Anonymity and Confidentiality: Do Survey Respondents Know the Difference?

Thomas J. Whelan
North Carolina State University

ABSTRACT
Making privacy assurances to participants is one of the requirements of informed consent in psychological research, where participants are told that their information is anonymous and/or confidential. Survey researchers often make distinctions between these two concepts, and many past studies have been designed around the premise that anonymous and confidential (i.e., identified) experimental conditions will yield different outcomes. However, are participants making the same conceptual distinction, or are they interpreting these two concepts as a single construct? This study demonstrated that respondents to a Web-based survey were able to distinguish appropriately between anonymity and confidentiality in different experimental conditions.

INTRODUCTION
One of the most frequent concepts examined in research concerning surveys is the effect of respondent identifiability. This notion has often been separated into two categories: anonymity and confidentiality. Some researchers treat these as two exclusive constructs (e.g., Ong & Weiss, 2000; Scott, 2005), others question whether or not anonymity and confidentiality might be the same thing (e.g., Rogelberg, Spitzmueller, Little, & Reeve, 2006), some researchers intentionally use the two terms interchangably (e.g., Boruch, 1971), and other researchers may talk about one concept while arguably describing the other (e.g., Ransdell, 1996). Despite the relative lack of agreement over these interrelated ideas, ideas about confidentiality and/or anonymity often emerge as a substantial concern of survey respondents (e.g., Cho & LaRose, 1999; Joinson, Woodley, & Reips, 2007; Sills & Song, 2002; Thompson & Surface, 2007; Thompson, Surface, Martin, & Sanders, 2003). From a practical viewpoint, investigating the issues surrounding survey privacy is not without due cause—privacy is quite clearly an instrumental part of proper research ethics (King & Spector, 1963; Scott, 2005).

The majority of research concerning survey privacy has often ignored the subjective perceptions of respondents, instead examining the differences between objectively defined anonymous respondents and confidential (or identified) respondents (e.g., Booth-Kewley et al., 1992; Fox & Schwartz, 2002; Joinson, 1999; Moore & Ames, 2003; Rosenfeld, Booth-Kewley, Edwards, & Thomas, 1996). The underlying logic in this research has been that assignment to conditions will yield different outcomes on variables of interest, and that any observed differences can be attributed to the experimental condition. However, the cumulative results of this type of research have yielded mixed findings (e.g., Singer, von Thurn, & Miller, 1995). In contrast, researchers such as Rogelberg et al. (2006) have called for research to explore whether anonymity and confidentiality are indeed different constructs, and further if respondents to a survey are able to distinguish between the two.
One of the more popular research topics in survey respondent behavior is how individuals react to the content of a given survey and how it might influence their willingness to disclose personal information. Particular attention has been paid to sensitive item content (e.g., Couper, Singer, & Tourangeau, 2003; Joinson, Woodley, & Reips, 2007; Lenert & Skoczen, 2002; Ong & Weiss, 2000; Tourangeau & Yan, 2007). A sensitive survey item can be conceptualized as including three components related to the intrusiveness of the item content, the threat of potential consequences for responding to the item, and the social undesirability of a respondent’s answer to the item (Tourangeau & Yan, 2007).

Researchers have studied the differences in the data quality generated by anonymous and confidential respondents for topics such as self-reports of drug use (Bjarnason & Adalbjarnardottir, 2000; O’Malley, Johnston, Bachman, & Schulenberg, 2000), erotophilia (Durant, Carey, & Schroeder, 2002), and sexual abuse as a child (Olson, Stander, & Merrill, 2004). Ong and Weiss (2000) stated that perceptions of privacy may be the most influential factor in a respondent’s decision to endorse sensitive items on a survey. As such, many researchers have advocated using deceptive methods of identifying individual surveys while leading the respondent to believe that they are anonymous (e.g., Dickson, Casey, Wyckoff, & Wynd, 1977; Manniche & Hayes, 1957). However, what constitutes sensitivity may not be the same for all respondents (Ong & Weiss, 2000). Accordingly, using only a single sensitive item, as done by Joinson et al. (2007), may not provoke privacy concerns in a respondent. As such, the current study utilizes a sizable behavioral inventory related to academic dishonesty comprised of a variety of items, which will be described in detail later, to ensure a breadth of coverage of a sensitive domain.

The aim of this study is not to propose a theoretical framework for the constructs of either anonymity or confidentiality. Instead, and of practical importance to both researchers and practitioners, this study focuses on whether or not survey respondents interpret the concepts of anonymity and confidentiality as synonymous or distinct. As such, instead of simply manipulating survey conditions and hypothesizing differences in an outcome variable, the present study extends the research on survey privacy by directly asking respondents to indicate how anonymous or confidential they feel in addition to varying the presence or absence of identifying information by condition.

Perceived Anonymity

Even though many disciplines such as psychology, computer science, and communications may define anonymity in objective terms, it is ultimately the subjective experience of anonymity that may lead survey respondents or experiment participants to inhibit their behavior or self-disclosure. Anonymity can be conceptualized as “the degree to which the identity of a message source is unknown and unspecified; thus, the less knowledge one has about the source and the harder it is to specify who the source is among possible options, the more anonymity exists” (Scott, 2005, p. 243). The concept of anonymity is relevant to both larger social contexts, such as crowds, or smaller contexts, such as computer-mediated communication between two people (Christopherson, 2007). Anonymity has not received much attention in the extant literature as a factor that may vary between individuals, even though the influence of anonymity has historically been examined repeatedly in research concerning privacy and survey methods (e.g., Ash & Abramson, 1952; Klein, Maher, & Dunnington, 1967; Olson, 1936; Rosen, 1960). Moreover, some authors contend that while anonymity is invoked often in research results, investigations of the full domain of the construct itself are comparatively scant (e.g., Anonymous, 1998).

As noted earlier, researchers frequently design studies around the premise that conditions differ in the anonymity afforded to respondents. However, such a manipulation may only be effective to the extent that the individuals completing a survey believe the anonymity assurances of the researcher and are cognizant of the fact that their identity either is not or will not be tied to other information that they might provide in the course of the research. It has been noted that some respondents simply do not believe the assurances made by researchers (Fox & Schwartz, 2002). Further, there may not be homogeneity across different research studies in respect to the type of assurances made to respondents (e.g., Sobal, 1984), so the effect of anonymity assurances on respondent perceptions may not be the same every time an individual agrees to complete a survey instrument. Accordingly, regardless of the actual, objective anonymity, it is
the perceptions of the individual that matter in respect to how anonymous a person feels (Christopherson, 2007; Pinsonneault & Heppel, 1997). For instance, Joinson et al. (2007) found that different techniques of authenticating a respondent’s identity may lead to both positive and negative outcomes for variables such as response rate and self-disclosure, depending on how salient the identification of the respondent by the researchers was perceived to be. The fact that identity may affect outcomes or respondent behaviors in different ways is not necessarily surprising, as privacy issues can be rather complex in how they are manifested (Scott, 2005).

While anonymity in general has garnered considerable research attention, there has been an increase in the attention paid to survey privacy with Web-based surveys (e.g., Buchanan, Paine, Joinson, & Reips, 2007; Lane, 2003; O’Neil, 2001; Stanton, 1998). Malin and Sweeney (2004) showed that even the use of encrypted data does not necessarily ensure anonymity, due to the data trails left by computer activity. However, a respondent’s anonymity concerns when completing Web-based surveys may or may not be grounded in reality. Most people do not have accurate knowledge of privacy technology and the vulnerabilities of the computer systems running the surveys; at the same time, they may have erroneous beliefs about what are and are not safe practices for online behavior (Graeff & Harmon, 2002; Jensen, Potts, & Jensen, 2005). Therefore, subjective perceptions of anonymity, in spite of assurances of objective anonymity, may have a significant impact on how people answer a Web-based survey.

For the present study, it is expected that respondents will be concerned about the anonymity of their responses to sensitive survey items, and thus when asked, will rate their perceptions of the anonymity of the survey as higher in a condition where they have not given their name compared to a condition in which they have identified themselves.

Hypothesis 1: Respondents in conditions of anonymous and identified survey conditions will have significantly different levels of perceived anonymity.

Perceived Confidentiality

For sensitive and/or personal information, protecting the privacy of respondents by keeping data confidential is of utmost concern to researchers (Ruebhausen & Brim, 1966; Wolf, Zandecki, & Lo, 2004). In practice, this often means that information about individuals is protected from inadvertent disclosure to others by physical means, such as a locked cabinet, legal means, such as signed confidentiality statements, or methodological means, such as the use of coded files (Easter, Davis, & Henderson, 2004). As the Internet continues to facilitate data collection through online survey administration, maintaining the confidentiality of records can become problematic. For instance, information from several sources can be combined and used to possibly identify respondents through online databases with greater ease compared to paper records (Leahey, 2007). It is the responsibility of the researchers collecting data to ensure that steps have been taken to prevent the accidental disclosure of data, and accordingly, researchers are obligated to make assurances to respondents that adequate security measures have been taken (Easter et al., 2004). Yet, confidentiality assurances may matter to respondents only when the content of the survey is considered to be sensitive (Singer et al., 1995). Therefore, there may be instances when an individual completing a survey is not actively concerned about their information being disclosed.

Very little research has explicitly addressed the underlying processes concerning how respondents might interpret the notion of confidentiality. When a survey requires respondents to be identified, the common sentiment is that the practice of identifying a respondent leads to invalid data, as an individual will be less forthcoming if they know their name can be tied to their responses (Corey, 1937). However, despite the fact that researchers such as Scott (2005) have proposed that confidentiality and anonymity are “either/or” concepts, that may not translate into how survey respondents think about the differences between the two ideas. For instance, if a respondent is in an identified survey condition in which he or she has no anonymity, it would logically follow that his or her perceptions of anonymity would be low. But would a respondent in an anonymous survey condition therefore rate low on confidentiality perceptions? What may be more likely is that confidentiality is a necessary (but arguably not sufficient) condition for anonymity. That is, for an individual’s identity to be completely
compromised (i.e., his or her information becomes publicly known), they cannot be anonymous; conversely, an individual’s identity may inadvertently become known, but it may still be held confidential by some given entity or individual and not made public.

**Research Question 1:** Will respondents in conditions of anonymous and identified survey conditions have different levels of perceived confidentiality?

**Evaluation Apprehension**

Already alluded to thus far in this paper is the important role that social evaluation can play in privacy. Anonymity, by definition, implies that there is no possibility of social evaluation, as identification would be required for an evaluation to occur. Evaluation apprehension was originally defined by Rosenberg (1965) as “an active, anxiety-toned concern that [the subject] win a positive evaluation from the experimenter, or at least that [the subject] provide no grounds for a negative one” (p. 29). Implicit in this concept is that the identity of an individual has to be known, at least by the researcher, for an evaluation to be made. The presumed relationship between anonymity and evaluation apprehension is illustrated by a study by Thomas et al. (1979), which operationalized high evaluation apprehension as “informing subjects that they were to be tested nonanonymously; low evaluation apprehension subjects were tested anonymously” (p. 194). With a condition of confidentiality, the two necessary pieces of information for a respondent to be evaluated (namely, identity and some other personal information) are available, so there is the possibility that an individual might anticipate some form of potential social evaluation when completing a survey which has been tied to their name. Thus, an influence of levels of evaluation apprehension may be perceived confidentiality; as an individual perceives a survey as more confidential, he or she will have less concern for social evaluation (Zimbardo, 1969). Thus, it is predicted that evaluation apprehension and perceived confidentiality will be significantly related constructs.

**Hypothesis 2:** Respondents’ evaluation apprehension will be negatively related to perceived confidentiality.

**Method**

**Participants**

Participants for this study were 405 undergraduate students from a large Southeastern university who participated in this study to receive course credit. With regard to gender, 60.7% of the sample was male, and the average age of participants was 19.1 years. There were 14 respondents who were excluded from further analysis due to missing data for one or more variables, for a final sample size of 391 participants.

**Procedure**

Participants in this study volunteered to complete a Web-based survey on academic dishonesty. Although this survey was conducted through a university experiment portal, an attempt to control the experimental surroundings was not made. This was done in an effort to add to the generalizability of the study by extrapolating to the varying respondent environments of field research, where the immediate surroundings of respondents completing Web-based surveys are often unable to be controlled. As such, respondents to the survey were able to access the survey materials and complete them from any location they chose where they had Internet access on a computer. Privacy assurances were made to each individual who participated in this study in informed consent materials, namely that the information in the study records would be kept strictly confidential and responses to the survey would be anonymous. As it is necessary to make privacy assurances to participants in accordance with IRB requirements, the informed consent terminology used in this study included both terms of interest, as it would be impossible to include neither.

After acknowledging they had read online informed consent materials, participants were randomly assigned to one of two experimental conditions via a Javascript algorithm embedded in the html code of the informed consent page. In the anonymous survey condition \( n = 280 \), participants did not indicate their identity prior to beginning the survey.
In the confidential survey condition \((n = 111)\), participants were asked to provide their name and student ID number prior to beginning the survey. Participants then completed a measure of evaluation apprehension, followed by the Academic Dishonesty Inventory (ADI) (Lucas & Friedrich, 2005), which provided sensitive item content for which respondents should feel some concern about survey privacy. The ADI asked participants to admit to incidences of dishonest behavior in school settings and academic misconduct that occurred within the past two years. Next, participants responded to items relating to perceptions of anonymity and confidentiality specific to the present survey. Demographic information was collected for each participant, and the name and student ID of each participant was requested for the purposes of assigning course credit for participation. Lastly, participants were directed to online debriefing materials. It should be noted that in the confidential survey condition, the name and student ID provided by each participant at the beginning of the survey was not actually recorded. Therefore, each respondent was technically anonymous to the researchers regardless of survey condition, so the two conditions in this study differ only in the presence or absence of self-identification.

The measures in this survey, which will be described in the following section, were presented in the order described above so that demographic information and name would be requested last in an attempt to prevent raising respondents’ suspicions about privacy assurances for any systematic reason other the experimental manipulation. It should also be noted that within each section of the survey, the items within each measure were presented in random order where applicable to counteract the potential for item-level order effects (see Feldman & Lynch, 1988).

**Measures**

**Evaluation apprehension** \((\alpha = 0.89)\). Twelve items from Leary (1983) were administered to assess evaluation apprehension. Reverse-coded items were reverse scored and responses to this scale were subsequently averaged. Scale scores ranged between 1 (“not at all characteristic of me”) and 5 (“extremely characteristic of me”), with higher values reflecting greater reported levels of evaluation apprehension. An example item is, “During the training program, I worried about what other people would think of me.”

| Table 1 |
| Means, Standard Deviations, and Intercorrelations |
|---------|-------|-------|-------|-------|-------|-------|
|         | \(M\) | \(SD\) | 1     | 2     | 3     | 4     |
| 1. Condition\(a\) | --    | --    | --    | --    |       |       |
| 2. Perceived Anonymity | 3.84  | 1.02  | -.11* | --    |       |       |
| 3. Perceived Confidentiality | 4.19  | 0.87  | -.08  | .55** | --    |       |
| 4. Evaluation Apprehension | 2.60  | 0.56  | -.01  | -.08  | -.11* | --    |
| 5. Reports of Cheating Behavior | 4.90  | 4.00  | .02   | -.08  | -.10* | .14** |

\* \(p < .05\), \** \(p < .01\)

\(a\) 1 = anonymous condition, 2 = identified condition
Anonymity and confidentiality perceptions. Two items were administered to measure perceptions of anonymity and confidentiality specific to the survey that respondents had just completed. Participants responded to two items that asked them to rate on a 5-point Likert-style scale (1 = strongly disagree, 5 = strongly agree) their certainty they that the survey was anonymous or confidential. The items are: “I feel certain that this survey is anonymous,” and, “I feel certain that this survey is confidential.”

Academic dishonesty. Academic dishonesty was measured using the ADI developed by Lucas and Friedrich (2005). Participants responded to 26 items that asked them to indicate “yes” or “no” in regard to whether or not they had engaged in academic misconduct at any time in the past two years. The amount of “yes” responses was subsequently summed to form a scale score. The highest scale score for the ADI in this sample was 20, and 12.5% of participants did not indicate a “yes” response for any of the items. Example items are, “Allowed own coursework to be copied by another student,” and, “Fabricated reference or a bibliography.”

Demographic variables. All participants responded to items requesting information about their gender and age.

Table 2

Results of t-tests for Survey Conditions

<table>
<thead>
<tr>
<th>Variable</th>
<th>Condition</th>
<th>N</th>
<th>M</th>
<th>SD</th>
<th>t-statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Perceived Anonymity</td>
<td>Anonymous</td>
<td>280</td>
<td>3.91</td>
<td>0.97</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Identified</td>
<td>111</td>
<td>3.67</td>
<td>1.13</td>
<td></td>
</tr>
<tr>
<td>Perceived Confidentiality</td>
<td>Anonymous</td>
<td>280</td>
<td>4.24</td>
<td>0.85</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Identified</td>
<td>111</td>
<td>4.08</td>
<td>0.93</td>
<td></td>
</tr>
<tr>
<td>Evaluation Apprehension</td>
<td>Anonymous</td>
<td>280</td>
<td>2.59</td>
<td>0.58</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Identified</td>
<td>111</td>
<td>2.58</td>
<td>0.51</td>
<td></td>
</tr>
<tr>
<td>Reports of Cheating Behavior</td>
<td>Anonymous</td>
<td>280</td>
<td>4.85</td>
<td>3.94</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Identified</td>
<td>111</td>
<td>5.04</td>
<td>4.18</td>
<td></td>
</tr>
</tbody>
</table>

* p < .05
be part of a continuum with anonymity, but may be a separate construct, in that participants will differentiate between confidentiality and anonymity.

The correlations in Table 1 showed that confidentiality perceptions were significantly and negatively related to evaluation apprehension. Thus, Hypothesis 2 was supported. Anonymity perceptions, while negatively correlated, were not significantly related to evaluation apprehension. This again suggests that anonymity and confidentiality are being perceived as distinct concepts by survey respondents, as they are highly correlated with each other yet are not both significantly correlated to evaluation apprehension.

Of interest, though not hypothesized, are the results of t-tests between conditions for evaluation apprehension and reports of academic dishonesty. While there were no differences between conditions for either of these variables, they were significantly correlated with each other and both were significantly correlated with confidentiality perceptions. As an exploratory analysis, a hierarchical regression was conducted to examine which, if any, of the variables in this study were predictors of reported academic dishonesty. All continuous predictor variables were mean-centered prior to analysis to reduce the effects of multicollinearity for the regression model (Aiken & West, 1991). For the first block of predictors, anonymity and confidentiality perceptions were used to predict reports of academic dishonesty while controlling for assignment to survey conditions. The regression model was not significant for the first block ($F[3, 387] = 1.44, p > .10, R^2 = 0.01$). Next, evaluation apprehension was added to the regression equation. The regression model was significant ($F[4, 386] = 2.66, p < .05, R^2 = 0.03$). The addition of evaluation made the full regression model significant ($\Delta R^2 = 0.02, F[1, 386] = 6.29, p < .05$). Table 3 contains a summary of the regression parameter estimates for the full model.

**Discussion**

This study demonstrates that respondents to a Web-based survey were able to distinguish appropriately between anonymity and confidentiality, two highly related concepts ($r = .55, p < .01$), in different experimental conditions. Moreover, the fact that there was variance on the items concerning anonymity and confidentiality shows that these ideas may present a relatively ignored confound to survey research, as not all individuals reported experiencing conditions of anonymous or confidential conditions to exactly the same degree. Further, evaluation apprehension was shown to be significantly and negatively related to confidentiality perceptions, but not significantly related to anonymity perceptions. This provides further evidence that when respondents are completing a survey, not only do they make a distinction between anonymity and confidentiality, but their ratings of either are not related in the same magnitude to another theoretically related construct. In addition, there was no significant difference between objective survey conditions for evaluation apprehension. This would suggest that it may not be the survey condition itself that influences a survey respondent’s concern for social evaluation, but their perception of how confidential their responses to the survey may be. This is of particular importance since confidentiality perceptions also did not significantly differ across conditions, which raises the question of whether or not evaluation apprehension can result in response distortion regardless of anonymity or confidentiality assurances made to respondents. Indeed, it would appear from the results of the exploratory regression analysis that evaluation apprehension does predict the amount of disclosure made by respondents. Therefore, it can be concluded that the processes underlying respondents’ reaction to sensitive items may not rely on assignment to anonymous or confidential survey conditions themselves, but on the extent to which

---

Table 3

*Regression Model Parameter Estimates*<sup>a</sup>

<table>
<thead>
<tr>
<th></th>
<th>β</th>
<th>t value</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Condition</td>
<td>0.01</td>
<td>0.27</td>
<td>0.79</td>
</tr>
<tr>
<td>2. Perceived Anonymity</td>
<td>-0.02</td>
<td>-0.40</td>
<td>0.69</td>
</tr>
<tr>
<td>3. Perceived Confidentiality</td>
<td>-0.07</td>
<td>-1.21</td>
<td>0.23</td>
</tr>
<tr>
<td>4. Evaluation Apprehension</td>
<td>0.13</td>
<td>2.51</td>
<td>0.01</td>
</tr>
</tbody>
</table>

<sup>a</sup> The dependent variable for the regression analysis was reports of academic dishonesty
low perceptions of confidentiality and higher levels of evaluation apprehension might contribute to an unwillingness to endorse socially undesirable behaviors.

The finding that there was no significant difference between conditions for reports of academic dishonesty, while not preferable given the demonstration of differences in self-reports of undesirable behavior in other conceptually similar studies such as Joinson et al. (2007) and Ong and Weiss (2000), is not inconsistent with the extant literature on survey privacy (e.g., Singer et al., 1995). Still, this study contributes to the literature on survey privacy and respondent behavior by demonstrating that even in the absence of a difference between conditions for reports of sensitive content, there are other factors that do differ and may influence how a respondent interprets and subsequently approaches a survey instrument.

Limitations and Future Research

Despite the contributions of this study to the understanding of survey respondents’ perceptions of privacy, it has several limitations that should be noted. First, the use of undergraduate students as a research sample may be suboptimal. In an applied setting, finding an effect for the perceptions of a survey respondent might be dependent on other variables not typically present in university research, such as the use of computer monitoring by an employer. Second, the study did not control for the physical environment of the participants. While this was an intentional limitation for reasons explained earlier, and arguably is indirectly accounted for by a portion of the variables that were measured in this study, it should be pointed out that this study would benefit from a replication using stronger experimental controls.

Future research can take into account the anonymity and confidentiality perceptions of participants to test the effects of intentional and unintentional differences in experimental conditions where privacy may be a concern of survey respondents. For instance, exploring the potential for anonymity or confidentiality in predicting outcomes could provide insight into other contextual variables in Web-based survey research that heretofore have been considered confounds. The boundaries of privacy perceptions as an explanatory variable should be explored to see where they may or may not serve to account for varying interpretations of what is and is not sensitive survey content. As noted by researchers such as Ong and Weiss (2000), the relative sensitivity of items can be subjective, so measuring perceived anonymity and confidentiality may be able to assist researchers in identifying when reports of behavior are low due to privacy concerns.

More research is also needed to properly define the constructs of anonymity and confidentiality through a subjective framework to determine what survey respondents see as different between the two constructs. Lastly, research should explore what types of threats to privacy or releasing identity affect anonymity or confidentiality perceptions, and how these threats may lead to differences in outcomes depending on the salience of a privacy threat.

References


*Direct all correspondence regarding this paper to Tom Whelan at tjwhelan@ncsu.edu*