

Development and application of a self-report measure for assessing sensitive information disclosures across multiple modes

Matthew D. Pickard¹ · D. Wilson² · Catherine A. Roster¹

Published online: 25 August 2017
© Psychonomic Society, Inc. 2017

Abstract Building on the literature that approaches self-disclosure as a decision-making process, we proposed a self-reported Sensitive Information Disclosure (SID) measure and tested the measure's reliability and validity in two studies across a variety of interview modes and settings. We used theory to identify potential dimensions of sensitive information disclosures, created potential scale items, performed two separate card sorts, and validated the resulting pool of items in two separate experiments. Participants answered the SID scale items following an interview involving sensitive information, potential risk, and after-disclosure vulnerability. Study 1 was a laboratory experiment conducted with 165 university students. Exploratory factor analysis results revealed a two-factor structure, Personal Discomfort and Revealing Personal Information. Study 2 replicated these procedures using confirmatory factor analysis to confirm the factor structure and demonstrate the scale's reliability and validity, with a sample of 77 students and 275 participants from Amazon's M-Turk. Together, these results demonstrate that the proposed 11-item SID scale has good convergent and discriminant validity as well as good reliability. A quasi-experimental application of the measure is illustrated using the substantive findings from Study 2. This research fills a gap in the literature by developing a topic-free scale to measure SID as a dependent variable. The ability to accurately measure sensitive information disclosure is an important and necessary step toward developing a more thorough understanding of how people feel and react

when asked to provide personal information in diverse interview settings.

Keywords Self-disclosure · Sensitive information · Sensitive question · Data collection methods

Gathering sensitive information from individuals is an integral but challenging task for researchers across multiple disciplines in the health, behavioral, and social sciences. Because of the risks involved, individuals are often reticent to disclose sensitive information, which can impact both the amount and validity of the data collected by sensitive-topic researchers (Gnambs & Kaspar, 2015; Lee & Renzetti, 1993). Factors shown to affect the disclosure of sensitive information include interview or survey mode (Kays, Gathercoal, & Buhrow, 2012; Tourangeau & Smith, 1996; Uriell & Dudley, 2009), topic content (Tourangeau & Yan, 2007), interviewer or respondent characteristics (Lind, Schober, Conrad, & Reichert, 2013), question format (Bradburn & Sudman, 1979; Roster, Albaum, & Smith, 2014), and the context or interview setting (Tourangeau, Couper, & Steiger, 2003), to name a few. Therefore, research investigating methods to maximize the willing disclosure of sensitive information represents an important research activity in and of itself.

Although advances in technology have expanded options for research investigators to collect sensitive information (e.g., Internet surveys, computer-assisted self-interviewing, and virtual worlds), metrics for assessing single-episode, vulnerability-oriented disclosures do not currently exist. The present study was undertaken to remedy this situation. We sought to create a measure for assessing sensitive information disclosures that could be used by sensitive-topic researchers to explore the effectiveness of alternative methods for data

✉ Catherine A. Roster
roster@unm.edu

¹ Anderson School of Management, University of New Mexico, Albuquerque, NM, USA

² University of Oklahoma, Norman, OK, USA

collection across or within a variety of interviewing modes, whether in-person or computer-assisted.

In the following sections, we review the self-disclosure literature, with an emphasis on existing measures. We discuss the numerous shortcomings in the current conceptualization and measures of self-disclosure as they pertain to comparing the effectiveness of different interview modes designed to elicit self-disclosures. Building on these shortcomings, we propose an expanded conceptualization of sensitive information disclosure (the Sensitive Information Disclosure [SID] scale) that is risk-oriented, an aspect missing from many current scales and one that is best assessed through self-report measures. We then describe the SID scale development process and report findings from two separate studies to test the reliability and validity of the scale. Substantive results from Study 2, in which the scale was used as the dependent variable in a multimode quasi-experiment, illustrate how the scale's unique properties can be used to derive insights previously masked in similar studies comparing disclosures of sensitive information across interview modes.

Background

Definitions of sensitive information and relationship to disclosure

Various definitions for sensitive-topic research have been offered, most involving some element of threat, vulnerability, or risk to the participant should the information be disclosed to a third party. For instance, Sieber and Stanley (1988, p. 49) defined sensitive research as studies “in which there are potential consequences or implications, either directly for the participants in the research or for the class of individuals represented by the research.” Lee and Renzetti (1993) similarly defined a sensitive topic as “one that potentially poses for those involved a substantial threat, the emergence of which renders problematic for the researcher and/or the researched the collection, holding, and/or dissemination of research data” (p. 5). Researchers often assume that topics involving social taboos or other illicit behaviors, such as drug use, lying, or cheating, are threatening because they conflict with social norms. Gnambs and Kaspar (2015) described sensitive questions as those that “address highly personal and sometimes even distressing topics which are often in conflict with social norms and frequently result in socially desirable answers or even non-response” (p. 1237).

“Sensitivity” is linked to social desirability, as in the above definition provided by Gnambs and Kaspar (2015). The two constructs are related, in the sense that sensitive topics often elicit socially desirable responses as people seek to portray themselves in a manner that adheres to social norms and “looks good to an interviewer.” What distinguishes the two

constructs is that social desirability arises from the sensitivity of the answer, not the sensitivity of the question (Krumpel, 2013). Answers to questions that suggest deviations from social norms are seen as socially undesirable, whereas self-reports of behaviors that conform to social norms are considered socially desirable. This leads to a tendency for respondents to underreport socially undesirable behaviors and to overreport socially desirable behaviors. Thus, “social desirability” refers to reporting an attitude or behavior subject to social norms in a matter that presents the respondent in a positive light. Sensitivity is a much broader concept. Lee and Renzetti (1993) emphasized that sensitivity “inheres less in the topic itself and more in the relationship between that topic and the social context within which the research is conducted” (p. 5). Tourangeau and Yan (2007) included social desirability as one of three distinct aspects of “sensitivity” in their definition, which also includes the dimensions of “intrusiveness” and “threat of disclosure,” both of which align more with the context in which questions are posed than with the response given by respondents. Therefore, both the sensitivity of the response and the feelings evoked by the context in which the question was asked are necessary for a comprehensive view of “sensitivity.”

What is consistent in the definitions of sensitive information above is that each implies an element of distress, controversy, or concern that poses at least a moderate threat to those from whom the information is elicited and/or the researchers involved in the collection of that information. For these reasons, people are naturally reticent to disclose sensitive information, and therefore may actively engage in attempts to misrepresent their true attitudes or behaviors in an attempt to circumvent revealing personal information, if they chose to disclose at all. These natural inclinations and behaviors on the part of respondents are problematic for sensitive-topic researchers because validity of people's responses depends on *both* the sensitivity of the topic and the degree to which people are willing to disclose sensitive information to others, including an interviewer (Jourard, 1971). Lee and Renzetti (1993, p. 3) noted that a problem with research that involves sensitive information is that the term “sensitive” is often treated in the literature as if it were self-explanatory. Researchers often assume from the nature of the topic that it is sensitive according to some culturally derived social standard, or by inferring sensitivity from respondents' behaviors, such as item omissions. Neither represents a pure indicator of a question's sensitivity, since what is regarded as sensitive to one person or group may be regarded by others as completely innocuous (e.g., de Jong, Pieters, & Stremersch, 2012), and item omissions can arise from multiple factors, some of which are unrelated to question sensitivity (Beatty & Hermann, 2002).

Lee and Renzetti (1993, p. 6) have offered a definition of sensitive information that associates risk with context, rather than with nature of the topic itself. These researchers suggest

four conditions in which research is more likely to be perceived as “risky” than others. These include: (1) where research intrudes into the private sphere or delves into a deeply personal experience; (2) where the study is concerned with deviance and social control; (3) where it impinges on the vested interests of powerful persons or the exercise of coercion or domination; and (4) where it deals with things sacred to those being studied that they do not wish profaned. None of these conditions in and of themselves are necessarily threatening, but depending on the context within which the research is conducted, each could contribute to higher levels of emotional distress and a heightened sense of risk for research participants. In turn, conditions surrounding the research context can impact how participants perceive and calculate if and how much to disclose.

Risk and self-disclosure

The more recent, and arguably the more advanced, literature on self-disclosure has theorized when, why, and how people disclose. We refer to this as the “decision-making perspective.” At the core of most disclosure decision-making models is *risk*—to self, to others, and/or to relationships (Afifi & Steuber, 2009; Fisher, 1984, 1986; Greene, 2009; Petronio, 2002). Individuals may fear being personally judged, embarrassed, or vulnerable, or fear exposing others to being judged, embarrassed, or vulnerable. Consequently, individuals guard sensitive information not only to protect themselves and others, but also to protect the relationships they have with others. Disclosure decision-making models suggest that individuals engage in a cost–benefit calculus when deciding whether and how much sensitive information to reveal. Disclosure occurs when the benefits outweigh the quantity and weight of the risks involved (Afifi & Steuber, 2009; Omarzu, 2000; Petronio, 2002; Smith, Dinev, & Xu, 2011).

Petronio (2002) identified five types of risks involved in disclosing personal information: security, stigma, face, relational, and role. Although these risks are not mutually exclusive, their delineation highlights the broad and complex spectrum of potential costs one must consider before disclosing sensitive information. It should be noted that risk is especially salient in interview contexts because the relationship between the interviewer and interviewee is usually one of low intimacy, trust, and rapport, leaving interviewees uncertain as to how their sensitive information will be treated. Few interpersonal conditions exist in typical interviews to mitigate the risk interviewees perceive in revealing sensitive information.

Importantly, risk translates to potential vulnerability, or susceptibility to harm. Therefore, when individuals calculate the cost-benefit outcome of disclosing certain information, they attempt to ascertain the potential harm that could result from disclosure. Most disclosure decision making models, including Petronio’s (2002) communication privacy management

model and Omarzu’s (2000) disclosure decision model, posit that an individual’s potential vulnerability is directly related to the sensitivity of the information. The more recent Afifi and Steuber (2009) revelation risk model (RRM) is also grounded in this risk assessment tenet. RRM explicitly agrees with CPM that “revealing sensitive information about the self is risky and makes people feel vulnerable” (Afifi & Steuber, 2009, p. 147). RRM posits that people’s willingness to disclose private information decreases with increasing risks; and, conversely, their willingness to reveal such information is increased when risks are low or mitigated. There is consensus across disclosure decision-making models that an individual’s vulnerability to harm arising from disclosure is directly related to the sensitivity of the information.

The difference between perceived risk and vulnerability is at the core of our conceptualization of SID. The possessor of sensitive information has substantial control over the actualization of harm—the possessor may simply choose to conceal the information. The risk associated with sensitive information is only actualized through its disclosure; and the result of sensitive information disclosure is a heightened state of vulnerability. Furthermore, we propose a difference between *perceiving* and *feeling* a risk. Whereas perceptions are largely cognitive-based, feelings are largely affect-based. The predisclosure perception of risk, which informs the risk–benefit calculus in disclosure decision making, is more cognitive; in contrast, the postdisclosure vulnerability is more emotional. Therefore, we argue that sensitive information disclosure should induce vulnerability, or a feeling of being at risk; otherwise, the information is not truly sensitive.

Few, if any, existing measures of self-disclosure capture the emotional, risk-based component of sensitive information disclosures prevalent in models explaining why people disclose. This is one problem with existing self-disclosure measures, but a number of other shortcomings also prevent existing self-disclosure measures from being useful for evaluating the effectiveness of different interviewing methods.

Existing self-disclosure measures

Existing self-disclosure measures have approached self-disclosure from three main perspectives—the *trait* perspective, the *state* perspective, and the *message* perspective. Below, we briefly review each of these perspectives and discuss the deficiencies of each for comparing and assessing different interview modes to elicit sensitive information.

Trait perspective scales The trait perspective treats self-disclosure as an enduring tendency rooted in individuals’ personalities, which predispose them to be open with others (Jourard & Lasakow, 1958; Jourard & Resnick, 1970; Marshall, 1970; Rickers-Ovsiankina, 1956; West & Zingle, 1969). Trait-based self-disclosure scales commonly assess a

matrix of topics and targets to determine how disclosers' tendencies change across a variety of sensitive/nonsensitive topics and familiar/distant targets (Jourard & Lasakow, 1958; Rickers-Ovsiankina, 1956). For instance, the Self-Disclosure Situations Survey takes into account varying situational details that influence individuals' tendency to disclose (Chelune, 1976; Marshall, 1970). Some scales are tailored to specific populations (West & Zingle, 1969). Recent variations attempt to measure individuals' tendency to conceal (rather than disclose) sensitive information (Kahn & Hessling, 2001; Larson & Chastain, 1990). Regardless of the variation, trait perspective measures view self-disclosure as an independent variable, which makes these measures inappropriate for comparing interview modes in which self-disclosure is a dependent variable.

State perspective scales The state perspective treats self-disclosure as a singular behavioral event that can be measured as a dependent variable (Suchman, 1965; Vondracek, 1969). Although this dependent variable focus is appropriate for comparing the effectiveness of different interview modes, almost without exception, state-based measures either (1) employ context- or topic-specific information in their assessments or (2) require independent raters to judge the disclosures. The scale that S. I. Vondracek and Vondracek (1971) used is an example of a context- or topic-specific scale. When these researchers evaluated “deep disclosure of a transgression,” they included “swimming in abandoned quarries” as an example of a deep disclosure. The second limitation becomes salient should there be a need to measure individuals' vulnerability following a disclosure, since independent raters lack accessibility to disclosers' psychological and emotional structure. Therefore, it is difficult for raters to assess the disclosers' vulnerability. Likewise, it is difficult for raters to assess the inherent risk disclosers perceive in connection with their potential disclosures because such risk perceptions tend to be subjective. This general lack of accessibility to disclosers' psychological and emotional structure results in disclosure assessments based on social norms and not upon the subjectivities of disclosers' situational details.

Message perspective scales The message perspective views self-disclosure as a unit of communication that contains sensitive information about oneself and, in general, seeks to evaluate the intimacy and amount of the information contained in the message in an objective manner (Cozby, 1973; Wheelless & Grotz, 1976). Cozby identified three basic parameters to evaluate a disclosure message—namely, breadth (i.e., topic coverage), depth (i.e., intimacy or sensitivity), and duration (i.e., length of verbalization). A deep disclosure necessarily contains sensitive information that would increase vulnerability if the information were to be divulged. Breadth is depth's necessary, horizontal counterpart to a complete disclosure,

since there is a difference between addressing all the necessary parts (breadth) and thoroughly divulging the specifics of each part (depth). There is no strong theoretical or empirical evidence that duration is indicative of the quality or quantity of a disclosure (Bloch & Goodstein, 1971; Chelune, 1975). A person may verbalize extensively without revealing private or sensitive information.

Importantly, a disclosure message can contain both verbalized and affective information (Chelune, 1975; Howell & Conway, 1990), and further, affective information can be manifest both verbally and nonverbally. For instance, there is a distinct difference between an individual stoically declaring, “I stole a car today,” and an individual shamefully and hesitantly admitting, “I stole a car today.” The first individual may feel considerably less risk in the disclosure than the second. Emotional information is generally perceived as sensitive information (Chelune, 1975; Howell & Conway, 1990; Pasupathi, McLean, & Weeks, 2009); therefore, deep disclosures often involve emotional information. Chelune (1975) argued that “the inclusion of the affective dimension... would make possible a more precise assessment of the total amount of information departed in a verbal communication” (p. 82). Therefore, in the self-report SID scale we developed, we sought to capture this affective dimension of making self-disclosures.

Goals for a proposed measure of sensitive information disclosures

The purpose of the present study was to develop and validate a psychometrically sound self-reported measure of SID for use in interview situations, whether administered by humans, avatars (i.e., virtual humans), computer-based questionnaires, or through paper-based questionnaires. Unlike the current measures, our proposed SID assessment measure is designed for settings in which parties are not closely related, as in general academic research studies investigating sensitive topics. We sought to create a scale that would be topic-free, thereby allowing the individual to infer how sensitive and risky his/her disclosure is as opposed to assuming vulnerability by the nature of the topic. In keeping with the decision-making risk-assessment perspective of self-disclosure, we also designed the scale to assess the related but distinct affective and perceptual consequences associated with disclosing sensitive information.

Development of items for the SID scale

To develop the SID scale, we identified potential dimensions of SID through a literature review, used each identified dimension to create a pool of scale items, then performed two separate card sorts to cull and refine the scale item pool. We then tested the remaining items in two separate laboratory

experiments. The details of the scale development and validation are described below.

Item generation and card sort

We began by reviewing theoretical models describing the disclosure process (e.g., Afifi & Steuber, 2009; Omarzu, 2000; Petronio, 2002). We specifically identified concepts linked with emotional, cognitive, and psychological reactions to disclosing (e.g., negative emotions or discomfort, private information, social risk, social control, depth of disclosure, and disclosure difficulty). Benefits of disclosure (e.g., catharsis) were also included in the list of potential SID dimensions. The objective of this stage was to ensure content validity; therefore, we included concepts that were tangentially related to SID and likely correlated with other concepts identified as potential SID dimensions. On the basis of our review, we created an initial pool of 42 scale items to capture each SID dimension, all fashioned as Likert statements (1 = *strongly disagree*, 5 = *strongly agree*).

To perform an initial assessment of construct validity for the pool of items, we enlisted the aid of both expert and non-expert judges in two separate card sorts. In the first card sort, we asked a panel of four expert judges trained in survey measurement procedures to examine the pool of items. We invited nine MIS Ph.D. students at a major southwestern U.S. university, who had recently completed a course on instrument development that covered card sorts, to participate in the initial card sort and serve as expert judges. Four of these Ph.D. students participated. We did not provide them monetary compensation or other incentive for their time; they participated in the card sort voluntarily. We instructed the judges to group the scale items into categories and provide their own label for each category; therefore, we did not constrain how they grouped the items or labeled the groups of items. The judges also provided feedback regarding wording, conceptual distinctions among the various groups of items, and so forth. We removed items that judges did not consistently group together (e.g., “I am glad to finally get this off my chest” and “I feel relieved after answering this question”) or closely duplicated other items or belonged to groups with many other similar items (e.g., “Answering this question made me feel violated”). The pool was culled to 30 scale items.

We then administered this refined pool of items to 72 untrained judges (with no knowledge of the study context or of scale development procedures) that we recruited from undergraduate business courses at a major southwestern U.S. university. We used the OptimalSort card sorting web application from Optimal Workshop¹ to administer and analyze this second card sort. As with the expert judges, we instructed the untrained judges to sort the items into categories and label

each category. The 72 judges took an average of just under 12 min to complete the exercise, and sorted the items into groups ranging from two to nine ($M = 5.21$ groups, $SD = 1.64$ groups). On the basis of the similarity matrix and dendograms (i.e., hierarchical cluster trees) produced by OptimalSort, we eliminated three items because the judges did not group them consistently or grouped them with items that were very similar. An additional eight items were removed from the pool for further testing by the researchers because upon reflection, they lacked content validity in relation to theoretical dimensionality of the SID scale, which was designed to capture vulnerability and risk associated with disclosing sensitive information (e.g., “It was difficult to answer this question because I didn’t know what it was asking” and “I found it difficult to answer this question because it was hard to remember the details”), or because they assumed an interviewer’s presence, as the intention was for the scale to be agnostic to methods (e.g., “I did not trust the interviewer sufficiently to answer this question truthfully”). The reliability and validity of the remaining 19 SID scales items were then tested in two separate experiments in which the items appeared in a post-survey following an interview.

Study 1

The analysis objectives for data from the first study were to explore the factor structure of the 19 SID items and assess their overall internal consistency. Procedures and results from Study 1 are described below.

Method

Participants A total of 165 individuals participated in the study. Volunteers were recruited from a variety of undergraduate courses at a public university in southwestern United States. Study participants were offered extra credit for their participation. Ninety participants were female, 72 were male, and three refrained from indicating their gender. The average age was 24.88 ($SD = 6.87$). Ethnicity and other demographic information were not elicited to help insure participant privacy.

Procedures The participants were randomly assigned to one of three interviewer conditions (human, automated avatar, or automated audio-only). In each interview condition, the participants were asked a series of 16 questions (see Table 1). All questions were posed in an open-ended format and were designed to elicit disclosures of personal information that varied in positive/negative valence. Examples of interview questions asked include “What do you feel most guilty about in your life?,” “Think of someone you deeply love. Explain to me why you love this person,” and “What do you like to do for

¹ <https://www.optimalworkshop.com/>

Table 1 Study 1 questions by block

Block 1	
Q1.	What do you fear most?
Q2.	What do you feel most guilty about in your life?
Q3.	Please provide me a range that captures the average number of times in a week that you lust after, that is have sexual desires towards, someone. Give me an upper and lower number.
Q4.	Describe the most boring thing you have ever done.
Q5.	What in life gives you the most joy?
Q6.	Think of someone you deeply love. Explain to me why you love this person.
Q7.	What do you like least about your physical appearance?
Q8.	Please provide me a range that captures the average number of times a month you don't wash your hands after using the restroom. Give me an upper and lower number.
Block 2	
Q9.	Please provide me with a range that captures your actual weight. Please provide an upper weight and a lower weight.
Q10.	Please provide me a range that captures the number of push-ups you can do. Provide me an upper and lower number.
Q11.	Tell me about a time that a friend or family member disappointed you or hurt your feelings.
Q12.	Tell me about a time when you have been ungrateful towards someone close to you.
Q13.	In what ways are you different in public than you are in private?
Q14.	Please provide me with a range in dollars within which your income or your family's income falls. Please provide a lower dollar amount and an upper dollar amount.
Q15.	What do you like to do for fun with your closest friends or family?
Q16.	Tell me about a time when you felt deep gratitude for something or someone.

fun with your closest friends or family?" Following the interview, the participants completed the 19 SID items in a post-survey. To avoid survey fatigue, we employed a split-ballot technique. This created two separate question blocks each containing eight questions. Each participant completed the SID items for two of 16 interview questions, one randomly selected from each block.

Results and discussion

Because the participants each completed the SID scale items for two randomly systematically chosen interview questions, one for Question Block 1 and one for Question Block 2, we conducted two separate exploratory factor analyses (EFAs). For our factor analysis, we used principal axis factoring with oblique (promax) rotation in recognition of the fact that our two dimensions, personal discomfort and (revealing) personal information, would naturally be somewhat correlated, as opposed to orthogonal. These results appear in Tables 2 and 3. In our separate question block analyses, we applied stringent criteria for the inclusion of items—specifically, (1) items that cross-loaded $>.40$ across factors in either Question Block 1 or 2 were to be excluded for further consideration, and (2) only those items that loaded $>.40$ within the same factor across both blocks were to be retained. Thus, scale items had to consistently meet our criteria for both question block EFAs to be included in the scale. As was recommended by Henson and Roberts (2006), we considered in our interpretation

both the factor pattern and factor structure matrices, and report both in our tables.

Eleven scale items were retained for further testing following our EFA procedures. These items formed two distinct factors: (1) *Personal Discomfort* (PD), and (2) *Revealing Personal Information* (PI). Table 4 shows the seven PD items and the four PI items that met our criterion for inclusion in the SID scale and further testing. The two-factor solution explained over 50% of variance for each question block. To assess reliability, we calculated coefficient omega with 1,000 bootstrap simulations using MBESS version 4.2.0 in R (R Development Core Team, 2014), as described by Dunn, Baguley, and Brunnsden (2014): coefficient omega for the seven-item PD dimension = $.92$, $SE = .008$, 95% CI $[.90, .93]$; coefficient omega for the four-item PI dimension = $.85$, $SE = .015$, 95% CI $[.82, .88]$. These results indicate that both subscales of the SID have good internal consistency reliability across both question blocks.

Study 2

In Study 2 we confirmed the factor structure of the SID scale with a larger and broader population sample, including non-students and different interview settings. In the initial stages of measurement design, small to modest sample sizes are

Table 2 Study 1 principal axis pattern/structure matrix for Sensitive Information Disclosure (SID) items with promax rotation, Block 1 questions

Variable	Factor 1 (Pattern/Structure)	Factor 2 (Pattern/Structure)	<i>h</i> ²
Hedg1	-.158/.319	.128/.160	.298
D1	-.049/-.129	.294/.318	.434
D2	-.032/.333	.453/.607	.421
D3	-.170/-.251	.346/.222	.383
D4	-.021/-.553	-.145/-.217	.653
NE1	.712/.840	.095/.516	.715
NE2	.864/.826	.112/.445	.721
NE3	.753/.775	-.044/.455	.614
PI1	.080/.521	.819/.848	.724
PI2	.029/.359	.851/.783	.640
PI3	.001/.520	.708/.811	.681
PI4	.020/.378	-.069/.441	.613
PI5	.117/.464	.161/.566	.539
PI6	.161/.578	.627/.675	.544
R1	.927/.851	.069/.536	.742
R2	.785/.821	.116/.565	.694
R3	.107/.529	.095/.402	.416
R4	.637/.764	-.070/.398	.610
R5	.816/.795	-.148/.393	.649
EV	7.52	2.32	9.84
% of variance	39.56	12.23	51.78

Percentage variance is postrotation. The eigenvalue of the third, unretained factor was .660. *h*² = communality coefficient. Loadings greater than .40 with no cross-loadings are italicized and were retained for that factor based on the results for Question 1

sufficient. However, larger samples are desirable in later stages of scale development to confirm the scale’s performance can be replicated with different samples and in diverse contexts (DeVeillis, 2003; Netemeyer, Bearden, & Sharma, 2003). In addition, we further ensured that the construct was conceptually distinct from other, related concepts. So as not to fatigue our respondents given the larger set of measures to be tested, we designed Study 2 so that the participants completed the SID scale for a single interview question that varied randomly in terms of valence, either positive or negative, as questions with different valences can evoke different reactions from interview participants in terms of emotional distress and perceived risk.

Method

Participants The participants included 352 individuals who completed the survey using different interview methods. Seventy-seven participants were recruited from a variety of undergraduate courses at a public university in the

Table 3 Study 1 principal axis pattern/structure matrix for SID items with promax rotation, Block 2 questions

Variable	Factor 1 (Pattern/Structure)	Factor 2 (Pattern/Structure)	<i>h</i> ²
Hedg1	-.084/.336	-.002/.196	.247
D1	-.142/-.042	.801/.595	.519
D2	-.041/.439	.552/.665	.508
D3	-.255/-.285	.558/.289	.353
D4	.025/.517	-.081/.225	.473
NE1	.873/.833	-.027/.345	.695
NE2	.845/.831	-.024/.350	.691
NE3	.639/.758	.142/.457	.596
PI1	.118/.483	.699/.782	.636
PI2	.168/.402	.571/.635	.421
PI3	-.058/.482	.645/.760	.652
PI4	.223/.411	.516/.593	.380
PI5	.196/.587	.390/.599	.522
PI6	.233/.509	.301/.484	.357
R1	.830/.865	.052/.427	.750
R2	1.02/.891	.055/.423	.811
R3	.117/.536	.076/.343	.411
R4	.644/.764	-.100/.278	.608
R5	.773/.793	.012/.363	.629
EV	7.70	1.99	9.69
% of variance	40.53	10.48	51.01

Percentage variance is postrotation. The eigenvalue of the third, unretained factor was .570. *h*² = communality coefficient. Loadings greater than .40 with no cross-loadings are italicized and were retained for that factor on the basis of results for Question 2

southwestern United States. Interviews were conducted using a human interviewer. Two-hundred and seventy-five participants were recruited from Amazon’s Mechanical Turk. The Mechanical Turk participants were either interviewed by an avatar (*N* = 148) or completed an online questionnaire (*N* = 127). The university participants were offered extra credit for their participation. The Mechanical Turk participants were paid for their time. Overall, 197 of the participants were female, 153 were male, and two refrained from indicating their gender. The average age was 32.77 years (*SD* = 10.34). Ethnicity and other demographic information were not elicited to help insure participant privacy. It is important to note that our methods for assigning participants to different interview settings was quasi-experimental, as opposed to a randomized, true experiment. A detailed description of our respondents by age and gender for each interview method is provided in Table 5.

Procedure In all three types of interviews (human, avatar, or online questionnaire), the interviewer asked three questions:

Table 4 Study 1 SID items that met the criteria in Question Blocks 1 and 2 following the EFA analysis

Item Description	Criteria Met*	
	Block 1	Block 2
FACTOR 1 – <i>Personal Discomfort</i> [PD]		
I felt embarrassed answering this question. [NE1]	*	*
I felt uncomfortable answering this question. [NE2]	*	*
Answering this question was scary. [NE3]	*	*
I felt vulnerable answering this question. [R1]	*	*
I felt exposed answering this question. [R2]	*	*
I felt threatened by this question. [R4]	*	*
Answering this question made me feel violated. [R5]	*	*
FACTOR 2 – <i>Revealing Personal Information</i> [PI]		
To provide a truthful answer, I provided a detailed response to this question. [D1]		*
I disclosed a lot in order to answer this question truthfully. [D2]	*	*
I provided a comprehensive answer to this question. [D3]		*
I revealed sensitive information in response to this question. [PI1]	*	*
I revealed private information answering this question. [PI2]	*	*
I revealed secrets in response to this question. [PI3]	*	*
I “bared my soul” when answering this question. [PI4]		*

*Criteria = loadings >.40 on the respective factor in Block 1 or 2, excluding items that failed to meet this criteria for either factor. Of these, only the 11 items that met the criteria in both blocks were retained for the model. For PD, seven items met the inclusion criteria; for PI, four items could be included. The component correlation between the two factors was .187

(1) “What do you like to do for fun with your closest friends of family?,” (2) “What do you feel most guilty about in your life?,” and (3) “Think of someone you deeply love. Explain to me why you love this person.” Question 1 was designed to be a nonsensitive lead-in question. Questions 2 and 3 were designed to be negatively valenced and positively valenced sensitive questions, respectively. The order of Questions 2 and 3 was randomized, and the participants only completed the SID scale for Question 2 or 3, whichever they were asked first.

Results and discussion

As was suggested in the prior literature (Gerbing & Anderson, 1988), we supplemented our EFA with a confirmatory approach on the factor structure implied by the EFA and other

conceptually related constructs. This strategy has several advantages over relying solely on an EFA. Importantly, a confirmatory approach has the ability to directly assess a construct’s dimensionality (Anderson & Gerbing, 1991; Gerbing & Anderson, 1988). Furthermore, establishing a construct’s dimensionality, as well as its reliability, is necessary but insufficient for fully establishing construct validity (Gerbing & Anderson, 1988). The researcher must also demonstrate that the construct is conceptually distinct from other constructs. The confirmatory analyses were performed using lavaan version 0.5-23.1097 (Rosseel, 2012) in R. We confirmed the two-factor structure suggested by the EFA and demonstrated that the SID scale items performed well in the context of five conceptually related constructs—*self-concealment* (Larson & Chastain, 1990), *self-monitoring* (Fenigstein, Scheier, & Buss, 1975), *private self-consciousness* (Fenigstein et al., 1975),

Table 5 Demographic composition of Study 2 respondents by gender and age within interview conditions

Interview Condition (<i>n</i>)	% Gender (M/F)	Age (mean/median)	Minimum Age	Maximum Age
Online survey (126)	49%/51%	36/32	19	69
Avatar interview (146)	63%/37%	32/29	18	64
Human interview (77)	55%/45%	29/26	20	68

Gender differences are significant ($\chi^2 = 5.57$, $df = 2$, $p < .10$). Age differences are also significant ($F = 12.84$, $df = 2$, $p < .01$) between interview conditions

Table 6 Confirmatory factor analysis loadings for the two-factor SID model in Study 2

Variable	Factor 1 PD	Factor 2 PI
NE1	.80	
NE2	.83	
NE3	.80	
R1	.74	
R2	.74	
R4	.78	
R5	.75	
D2		.65
PI1		.78
PI2		.86
PI3		.71

$N = 352$. Loadings estimated using maximum likelihood estimation. PD = Personal Discomfort; PI = (revealing) Personal Information

public self-consciousness (Scheier & Carver, 1985), and *social anxiety* (Scheier & Carver, 1985). The confirmatory factor loadings for the two SID scale factors are provided in Table 6.

We then performed a confirmatory factor analysis (CFA) with the SID items forming the two-factor structure indicated in our prior results, including the measurement items from each of the other constructs. We then fit the model to the data. The confirmatory model converged successfully and exhibited acceptable fit to the data [$\chi^2(385) = 615.91, p < .001, \chi^2/df = 1.60, CFI = .96, TLI = .95, RMSEA = .04, SRMR = .05$]. Satisfied that the model was a good fit to the data, we calculated correlations, reliabilities, and the average variance extracted (AVE) to further aid in establishing factorial validity. To demonstrate factorial validity, convergent validity for the construct must be demonstrated with an AVE > .5 (Hair, Black, Babin, & Anderson, 2010). In addition to convergent validity, factorial validity requires discriminant validity. This is established when the square root of a construct's AVE is higher than the correlation between that construct and all other constructs in the model (Hair et al., 2010). Finally, to establish reliability, the composite reliability value for each latent variable should be $\geq .7$ (Fornell & Larcker, 1981; Nunnally & Bernstein, 1994). These metrics are summarized in Table 7. Using the guidelines described above, we note that the reliability, AVE, and correlations shown in Table 7 indicate excellent measurement properties.

Summary

The CFA results afford us two conclusions. First, the two-dimensional structure of the SID construct appears to fit the data quite well. The data therefore support the conclusion that SID is a second-order construct with two reflective, first-order constructs—personal discomfort and revealing personal

Table 7 Construct reliabilities, average variance extracted, and correlations for the confirmatory model (Study 2)

	CR	AVE	1	2	3	4	5	6	7
1. SelfCncl	.80	.57	.75						
2. SelfMon	.90	.66	.01	.81					
3. PrivSC	.80	.58	.23	.18	.76				
4. PubSC	.82	.54	.33	.00	.59	.73			
5. SocAnx	.84	.63	.43	-.28	.32	.54	.79		
6. PD	.90	.60	.50	.02	.15	.30	.28	.77	
7. PI	.83	.58	.15	.01	.05	.09	.08	.15	.76

$N = 352$; CR = composite reliability; AVE = average variance extracted; italicized values along the diagonal show the square root of each AVE value. SelfCncl = self concealment; SelfMon = self monitoring; PrivSC = private self-consciousness; PubSC = public self-consciousness; SocAnx = social anxiety; PD = Personal Discomfort; PI = (revealing) Personal Information

information. Second, the SID construct performs well in terms of both factorial validity and reliability in the context of several conceptually similar constructs. This establishes good construct validity.

Together, the results from our EFA and CFA analyses provide support for a two-factor, 11-item SID scale that demonstrates very good to excellent reliability and validity indices across multiple samples, interview methods, interview questions, and topic valence. The SID scale described in this article is conceptually unique relative to existing self-disclosure measures, in that the scale is designed to capture a participant's *postdisclosure felt* risks and heightened sense of vulnerability, which can be quite different from the perceived risks that inform participants' *cognitive* risk–benefit calculations *prior* to disclosing sensitive information. Because the risks associated with sensitive information disclosure are only truly *felt* through the act of disclosure (as the possessor of sensitive information can simply chose not to disclose, or give deceptive responses following a cognitive assessment of cost/benefits), we expect SID scores to increase when participants make disclosures in interview settings that increase their vulnerability. Specifically, in terms of the two SID factors, PI and PD, we expect participants, in general, to disclose *more* PI in *less* threatening interview modes (e.g., human interviewer) and *less* PI in *more* threatening interview modes (e.g., online survey). This is consistent with empirical findings in the literature (Gnambs & Kaspar, 2015; Tourangeau & Yan, 2007). In contrast, we expect participants to, in general, feel *more* discomfort and vulnerability after disclosing PI in *more* threatening interview modes and to feel *less* discomfort and vulnerability after disclosing PI in *less* threatening interview modes. We test these expectations in an application using data from Study 2 described below.

Application of the SID measure using the Study 2 substantive findings

We analyzed the substantive findings obtained from the 352 participants in Study 2 to illustrate how the proposed SID measure can be used to compare effectiveness of alternative interview modes used to collect sensitive information. Methods and procedures are described in the previous section. The results were analyzed using a two-way factorial MANOVA. The two categorical independent variables were topic (Guilt, Love), fashioned to represent negatively and positively valenced private information, respectively, and interview mode (online survey, avatar interviewer, human interviewer). For dependent variables, we created two separate averaged summated scores to represent the two factors in the SID scale, a PD score and a PI score. Gender and age were entered as covariates for exploratory purposes. Because our intention was to demonstrate an exploratory application of the scale, we did not craft formal hypotheses. Instead, we sought to answer the research question, “Does the degree of disclosure and affect generated by being asked to reveal sensitive information differ by topic valence and interview mode?” Of primary concern was the utility of our proposed SID measure in a multimode experimental study involving potentially sensitive interview topics, as opposed to shedding light on substantive issues regarding hypothesized differences in self-disclosure that might be offered in a study designed specifically to expose differences between groups based on multimode interview conditions. Due to the quasi-experimental nature of Study 2 and potential confounds arising from different sample populations and lack of random assignment to interview conditions, our results are offered merely as exploratory findings and serve primarily to illustrate how the SID scale operates.

Results and discussion

In short, the answer to our exploratory research question is “yes”: Both topic valence and interview mode affected the degree of disclosure and discomfort felt by study participants. Descriptive statistics appear in Table 8. The multivariate test results revealed a significant main effect of topic valence (Wilks’s $\lambda = 5.04$, $p < .01$), as well as a significant main effect of interview condition (Wilks’s $\lambda = 3.21$, $p \leq .01$). The Topic \times Interview Condition interaction was moderately significant (Wilks’s $\lambda = 2.10$, $p < .10$). Age was not significant (Wilks’s $\lambda = 2.17$, $p > .10$), and gender just met significance at the 90% confidence level (Wilks’s $\lambda = 2.37$, $p = .10$).

In light of the moderate significance found for the higher order interaction in our multivariate analysis of variance (MANOVA), we reverted to univariate ANOVAs for each dependent variable, separately, to aid our interpretation of effects. The findings appear in Tables 9 and 10. As for the main effect of question, the positive-valence question, “Think of someone you deeply love. Explain to me why you love this person,” created significantly less personal discomfort than did the negative-valence question “What do you feel most guilty about in your life?” Overall, in terms of interview modes, the online survey condition induced significantly less personal discomfort than the avatar condition. Though not significant, the mean of personal discomfort was lower for the online survey condition than for the human condition as well. The general direction of personal discomfort, therefore, aligned with our expectations that respondents would report *more* discomfort in the *more* threatening interview conditions. Respondents also reported significantly more personal discomfort with the guilt question than with the love question. We found no Topic Valence \times Interview Mode interaction for personal discomfort.

Table 8 Descriptive statistics for interview conditions by topic for sensitive information disclosure (SID) factors

SID Factors	Topic	Interview Condition	<i>n</i>	Mean*	<i>SD</i>
Personal Discomfort (PD)	Guilt	Online survey	62	3.03	1.54
		Avatar interview	73	3.86	1.54
		Human interview	35	3.38	1.44
	Love	Online survey	64	2.72	1.48
		Avatar interview	73	3.18	1.49
		Human interview	41	2.85	1.45
Revealing Personal Information (PI)	Guilt	Online survey	62	3.34	1.57
		Avatar interview	73	3.63	1.59
		Human interview	35	3.61	1.31
	Love	Online survey	64	3.56	1.65
		Avatar interview	73	3.32	1.52
		Human interview	41	2.59	1.09

* Average score based on seven PD items and four PI items, where 1 = *strongly disagree*; 5 = *strongly agree*

Table 9 Univariate two-way factorial ANOVA statistics for sensitive information disclosure factors

Dependent Variables	Model Variables	<i>F</i>	<i>p</i>	Eta-Squared
Personal Discomfort (PD)	Topic	7.62	.006**	.022
	Interview Condition	5.01	.007**	.029
	Topic × Interview Condition	0.52	.597	.003
Revealing Personal Information (PI)	Topic	4.26	.040*	.012
	Interview Condition	1.48	.228	.009
	Topic × Interview Condition	3.89	.021*	.022

* <.05, ** <.01

However, there was a significant interaction effect for Topic Valence × Interview Mode for the dependent variable PI. For the love question, respondents reported higher levels of revealing personal information in the online survey than either the avatar or human conditions. When answering the negative-valenced guilt question, respondents reported higher levels of revealing personal information in the human and avatar interview conditions than in the online survey condition.

Though there was not a significant difference in the interview conditions, the general trend of the means are as we expected—that is, respondents reported revealing *more* personal information in the *less* threatening interview modes. Of particular interest is the avatar interview mode in which respondents reported feeling the highest personal discomfort as well as revealing the most amount of personal information, nearly equivalent to the personal information they reported to reveal in the online survey mode. One potential explanation of this is that, in the moment of the interview, the respondents did not feel threatened or judged by the nonintelligent avatar, much like in the online survey condition; therefore, they

revealed more personal information than in the human interview condition. However, when completing the postinterview SID measurement items and contemplating on the amount of personal information they revealed to the human-like avatar, they felt more vulnerable. In other words, their higher levels of reported personal discomfort could be a reflection of their feelings of vulnerability when asked to disclose sensitive information to a human-like agent. An alternative explanation is that respondents simply felt uncomfortable talking with a human-like, but not human, interviewer, who was incapable of expressing empathy. In a scenario-based survey of 14 potentially sensitive topics conducted by Pickard, Roster, and Chen (2016), participants were asked whether they would be more likely to disclose sensitive information to a human or avatar interviewer, and to explain “why” in open-ended probes. Participants who selected disclosure to a human interviewer gave reasons such as “it feels more like a conversation” and “humans have compassion and understanding,” which highlight the social dimension of sensitive-topic disclosures emphasized by Lee and Renzetti (1993, p. 5). On the other hand,

Table 10 ANOVA and *t*-test results for the sensitive information disclosure factors

Independent Variables	Dependent Variables Mean (<i>SD</i>)	<i>t</i>	<i>p</i>	Cohen's <i>d</i>
Topic	Personal Discomfort	3.18	.002**	0.24
Guilt	3.46 (<i>1.57</i>)			
Love	2.94 (<i>1.48</i>)			
	Revealing Personal Information	1.88	.062*	0.14
Guilt	3.54 (<i>1.52</i>)			
Love	3.24 (<i>1.52</i>)			
		<i>F</i>	<i>p</i>	Eta-Squared
Interview Mode	Personal Discomfort	6.23	.002**	.035
Online survey	2.88 (<i>1.51</i>)			
Avatar interviewer	3.52 (<i>1.56</i>)			
Human interviewer	3.10 (<i>1.48</i>)			
	Revealing Personal Information	2.05	.131	.012
Online survey	3.46 (<i>1.61</i>)			
Avatar interviewer	3.48 (<i>1.55</i>)			
Human interviewer	3.08 (<i>1.29</i>)			

* <.10, ** <.01

the participants in the Pickard et al. study who indicated a preference for making disclosures to avatars stated reasons such as “an avatar cannot judge you,” and “I feel I would be able to say more easily without the pressure of an actual person.” These reasons could account for why respondents reported revealing more personal information to the avatar than to the human, despite their higher level of personal discomfort. The social dimension of interviews could also potentially explain why our respondents, in retrospect, stated they revealed more to both human or avatar interviewers than in the noninterviewer online survey, which was less threatening, but also devoid of any human or human-like social pressure to respond.

General discussion

In this article, we have developed a risk-oriented scale to measure sensitive information disclosures. Such a scale is needed because the majority of existing scales treat self-disclosure as a trait, making them unsuitable for situations such as interviews, in which the elicitation of sensitive information is the phenomenon of interest. Existing state-based scales are usually context-specific and are designed for use by third parties. This creates a challenge of measuring a subjective phenomenon in an objective way. What is sensitive to one person in a specific situation may not be judged sensitive to a different person in the same situation or even to that same person in a different situation. It is difficult for raters to evaluate sensitivity of a disclosure because they do not have access to the discloser’s psychological structure, which determines the sensitivity of a topic. The SID scale developed in this article helps fill these gaps. Importantly, it treats SID as a dependent measure, making it appropriate for research comparing the effectiveness of different methods and modes for eliciting sensitive information. Because it is a topic-free scale, it can be used across a variety of interviewing modes and contexts. The SID scale is also relatively short and easy to administer.

To our knowledge, this is the first scale that aligns with the disclosure decision-making literature and incorporates the relationship between vulnerability and risk. Decision perspective theories hold that a risk–benefit calculus is at the core of individuals’ decisions whether to disclose sensitive information. From this perspective, disclosure of truly sensitive information exposes individuals to increased risk, making them more vulnerable. According to the decision-making perspective of self-disclosure, the strength of the boundaries individuals erect correlates to the sensitivity of the information they seek to protect. This core concept in the disclosure decision-making research has not yet been captured in a measurable form. Of the two factors that emerged from the SID scale development, the personal discomfort factor corresponds to the vulnerability aspect we sought to measure. It captures

individuals’ emotional reaction (e.g., uncertainty, embarrassment, discomfort, fear, and vulnerability) to disclosing sensitive information. In contrast, the private information dimension captures individuals’ perception of the risk incurred by disclosing such information. In other words, the private information dimension measures whether anything sensitive was disclosed, whereas the personal discomfort dimension measures the discomfort felt as a result of the sensitive information that was disclosed. In this respect, the SID scale poses a clear advantage over existing scales because vulnerability is more likely to arise from the context in which the information is revealed, rather than from the nature of the topic, *per se*.

Our application of the SID scale in an exploratory multi-mode interview study revealed how the scale can be used to gather insights that are typically masked in studies of this nature. Due to potential confounds arising from our quasi-experimental treatment groups, any substantive findings must be regarded with caution. However, our application serves to illustrate the unique properties of the SID scale and how it operates. Because the SID separates the PI and PD dimensions of self-disclosure, our results were able to capture more clearly the relationship between sensitivity of the information requested and the affective and risk-related consequences participants felt after making disclosures. When information was regarded as less sensitive, vulnerability decreased, and participants felt more comfortable revealing private information, especially in the online survey condition, in which they retained greater anonymity. On the other hand, when the topic sensitivity increased as it did in the negative valence question, participants felt more vulnerable, especially in the human and avatar interviewer conditions. These results reveal how important it is to consider the degree of control the individual has over whether to reveal or not reveal private information. It appears that as vulnerability increased, so did participants’ perceptions of lack of control over their ability to safely divulge private information. The vulnerability respondents felt varied by interview mode for the PI dimension, indicating that the context intensified (or deintensified) participants’ feelings regarding the risk incurred from revealing private information. Simply asking participants to rate the sensitivity of the information requested would not fully capture how vulnerability informs cognitive risk assessments, nor the emotions elicited by making these requests under different interview conditions.

The SID scale builds on the assumption that the extent to which individuals feel personal discomfort in response to a disclosure is correlated with the extent to which they disclose truly sensitive information. A perception that information is sensitive is a necessary, but insufficient, condition for sensitive information disclosures. For example, overweight or anorexic people will likely find a question about their weight to be much more sensitive than people of average or healthy weights. A physically average individual responding to a question about his or her weight would not be considered to

have engaged in sensitive information disclosures because divulging such information is likely not psychologically risky. Further, it is possible for self-confident, overweight individuals to not be bothered by their weight. Such individuals would not perceive any social risk in revealing their true weight. We would again expect such disclosure to score low on our SID scale, even though such a disclosure might open them to public ridicule; they are apathetic to what other people think about their weight. Yet, in obvious contrast, those individuals who perceive their true weight as sensitive information cannot be considered to engage in disclosure if they do not reveal their true weight. Thus, the PI and PD dimensions comprise two complementary parts of the whole self-disclosure picture.

Future research

This aforementioned dichotomy between the PD and PI dimensions of the SID scale suggests several opportunities for future research. First, because the SID scale produces self-reported data, it can also measure the subjective vulnerability a disclosure induces. Although it is necessary for a risk-oriented SID measure to be self-reported, it is important to acknowledge that self-report data are notoriously prone to social desirability bias, especially when sensitive information is gathered (Gnamb & Kaspar, 2015; Tourangeau & Yan, 2007). Future research could utilize the SID scale as one of several measures that includes those capable of providing validation of truthfulness. For example, participants could be weighed after they were asked to disclose their weight. This would enable objective evaluation of the honesty of participants' responses, an obvious difficulty when soliciting from participants potentially sensitive disclosures based on attitudes or unobservable recollections. Such studies would allow researchers to directly assess the correlation between participants' honesty and the amount of after-disclosure negative emotion or vulnerability reported by participants. Additionally, studies that incorporate a qualitative analysis of the verbatim responses could provide insights into the correlation between self-report measures and elements of disclosure such as the breadth, depth, and content of disclosures.

Although the SID scale obviously enables researchers to measure SID as a dependent variable and thereby compare the effectiveness of different interviewing techniques, future research opportunities also abound in distinguishing between the elements that can cause non-disclosure. For example, people may not want, or be able, to recall a fitting answer to the question at hand. We suspect that some of our participants simply could not think of something they felt guilty about in the time allotted. Some participants may have arrived at a fitting answer if given enough time, whereas others may truly not feel guilty about anything. When evaluating the effectiveness of different methods of eliciting disclosure, it is important that such nonresponses not be confused with a lack of

willingness to disclose. Similarly, some people are better able to articulate sensitive information than others and some are more naturally prone to be open than others. Being able to independently measure these potential confounds will allow researchers to systematically determine whether their interview manipulations (e.g., environment, question phrasing, and interviewer characteristics) are effectively increasing individuals' willingness to disclose.

A potential limitation of the SID scale is that the measures are directed toward a specific question. For studies involving numerous sensitive questions/topics, this could be problematic as it would require repeating the scale multiple times. One solution would be the split-ballot method we employed in Study 1 that randomized the questions respondents rated on the SID scale. In many cases, sensitive information disclosure research deals with a limited set of topics, sometimes only one (e.g., drug use, sexual practices, income, etc.), which are assessed for sensitivity at the topic level. Future research could determine whether the SID scale items could be modified to replace wording of "this question" with "this interview" (allowing for a mix of topics), or even "questions about X topic," without compromising the scale's reliability or validity. Also, although the scale is relatively short (i.e., 11 items), future research could determine whether the scale could be shortened, particularly the Personal Discomfort dimension.

Another potential limitation of the SID scale is the necessity of analyzing its two factors, PI and PD, separately. Use of an aggregated SID measure without further analysis of the two factors could lead to the misinterpretation of findings, because the correlation between the two factors is expected to differ on the basis of how people assess, post-hoc, their feelings of vulnerability and the admissions they have actually made in situations in which they have been asked to reveal information. The correlation between the two dimensions of PI and PD is expected to be higher and positive to the extent that people actually disclosed personal information in interview conditions that were more threatening, which may seem counterintuitive, but actually reflects their discomfort about making revelations of a sensitive nature. Our application study revealed that the PI dimension moves consistent with the existing empirical findings, in that the means for revealing personal information were higher in both the online survey and avatar interview conditions (see Table 10). The PD dimension then sheds further light on the intensity and depth of the disclosure by measuring vulnerability. Personal discomfort was lowest in the online survey condition. Overall, these results correspond with our expectations as well as the extant disclosure literature.

Our SID scale solves an important problem for sensitive-information researchers, which is the lack of a relatively easy-to-administer self-report scale that can be used to measure sensitive information disclosure across multiple contexts and topics. However, research of sensitive information disclosures

involves numerous complexities and challenges for researchers, and our scale only partially resolves issues in this area. We recommend that future studies include additional measures alongside the SID. Among those most important to include are measures of interviewer/research agent trust, which may fully or partially mediate the relationship between interview conditions and sensitive information disclosures. By separating the two dimensions of SID, personal discomfort and revealing personal information, future research studies could more easily and more fully investigate the subtle nuances arising from both interviewer characteristics and interview conditions that create/inhibit trust and ultimately lead to the disclosure of sensitive information.

Conclusion

The present research has introduced the concept that sensitive information is risky to disclose and thereby should induce a feeling of vulnerability in the discloser when it is revealed. We proposed and developed a topic-free scale to measure sensitive information disclosure as a dependent variable. The SID scale fills a gap in the disclosure literature and can enable researchers and practitioners to effectively compare different data collection methods (e.g., Internet surveys, computer-assisted self-interviewing, and virtual worlds) with respect to eliciting sensitive information. The ability to accurately measure sensitive information disclosure is an important and necessary step toward developing a more thorough understanding of why people do or do not respond truthfully when asked to provide information.

References

- Afifi, T., & Steuber, K. (2009). The revelation risk model (RRM): Factors that predict the revelation of secrets and the strategies used to reveal them. *Communication Monographs*, *76*, 144–176. doi:10.1080/03637750902828412
- Anderson, J. C., & Gerbing, D. W. (1991). Predicting the performance of measures in a confirmatory factor analysis with a pretest assessment of their substantive validities. *Journal of Applied Psychology*, *76*, 732–740.
- Beatty, P., & Hermann, D. (2002). To answer or not to answer: Decision processes related to survey item nonresponse. In R. M. Groves, D. A. Dillman, J. L. Eltinge, & R. J. A. Little (Eds.), *Survey nonresponse* (pp. 71–86). New York, NY: Wiley.
- Bloch, E. L., & Goodstein, L. D. (1971). Comment on “Influence of an interviewer’s disclosure on the self-disclosing behavior of interviewees”. *Journal of Counseling Psychology*, *18*, 595–597.
- Bradburn, N. M., & Sudman, S. (1979). *Improving interview method and questionnaire design*. San Francisco, CA: Jossey-Bass.
- Chelune, G. J. (1975). Self-disclosure: An elaboration of its basic dimensions. *Psychological Reports*, *36*, 79–85.
- Chelune, G. J. (1976). Self-disclosure situations survey: A new approach to measuring self-disclosure. *JSAS Catalog of Selected Documents in Psychology*, *6*, 111–112.
- Cozby, P. C. (1973). Self-disclosure: A literature review. *Psychological Bulletin*, *79*, 73–91.
- de Jong, M. G., Pieters, R., & Stremersch, S. (2012). Analysis of sensitive questions across cultures: An application of multigroup item randomized response theory to sexual attitudes and behaviors. *Journal of Personality and Social Psychology*, *103*, 543–564.
- DeVeillis, R. F. (2003). *Scale development: Theory and applications*. Thousand Oaks, CA: Sage.
- Dunn, T. J., Baguley, T., & Brunsten, V. (2014). From alpha to omega: A practical solution to the pervasive problem of internal consistency estimation. *British Journal of Psychology*, *105*, 399–412.
- Fenigstein, A., Scheier, M. F., & Buss, A. H. (1975). Public and private self-consciousness: Assessment and theory. *Journal of Consulting and Clinical Psychology*, *43*, 522–527.
- Fisher, D. (1986). Decision-making and self-disclosure. *Journal of Social and Personal Relationships*, *3*, 323–336.
- Fisher, D. V. (1984). A conceptual analysis of self-disclosure. *Journal for the Theory of Social Behaviour*, *14*, 277–296. doi:10.1111/j.1468-5914.1984.tb00498.x
- Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, *18*, 39–50.
- Gerbing, D. W., & Anderson, J. C. (1988). An updated paradigm for scale development incorporating unidimensionality and its assessment. *Journal of Marketing Research*, *25*, 186–192. Retrieved from www.jstor.org/stable/3172650
- Gnambs, T., & Kaspar, K. (2015). Disclosure of sensitive behaviors across self-administered survey modes: A meta-analysis. *Behavior Research Methods*, *47*, 1237–1259. doi:10.3758/s13428-014-0533-4
- Greene, K. (2009). An integrated model of health disclosure decision-making. In T. D. Afifi & W. A. Afifi (Eds.), *Uncertainty, information regulation in interpersonal contexts: Theories and applications* (pp. 226–253). New York, NY: Routledge.
- Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2010). *Multivariate data analysis*. Upper Saddle River, NJ: Prentice Hall.
- Henson, R. K., & Roberts, J. K. (2006). Use of exploratory factor analysis in published research: Common errors and some comment on improved practice. *Educational and Psychological Measurement*, *66*, 393–416.
- Howell, A., & Conway, M. (1990). Perceived intimacy of expressed emotion. *Journal of Social Psychology*, *130*, 467–476.
- Jourard, S. M. (1971). *Self-disclosure: An experimental analysis of the transparent self*. New York, NY: Wiley.
- Jourard, S. M., & Lasakow, P. (1958). Some factors in self-disclosure. *Journal of Abnormal and Social Psychology*, *56*, 91–98.
- Jourard, S. M., & Resnick, J. L. (1970). Some effects of self-disclosure among college women. *Journal of Humanistic Psychology*, *10*, 84–93. doi:10.1177/002216787001000107
- Kahn, J. H., & Hessling, R. M. (2001). Measuring the tendency to conceal versus disclose psychological distress. *Journal of Social and Clinical Psychology*, *20*, 41–65.
- Kays, K., Gathercoal, K., & Buhrow, W. (2012). Does survey format influence self-disclosure on sensitive question items? *Computers in Human Behavior*, *28*, 251–256. doi:10.1016/j.chb.2011.09.007
- Krumpel, I. (2013). Determinants of social desirability bias in sensitive surveys: A literature review. *Quality and Quantity*, *47*, 2025–2047.
- Larson, D. G., & Chastain, R. L. (1990). Self-concealment: Conceptualization, measurement, and health implications. *Journal of Social and Clinical Psychology*, *9*, 439–455.
- Lee, R. M., & Renzetti, C. M. (1993). The problems of researching sensitive topics: An overview. In C. M. Renzetti & R. M. Lee (Eds.), *Researching sensitive topics* (pp. 3–13). Newbury Park, CA: Sage.
- Lind, L. H., Schober, M. F., Conrad, F. G., & Reichert, H. (2013). Why do survey respondents disclose more when computers ask the questions? *Public Opinion Quarterly*, *77*, 888–935.

- Marshall, M. J. (1970). *The effects of two interviewer variables in self-disclosure in an experimental interview situation*. Unpublished paper.
- Netemeyer, R. G., Bearden, W. O., & Sharma, S. (2003). *Scaling procedures: Issues and applications*. Thousand Oaks, CA: Sage.
- Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory*. New York, NY: McGraw-Hill.
- Omarzu, J. (2000). A disclosure decision model: Determining how and when individuals will self-disclose. *Personality and Social Psychology Review*, 4, 174–185.
- Pasupathi, M., McLean, K. C., & Weeks, T. (2009). To tell or not to tell: Disclosure and the narrative self. *Journal of Personality*, 77, 89–124. doi:10.1111/j.1467-6494.2008.00539.x
- Petronio, S. S. (2002). *Boundaries of privacy: Dialectics of disclosure*. Albany, NY: SUNY Press.
- Pickard, M. D., Roster, C. A., & Chen, Y. (2016). Revealing sensitive information in personal interviews: Is self-disclosure easier with humans or avatars and under what conditions? *Computers in Human Behavior*, 65, 23–30.
- Development Core Team, R. (2014). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.
- Rickers-Ovsiankina, M. A. (1956). Social accessibility in three age groups. *Psychological Reports*, 2, 283–294.
- Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, 48, 1–36.
- Roster, C. A., Albaum, G., & Smith, S. M. (2014). Topic sensitivity and internet survey design: A cross-cultural/national study. *Journal of Marketing Theory & Practice*, 22, 91–102.
- Scheier, M. F., & Carver, C. S. (1985). The self-consciousness scale: A revised version for use with general populations. *Journal of Applied Social Psychology*, 15, 687–699.
- Sieber, J. E., & Stanley, B. (1988). Ethical and professional dimensions of socially sensitive research. *American Psychologist*, 43, 49–55.
- Smith, H. J., Dinev, T., & Xu, H. (2011). Information privacy research: An interdisciplinary review. *MIS Quarterly*, 35, 989–1016.
- Suchman, D. S. (1965). *A scale for the measurement of revealingness in spoken behavior*. Unpublished master thesis.
- Tourangeau, R., Couper, M., & Steiger, D. (2003). Humanizing self-administered surveys: Experiments on social presence in web and IVR surveys. *Computers in Human Behavior*, 19, 1–24.
- Tourangeau, R., & Smith, T. W. (1996). Asking sensitive questions: The impact of data collection mode, question format, and question context. *Public Opinion Quarterly*, 60, 275–304.
- Tourangeau, R., & Yan, T. (2007). Sensitive questions in surveys. *Psychological Bulletin*, 133, 859–888.
- Uriell, Z. A., & Dudley, C. M. (2009). Sensitive topics: Are there modal differences? *Computers in Human Behavior*, 25, 76–85.
- Vondracek, F. W. (1969). The study of self-disclosure in experimental interviews. *Journal of Psychology*, 72, 55–59.
- Vondracek, S. I., & Vondracek, F. W. (1971). The manipulation and measurement of self-disclosure in preadolescents. *Merrill-Palmer Quarterly of Behavior and Development*, 17, 51–58.
- West, L. W., & Zingle, H. W. (1969). A self-disclosure inventory for adolescents. *Psychological Reports*, 24, 439–445.
- Wheeless, L. R., & Grotz, J. (1976). Conceptualization and measurement of reported self-disclosure. *Human Communication Research*, 2, 338–346. doi:10.1111/j.1468-2958.1976.tb00494.x