either positive or negative.

For instance, continuing with our attitudes toward Canada example, the primes could once again be words related to Canada (e.g., maple leaf, Ottawa). The neutral stimuli could be letters from an ancient language unknown to the respondents. For each trial, the prime word (e.g., maple leaf, Ottawa) would very briefly appear, rapidly followed by a letter from the ancient language, which itself is presented only briefly. Respondents are then queried to judge if they feel more positive or negative toward the letter that was just presented.

The logic behind the AMP is simple; the primed word will trigger an evaluative response within the respondent, which, because of the very brief presentations of both the prime and the target of judgment, will subsequently be misattributed to the neutral stimulus. Hence, the task works via affect misattribution, as the evaluation of the ambiguous stimulus is directly influenced by an individual's evaluation of the primes representing the topic of interest (e.g., Canada). Thus, in the case of our example, positive attitudes toward Canada would be expected to produce a response pattern in which people tend to report being positive toward most of the letters that are preceded by Canada-related words. In contrast, negative attitudes toward Canada would be expected to produce a response pattern in which most of the letters preceded by Canada-related words would be judged negative.

Like the IAT and EP, the AMP is opaque in its intent in that people are never asked to judge the primes (i.e., the stimuli directly related to the topic of interest). Additionally, because of the very rapid presentation of stimuli, it is very difficult for people to exert intentional control over their responses. Indeed, instructing respondents to not allow the primes to have any effect on their judgments of targets has little actual impact on their judgments of the neutral stimuli (e.g., ancient letters). Because the procedure involves very

precise timing in presenting stimuli, it requires computers in order to be administered. However, the procedure makes no use of the reaction time of the respondents, but instead simply the proportion of positive versus negative responses to the letters or other neutral stimuli. Thus, it is likely less sensitive to distractions. Additionally, the measure can be used with comparatively few trials and thus can be completed in just a few minutes. Hence, the simpler nature of the AMP makes it a potential candidate for wider adoption in a variety of circumstances.

Concluding Thoughts on Existing Indirect Measures of Attitudes

As illustrated in our review, the use of indirect measures has a long history in social psychology and related disciplines. In some respects, the reasons for developing these measures arose in response to challenges that parallel the sort of issues faced by the CAF and other national security organizations when they attempt to gauge the efficacy of their own influence operations or those of their adversaries (e.g., concerns that target audiences might be unwilling to honestly report their attitudes). Specifically, these existing measures were designed to assess attitudes in audiences and/or contexts where people might be unwilling or unable to respond to overt attitude measures.

That being said, there are important practical differences in how these existing measures have been applied in social science research and the likely contexts and audiences for which they would need to be used in national security settings. In many situations, the contexts and audiences in national security settings present far more challenging practical constraints, and thus one cannot assume that respondents will have either the ability or the motivation to undertake lengthy measurement procedures, even when they are unaware of the intent of these procedures. For instance, soldiers of an adversary are unlikely to have the opportunity or inclination to complete a twenty-minute IAT procedure assessing their attitudes toward surrender. Thus, comparatively few of these indirect measures are likely to be suitable in their current form for use in national security settings. That being said, many of the core concepts and procedures underlying these existing indirect measures could provide a foundation for developing indirect measures that might be suitable for these more demanding contexts and audiences (e.g., enemy soldiers on a battlefield). It is this possibility to which we turn our attention in the next section of this chapter.

Developing Indirect Measures for National Security Settings

Conceptually, indirect attitude measures share a number of features. Most notably, people are never directly asked to report their attitude, making it difficult to deduce what exactly these techniques are measuring. In addition, many of these procedures are designed to assess attitudes without giving people a chance to intentionally adjust or consider their responses, which can be important for gathering information in areas where expressing one's true attitude may carry negative consequences and/or when one wants to assess people's instant "gut" reactions. That said, the contemporary techniques that are particularly salient here are computer-based assessments that require attention and time, which can often not be guaranteed in field settings. In some cases, it might be possible to overcome these practical challenges simply by presenting these tasks in creative ways that might be likely to engage people to expend the effort to complete the procedures. Thus, with some minor adaptations, existing measures could be rendered suitable in some circumstances. In other cases, more fundamental changes might be necessary that ultimately involve creating new indirect measures. However, even in these cases, the existing measures might provide a conceptual and/or procedural starting point upon which to base these new measures.

Potential Adaptations of Existing Indirect Measures

Even if the exact procedures for the techniques previously discussed cannot be precisely replicated for use in some field settings, the core procedural features could be utilized in many settings where online-based administration of measures is feasible. As we have noted, with the explosion of social media platforms for communication, much of the social influence conducted by adversaries and the government organizations tasked with countering them is likely to occur in online settings or via other forms of digital communication. Many of the methods previously discussed could be administered in these settings, and indeed social scientists have been collecting data using indirect measures in online settings for many years (e.g., the Project Implicit Website at https://implicit.harvard.edu/implicit/; see also Xu et al., 2014). The primary challenge is finding ways to "frame" the purpose of these tasks such that they are at best likely to encourage people to devote time to completing them, and at worst do not cause the target audience to actively avoid responding to these measures. That is, these measures must be opaque not only in terms of what

they are measuring, but also who is sponsoring them and the purpose for which the information is being used.

In considering existing measures for adaptation, perhaps the easiest might be the error-choice technique (EC) and the affect-misattribution procedure (AMP), because neither measure requires high-resolution response-time data and both are comparatively short in duration. These procedures could be administered in online environments and likely completed even in contexts where people have some outside auditory distractions. However, plausible cover stories would need to be provided for the purposes of such measures. For example, the AMP could be presented under the guise of a game that informs the respondent of a certain skill based on their evaluation of neutral stimulus. Judging unknown letters might be framed as a measure of people's ability to learn or intuit new languages or symbol systems. In the case of the EC, it could be framed as a test of people's knowledge of certain topics or general trivia knowledge. Importantly, just as such Internet games often include prizes for performance, similar prizes could be offered to induce people to undertake these tasks. Such games could be advertised on social media, where they would be exposed to many people within a specified geographical area, interest group, or other designation to allow for widespread but precise data collection.

Similarly, if the target group of interest is likely to be accessing measures in contexts where distractions are comparatively modest and they might have time to complete lengthier measures (e.g., a home or a workplace setting), reaction time–based measures such as EP and IAT could be feasible. These tasks could be advertised as "reaction time" or "brain age" tests for participants to assess their cognitive speed. Once again, incentives could be offered and advertisements on social media outlets could be targeted at designated groups.

Potential New Indirect Attitude Measures

In other cases, it might not be feasible to adapt existing measures, or it might be useful to develop new measures to supplement existing ones. In these cases, following the general logic of traditional indirect measures based on behavioural observations could be an avenue for developing new measures. However, the opportunities for collecting behavioural data are far richer now than was the case in the 1950s and '60s, when these approaches were originally developed. The vast majority of countries now either have widespread Internet access or are approaching that point; using this medium to gather

behavioural information could be invaluable due to the unprecedented reach it enjoys among the potential audience.

Researchers could construct websites focusing on a central topic of interest and advertise them via social media. A target audience's engagement with the content of these websites could be measured by counting the number of visits to a site, average time spent on a website, and registered email subscriptions. Additionally, activity can be monitored for various pages of the website covering different types of content to compare which content is engaged with more and can therefore be interpreted as reflective of attitudes. In addition to advertising the website through social media, flyers with QR codes could be posted or distributed to a target audience.

In addition to websites, social media could be directly engaged to assess user attitudes. Many social media outlets have built-in measures of community engagement (e.g., Facebook "likes," Twitter "likes" and "retweets," YouTube views and subscriptions, and Reddit "upvotes") where the degree of community engagement and valenced evaluations of content can be directly ascertained. For example, Facebook is one of the most used social media outlets worldwide and has many different methods with which users can engage with people. Creating and advertising a "Facebook page" that represents a certain belief or idea would allow a researcher to assess a target audience's engagement with said beliefs by measuring the number of people who follow that page and "like" its posts. Similarly, comments on said posts can be coded for valence and intensity to assess attitudes toward them (see Rockledge et al., 2018). Overall, Facebook has the potential to be a versatile and far-reaching tool for data and information collection.

Twitter is another highly popular social media outlet that measures an online community's engagement with short messages or images via "likes," which indicate approval of a message, and "retweets," where a user reposts another user's message to their own social network. Both of these responses can be gauged to assess the degree of exposure and agreement with the associated public posts. Like Facebook, Twitter users can reply to posts while simultaneously spreading them to their own social network. Thus, engagement allows the message to be more visible to more people, creating a snowball effect for data collection.

In addition to the previous outlets, researchers can use YouTube to upload videos containing certain messages or arguments and track engagement through view count, subscriptions to the channel that posts the video

(indicating that the user wishes to see more content of the same nature), and monitoring the like/dislike ratio and comments on the video itself. Videos also allow for richer stimuli to be tested on social media users for assessing attitudinal responses.

In summary, the Internet offers a vast array of options for presenting members of a target audience with opportunities for engaging in behaviours related to a given topic of interest that might be used to reliably infer those people's attitudes. However, recent developments in data analytics might permit this method to achieve even higher levels of accuracy than was possible with earlier behavioural observation techniques. More specifically, an emerging literature in the social sciences has focused on developing computational algorithms that can be used to infer specific attributes of people from their "digital footprints" (i.e., their online activities). Thus, large of arrays of online behavioural responses can be combined using formal computational algorithms optimized for accuracy of prediction.

For example, inferences regarding personality traits on the basis of social media content can be made using computer-based algorithms that outperform the judgments of laypeople examining the same social media content (see meta-analytic summaries by Azucar et al., 2018, and Hinds & Joison, 2019; see also Park et al., 2015). However, inferences are not confined to personality traits. Research suggests that prediction algorithms can be used to infer a variety of other characteristics such as sexual orientation, ethnicity, religious and political views, intelligence, happiness, age, and gender (e.g., Kern et al., 2016; Kosinski et al., 2013; Settanni et al., 2018). It is also possible to infer more specific features of people's attitudes such as their emotionality and extremity (e.g., Rockledge et al., 2018). Thus, it might be possible to construct websites and/or create social media content to elicit behavioural responses in a target audience and then develop specific computational algorithms to optimize the value of this information for inferring attitudes on the topic of interest.

Of course, not all situations in which social influence is assessed will be amenable to Internet-based data collection. For example, the CAF will still find itself confronting situations in which the efficacy of its influence operations or those of its adversaries must be assessed in places such as a physical battlefield. In these contexts, adapting traditional behavioural observation measures might still be possible. Following the general logic of procedures such as the LLT, it might be possible to develop tangible physical

communications (e.g., leaflets, posters) or other actions that imply a certain attitudinal position and then create contexts where people have the possibility to engage in behavioural response that either facilitate or inhibit these efforts. One might then infer the prevalence of attitudinal positions at an aggregate level, or, if precise behavioural data can be collected on individuals, perhaps even at an individual level. Importantly, one could in principle develop computational algorithms that combine responses to a variety of these focal behavioural actions so as to enhance the accuracy of inferring attitudes from such behaviours, just as they are used to more accurately infer attributes on the basis of online behaviours. Developing "standardized behavioural opportunity" protocols that mimic essential features of techniques such as the LLT and that can be applied with only modest modification across a range of situations constitutes one of the great challenges and potential opportunities for enhancing the ability to evaluate social influence operations. Equally important and promising is the effort to developing more sophisticated and efficient data analytic procedures for inferring information from this behavioural observation data.

Concluding Thoughts

While many promising methods of indirect attitude assessment have been developed over the years, the research focusing on the application of these methods to field settings, particularly of the sort often faced by militaries and other national security organizations, has been relatively sparse. Indeed, some of these techniques are dependent on controlled environments to minimize distractions and involve relatively lengthy procedures that can become tedious. Given these facts, many of the current indirect measures of attitudes are likely to be more suitable to relatively controlled environments, and particularly to audiences that are at least reasonably motivated to be co-operative.

That being said, these challenges are by no means insurmountable, and this research literature has the potential to provide valuable contributions to the efforts of government security organizations seeking to better assess the impact of their own social influence operations and those of their adversaries. A few of these procedures might, with only modest adaptations, be employed in some relevant field settings. Likewise, established indirect attitude assessment techniques employ general principles that can be retained and transferred to new mediums and designs that could be suitable for an even wider range of naturalistic environments. Of course, the potential adaptations and

innovations we have discussed are at this point speculative. Future research would need to be conducted to fully develop the procedural details of these adaptations and new approaches and to evaluate their validity. Thus, if a robust capability in assessing the impact of social influence operations is to be developed by the CAF and other Canadian government organizations tasked with conducting and countering such activities, a sustained commitment to empirically investigating indirect measures will need to be undertaken. Such challenges are unlikely to be addressed by the academic community on its own.

Equally important, the CAF and other relevant organizations will also need to make a sustained commitment to carefully consider the doctrinal issues that arise from utilizing such measures. For example, our speculations regarding the alternative ways in which existing measures such as the EC and AMP might be presented involve the active deception of respondents. Indirect measures necessarily involve some level of deception and/or ambiguity, the cost of which will have to be weighed against the potential benefits of obtaining such information.

Additionally, there are important operational considerations that must be addressed. If such techniques are employed, there is the distinct possibility that adversary governments and organizations will condemn such techniques of information gathering. As such, they may intervene to stop or corrupt data collection. This can be done directly by having websites taken down, engaging in cyber-attacks, or feeding fake/useless information through the data-collection streams. Indirect methods of shutting down such research can also be employed (e.g., disabling Internet access in areas of interest, or warning people to be suspicious of new surveys and pages on their social media pages). These practical challenges will need to considered and tactics for coping with them developed accordingly.

In summary, because these techniques are novel to the CAF and other security organizations, many implications, as well as the potential challenges of indirectly assessing attitudes in national security applications, remain unknown. More empirical research and doctrinal development are required if the potential of these techniques is to be fully realized and the related risks fully appreciated. Ultimately, it is important to grapple with not only the questions of if and how these measures can be used, but also whether they should be used at all, and if so, for whom and under what circumstances.

Authors' Note

Preparation of this chapter was supported in part by an Insight Development Grant (430-2019-00099) from the Social Science and Humanities Research Council of Canada (SSHRC) of Canada to the first and third authors and Insight Grants (435-2015-0114, 435-2022-0034) from SSHRC to the third author.

NOTES

1 It should be noted that in most influence operations in national security settings (as well as other applied settings), the ultimate goal of operations is some form of behavioural outcome, be it a very specific target behaviour or a broad pattern of behavioural responses across an array of relevant behaviours. Generally, attitude change is a necessary but not sufficient condition for achieving broad and enduring behavioural change. Thus, the assessment of attitude change can provide a preliminary evaluation of the likelihood of success of an influence operation, but not a definitive verdict on its ultimate efficacy. The topic of when and why attitudes predict behaviour and how to assess the likelihood that attitude change might be expected to translate into changes in behaviour is itself the subject of a large research literature that goes beyond the scope of this chapter (see Fabrigar et al., 2019; Fabrigar et al., 2010).

2 Throughout this chapter, we primarily use the terms "direct/indirect attitude measures" to differentiate between traditional attitude measures that overtly ask people to report their attitudes and more subtle forms of attitude measurement that never overtly ask people to report their attitudes. This terminology has been the traditional set of labels for differentiating between overt and subtle attitude measures, but it is less commonly used in contemporary discussions of attitude measurement. Instead, the terms "explicit/implicit" have become more popular. In many discussions, the manner in which these two sets of terms have been used can be considered interchangeable. However, in some contemporary discussions (e.g., Gawronski & De Houwer, 2014), the term "implicit measure" has been used in a somewhat more restrictive manner to refer to indirect measures that are presumed to reflect comparatively automatic psychological processes that operate outside people's intentional control. For this reason, we use the "direct/indirect" terms, which refer to the overtness of the measure and convey no formal assumptions regarding the nature of the psychological process it reflects.

3 Another alternative approach to direct measures of attitudes is the use of physiological responses. A number of physiological measures of attitudes have been proposed (Blascovich, 2014), some of which have been found to function reasonably well. Because such measures are unlikely to be feasible in the field settings in which one might expect to use attitude measures for the purposes we discuss, we do not analyze these measures in this chapter.

REFERENCES

Azucar, D., Maregndo, D., & Settanni, M. (2018). Predicting Big 5 personality traits from digital footprints on social media: A meta-analysis. *Personality and Individual Differences, 124*, 150–9. https://doi.org/10.1016/j.paid.2017.12.018

Blascovich, J. (2014). Using physiological indexes in social psychological research. In H. T. Reis & C. M. Judd (Eds.), *Handbook of research methods in social and personality psychology* (2nd ed., pp. 101–22). Cambridge University Press.

Bushman, B. J., & Bonacci, A. M. (2004). You've got mail: Using e-mail to examine the effect of prejudiced attitudes on discrimination against Arabs. *Journal of Experimental Social Psychology, 40*(6), 753–9. https://doi.org/10.1016/j.jesp.2004.02.001

Coffin, T. E. (1941). Some conditions of suggestion and suggestibility: A study of certain attitudinal and situational factors influencing the process of suggestion. *Psychological Monographs, 53*(4), 125–7. https://doi.org/10.1037/h0093490

Fabrigar, L. R., MacDonald, T. K., & Wegener, D. T. (2019). The origins and structure of attitudes. In D. Albarracin & B. T. Johnson (Eds.), *Handbook of attitudes and attitude change*, vol. 1 (2nd ed., pp. 109–57). Routledge.

Fabrigar, L. R. & Porter, R. D. (2008). *Evaluating the psychology in psychological operations: An assessment of the state of psychological knowledge in Canadian Forces' PSYOPS training* (Research Report 2008-01). Royal Military College of Canada.

Fabrigar, L. R., Wegener, D. T., & MacDonald, T. K. (2010). Distinguishing between prediction and influence: Multiple processes underlying attitude-behavior consistency. In C. R. Agnew, D. E. Carlston, W. G. Graziano, & J. R. Kelly (Eds.), *Then a miracle occurs: Focusing on behavior in social psychological theory and research* (pp. 162–85). Oxford University Press.

Fazio, R. H., Jackson, J. R., Dunton, B. C., & Williams, C. J. (1995). Variability in automatic activation as an unobtrusive measure of racial attitudes: A bone fide pipeline? *Journal of Personality and Social Psychology, 69*(6),1013–27. https://doi.org/10.1037/0022-3514.69.6.1013

Gawronski, B. & De Houwer, J. (2014). Implicit measures in social and personality psychology. In H. T. Reis & C. M. Judd (Eds.), *Handbook of research methods in social and personality psychology* (2nd ed., pp. 283–310). Cambridge University Press.

Gerasimov, V. (2016). The value of science is in the foresight: New challenges demand rethinking the forms and methods of carrying out combat operations. *Military Review*, January–February, pp. 23–9.

Goldstein, F. L., & Findley, B. F. (1996). *Psychological operations: Principles and case studies*. Air University Press.

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of*

Personality and Social Psychology, 74(6), 1464–80. https://doi.org/10.1037/0022-3514.74.6.1464

Guttman, L. (1944). A basis for scaling qualitative data. *American Sociological Review, 9*(2), 139–50. https://doi.org/10.2307/2086306

Hammond, K. R. (1948). Measuring attitudes by error-choice: An indirect method. *Journal of Abnormal and Social Psychology, 43*(1), 38–48. https://doi.org/10.1037/h0059576

Hinds, J., & Joinson, A. (2019). Human and computer personality prediction from digital footprints. *Current Directions in Psychological Science, 28*(2), 204–11. https://doi.org/10.1177/0963721419827849

Karpinski, A. & Steinman, R. B. (2006). A single category implicit association test as a measure of implicit cognition. *Journal of Personality and Social Psychology, 91*(1), 16–32. https://doi.org/10.1037/0022-3514.91.1.16

Kern, M. L., Park, G., Eichstaedt, J. C., Schwartz, H. A., Sap, M., Smith, L. K., & Ungar, L. H. (2016). Gaining insights from social media language: Methodologies and challenges. *Psychological Methods, 21*(4), 507–25. https://doi.org/10.1037/met0000091

Kidder, L. H., & Campbell, D. T. (1970). The indirect testing of social attitudes. In G. F. Summers (Ed.), *Attitude measurement* (pp. 333–85). Rand McNally & Co.

Kim, Y. M., Hsu J., Neiman D., et al. (2018). The stealth media? Groups and targets behind divisive issue campaigns on Facebook. *Political Communication, 35*(4), 515–41. https://doi.org/10.1080/10584609.2018.1476425

Kosinski, M., Stillwell, D., & Graepel, T. (2013). Private traits and attributes are predictable from digital records of human behavior. *Proceedings of the National Academy of Sciences, 110*(15), 5802–5. https://doi.org/10.1073/pnas.1218772110

Kremer, J., Barry, R., & McNally, A. (1986). The misdirected letter and the quasi-questionnaire: Unobtrusive measures of prejudice in Northern Ireland. *Journal of Applied Social Psychology, 16*(4), 303–9. https://doi.org/10.1111/j.1559-1816.1986.tb01142.x

Kubany, A. J. (1953). A validation study of the error-choice technique using attitudes on national health insurance. *Educational and Psychological Measurement, 13*(2), 157–63. https://journals.sagepub.com/doi/pdf/10.1177/001316445301300201?casa_token=CMyHev5uD3MAAAAA:A9fppFAINUwxlaVfTVe9KrY2WCx5Z70mkLoZtpX9NxOoxrWtVeiuIcDmg1U5EfgdUezVtXkOmGp-Lw

Lewis, P., Hilder, P. (2018, 23 March). Leaked: Cambridge Analytica's blueprint for Trump victory. *The Guardian*. https://www.theguardian.com/uk-news/2018/mar/23/leaked-cambridge-analyticas-blueprint-for-trump-victory

Likert, R. (1932). A technique for the measurement of attitudes. *Archives of Psychology, 22*(140), 5–55.

Milgram, S., Mann, L., & Harter., S. (1965). The lost-letter technique: A tool of social research. *Public Opinion Quarterly, 29*, 436–7. http://www.communicationcache.

com/uploads/1/0/8/8/10887248/the_lost-letter_technique-_a_tool_of_social_research.pdf

Newcomb, T. M. (1940). Labor unions as seen by their members: An attempt to measure attitudes. In G. W. Hartmann, & T. M. Newcomb (Eds.), *Industrial conflict* (pp. 313–38). Cordon.

Newcomb, T. M. (1946). The influence of attitude climate upon some determinants of information. *Journal of Abnormal and Social Psychology, 41*(3), 291–302. https://doi.org/10.1037/h0058824

Osgood, C. E., Suci, G. J., & Tannenbaum, P. H. (1957). *The measurement of meaning.* University of Illinois Press.

Park, G., Schwartz, H. A., Eichstaedt, J. C., Kern, M. L., Kosinski, M., Stillwell, D. J., Ungar, L. H., & Seligman, M. E. P. (2015). Automatic personality assessment through social media language. *Journal of Personality and Social Psychology, 108*, 934–52.

Payne, B. K., Cheng, C. M., Govorun, O., & Stewart, B. D. (2005). An inkblot for attitudes: Affect misattribution as implicit measurement. *Journal of Personality and Social Psychology, 89*, 277–93. https://doi.org/10.1037/0022-3514.89.3.277

Petty, R. E., Fazio, R. H., & Briñol, P. (Eds.). (2009). *Attitudes: Insights from the new implicit measures*. Psychology Press.

Paulhus, D. L. (1991). Measurement and control of response bias. In Robinson, J. P., Shaver, P. R., Wrightsman, L. S. (Ed.), *Measures of personality and social psychological attitudes*, vol. 1 (pp. 17–51). Academic Press.

Porter, R. D. (2010). *Resurrecting the error choice technique: The premature demise of an indirect measure of attitude?* [Unpublished PhD dissertation]. Queen's University.

Proshansky, H., & Murphy, G. (1942). The effects of reward and punishment on perception. *Journal of Psychology, 13*(2), 295–305. https://doi.org/10.1080/00223980.1942.9917097

Rockledge, M. D., Rucker, D. D., & Nordgren, L. F. (2018). The evaluative lexicon 2.0: The measurement of emotionality, extremity, and valence in language. *Behavioral Research Methods, 50*, 1327–44. https://link.springer.com/article/10.3758/s13428-017-0975-6

Seaboyer, A. (2016). *Social media messaging for influence in national security.* Defence Research & Development Canada. Report number DRDC-RDDC-2016-C257. https://www.researchgate.net/publication/331071944_Social_Media_Messaging_for_Influence_in_National_Security.

Seaboyer, A. (2018). *Influencing techniques using social media*. Defence Research & Development Canada. Report number DRDC-RDDC-2018-C177. https://www.researchgate.net/publication/331069575_Influence_Techniques_Using_Social_Media.

Settanni, M., Azucar, D., & Marengo, D. (2018). Predicting individual characteristics from digital traces on social media: A meta-analysis. *Cyberpsychology, Behavior, and Social Networking, 21*(4), 217–28. https://doi.org/10.1089/cyber.2017.0384

Singer, P. W., Brooking, E. T., (2018). *LikeWar: The weaponization of social media*. Eamon Dolan/Houghton Mifflin Harcourt.

Smith, G. H. (1947). Beliefs in statements labelled Fact and Rumor. *Journal of Abnormal & Social Psychology, 42*(1), 80–90. https://doi.org/10.1037/h0057845

Stern, S. E., & Faber, J. E. (1997). The lost e-mail method: Milgram's lost-letter technique in the age of the Internet. *Behavior Research Methods, Instruments, & Computers, 29*(2), 260–3. https://link.springer.com/article/10.3758/BF03204823

Thurstone, L. L. (1928). Attitudes can be measured. *American Journal of Sociology, 33*(4), 529–54.

Vaughan-Johnston, T. I., Fowlie, D. I., & Jacobson, J. A. (2021). *Facilitating scientific communication between strangers: A preregistered lost email experiment* [Unpublished manuscript]. Queen's University.

Webb, E. J., Campbell, D. T., Schwartz, R. D., & Sechrest, L. (1966). *Unobtrusive measures: Nonreactive research in the social sciences*. Rand McNally.

Weschler, I. R. (1950a). An investigation of attitudes toward labour and management by means of the error-choice method. *Journal of Social Psychology, 32*(1), 51–62. https://www.tandfonline.com/doi/pdf/10.1080/00224545.1950.9919030?casa_token=vE9GWsxIqOYAAAAA:gGl7XiEr1wFKNoL3Lil7F2Ef-0E0e-1-TqsCUZ_0Ry-UDVlohuLQe5a8uNUImNnDn6nrspCm--OMNQ

Weschler, I. R. (1950b). A follow-up study on the measurement of attitudes toward labour and management by means of the error-choice method. *Journal of Social Psychology, 32*(1), 63–9. https://www.tandfonline.com/doi/pdf/10.1080/00224545.1950.9919031?casa_token=IWdQMntASEgAAAAA:wFLCZ1Jze77ddK7CjT0md7IDB8eDFgNzgO3gSNgntupiy1-PfkagtYQUczWF3Sw0M5VtUwDZ9Pvvuw

Xu, K., Nosek, B., & Greenwald, A. G. (2014). Psychology data from the race implicit association test on the Project Implicit demo website. *Journal of Open Psychology Data, 2*(1), e3. https://doi.org/10.5334/jopd.ac