Introduction

Research relying on qualitative research of various types has become widespread in marketing research, particularly in consumer behavior studies, for qualitative methods provide researchers with insight into phenomena at a depth that cannot be achieved with quantitative data. Notable studies utilizing qualitative techniques include Arnould and Price (1993), Clesi et al. (1993), Gilly and Wolfenbarger (1998), Hirschman (1992), and Thompson (1997). In addition to purely qualitative methods, pluralistic research, where qualitative methods are used in conjunction with quantitative methods, has gained in popularity due in large part to the admonitions of scholars such as Deshpande (1983) and Zinkhan and Hirschheim (1992). Examples of pluralistic research include Price et al. (1995), Mick and Fournier (1998), and Sirri et al. (1996). In short, qualitative research is indispensable whenever a marketing researcher investigates a domain that is unknown or has received relatively little attention to date.

It is self-evident that the makeup of a researcher’s data determines the character of his/her findings. With quantitative data, there are established procedures such as exploratory factor analysis, Cronbach’s alpha, confirmatory factor analysis, and statistical tests that collectively identify the data’s quality level. However, with qualitative data, quality is not amenable to such techniques, and, while judgemental quality checks are used, there may be opportunity for subtle patterns to exist in the data and possibly influence the researcher’s interpretations.

This paper describes and tests how qualitative data may harbor subtle biases depending on factors such as the gender of the informant, the topic, or even the cover story used by the researcher. Due to the disparate qualitative inquiry techniques available, the paper focuses on a single technique, namely critical incidents technique (CIT), as this method is attractive to academic marketing researchers. The paper begins by describing the CIT elicitation procedure and identifying cases where this method has been used as a singular marketing research qualitative technique as well as where it has been used in pluralistic research.

To substantiate our claims of tacit data quality errors, we introduce self-disclosure
theory, and we refer to research findings documenting systematic differences in subjects' self-disclosures. Given these, we develop hypotheses on how qualitative data generated with CIT may be affected by various factors, and we describe the findings of an experiment designed to test these hypotheses. The paper concludes with recommendations for a researcher who wishes to gather richer narrative texts from informants while using CIT.

Critical incident technique

Flanagan (1954) originated CIT as a means of identifying a significant, or critical, behavior or factor that contributes to the success or failure of some human event. Early uses of CIT involved military training such as selecting pilot trainees, bombing procedures, and combat leadership (Flanagan, 1954). Since that time, marketing researchers such as Bitner and her colleagues (Bitner et al., 1990; Bitner et al., 1994; Mohr and Bitner, 1995) have adapted CIT to identify critical aspects of service encounters. Similarly, Grove and Fisk (1997) used CIT to identify events that detract or augment customers' service experiences. At the same time, Kelley et al. (1993) used CIT to develop a typology of retail failures and recovery strategies.

In its marketing and consumer behavior applications, the earmarks of CIT are:
(1) focus on a particular event or phenomenon;
(2) elicitation of informants' detailed narrative text records of their experiences with the phenomenon;
(3) content analysis of these narrative texts; and
(4) thematic interpretation based on the content analysis.

For example, researchers have used CIT to gain an understanding of specific consumer behavior situations. Mick and DeMoss (1990a, 1990b), for instance, used CIT to investigate self-gifts, and Rook (1987) used this technique to study impulse purchases. Keaveney (1995) used CIT to understand why customers switch service providers, and Gabbott and Hogg (1996) used CIT to gain insight into consumers' perceptions of service quality in the healthcare industry. Finally,

CIT is sometimes used as a foundation for subsequent quantitative research as for example O’Guinn and Faber (1989) who used CIT as part of their investigation of compulsive buying.

Whatever its application, CIT is an appealing qualitative method for marketing researches for good reasons. For one, it is easily adapted to research seeking to understand marketing experiences encountered or recalled by informants. Second, there are examples of its use in both purely qualitative research and pluralist marketing studies, and thus there is precedent for its use. Third, it can be administered quickly and easily to large groups of informants. Finally, with the use of computerized text analysis programs, large amounts of narrative text generated by CIT can be analyzed expeditiously.

Self-disclosure

An informant's narrative text can be considered to be self-disclosure, defined as "any information exchange that refers to the self, including personal states, dispositions, events in the past, and plans for the future" (Derlega and Grzelak, 1979, p. 152). A diverse set of researchers is interested in self-disclosure; a considerable body of empirical research exists, and much information has been collected regarding the factors that affect the character of a person's self-disclosures. However, no research to date has sought to verify that these systematic data quality differences exist in marketing research studies utilizing CIT. Three areas that are particularly relevant to marketing researchers are informant factors, target characteristics and topic of study.

Informant factors

Informant factors are characteristics of the informer associated with variations in the amount of self-disclosure. Informant gender is one such characteristic. Research suggests that women divulge more about themselves than do men (Chelune, 1976; Alloy et al., 1985). This gender difference has been observed in a large number of self-disclosure studies beginning with early discoveries by Jourard and Lasakow (1958), Jourard (1961), and Jourard and Richman (1963), and it has been further documented in reviews (Cozby,
1973) and meta-analysis (Dindia and Allen, 1992). This tendency in women is explained by sex role socialization, wherein women are taught to be expressive and men are expected to be nonexpressive (Jourard, 1971b). A variation is masculinity-femininity where males are held to be instrumental and task-oriented while females are reactive and communal (France et al., 1984); that is, females interact and provide more information than do males.

Thus, the first hypothesis of our study is stated as follows:

**H₁:** In a marketing study utilizing CIT, female informants will disclose more information about themselves than will male informants disclose about themselves.

**Target characteristics**

**Target familiarity**

Target characteristics, or characteristics of the information recipient, affect self-disclosure. More self-disclosure is provided to targets that are similar to the informant than to targets who are dissimilar. Both social exchange theory (Homans, 1950) and similarity theory (Byrne, 1971) explain the target factor differences in self-disclosure. Social exchange theory, sometimes termed the "dyadic effect", (Jourard, 1971a) and reciprocity (Bloch and Goldstein, 1971) hold that more self-disclosure occurs between friends than it does between strangers.

Friends sharing common interests or beliefs exchange information about themselves as a means of building intimacy, also, it is believed the history of the relationship tends to solidify these similarities. At the same time, individuals limit the exchange of information about themselves with strangers as a means of maintaining social distance (Cline and Musolf, 1985). In fact, some researchers hold that people in general subscribe to social norms discouraging self-disclosure to strangers (Gerdes et al., 1981). Again, much research supports this contention (see for example Hansen and Schultz, 1982; Hatch and Leighton, 1986; or Gitter and Black, 1976), leading to our second hypothesis.

**H₂:** In a marketing study utilizing CIT, informants will disclose more information about themselves to a person with whom they are familiar than they will disclose to a stranger.

**Target similarity**

In attempts to further explain the findings of the differences in self-disclosure across targets, other research has investigated instances where the operationalization of similarity has been same-sex pairs, as opposed to different-sex pairs, regardless of the relationship in place. Jourard (1971a), for instance, found that more disclosure occurred with same-sex friends than with different-sex friends, and Hatch and Leighton (1986) also report this finding. Cline and Musolf (1985) cite studies where females divulge more to female targets than to male targets, and Burnard and Morrison (1994), in a replication of an early Jourard study, found that a same-sex friend garnered the greatest number of disclosures by female nursing students. From slightly different contexts, Balswick and Balkwell (1977) found that female adolescents disclosed more to mothers while male adolescents disclosed more to fathers, and Lord and Velicer (1975) found more disclosure between same-sex siblings, but no differences for same-sex friends. These findings lead to our third hypothesis:

**H₃:** In a marketing study utilizing CIT, informants will disclose more information about themselves to a person of the same sex than they will disclose to a different-sex person.

**Disclosure topic**

Finally, the sensitivity and interest of the respondent for the topic of discussion has been held to also affect the quality of information given by respondents. Some topics elicit more self-disclosure than do other topics. Topics that are personal or intimate tend to elicit less disclosure than do topics that are public or superficial (Gitter and Black, 1976). However, topic-specific disclosure has received little systematic self-disclosure research attention, and Dindia and Allen (1992) were unable to analyze its effects in their meta-analysis for this reason.

Informant and target considerations aside, it seems logical to expect that topics for which informants have high involvement should elicit more disclosure than should low involvement topics. Specifically, high involvement products and services are those where the consumer has more interest, information, and experience than with low involvement products and services. At the same time, high involvement situations are
more complex, more complicated, and typically have more dimensions than do low involvement situations. Indeed, several items in the Zaichkowsky (1985) involvement scale (important, means a lot to me, interested, relevant, etc.) strongly suggest that high involvement objects are more substantive than are low involvement objects. Thus, high involvement service and product experiences are typically richer, deeper, and longer than are low involvement ones; that is, high involvement topics are more information-laden than are low involvement topics. This logic leads to the last hypothesis.

H4: In a marketing study utilizing CIT, informants will disclose more about a high involvement topic than they will disclose about a low involvement topic.

Method

Identification of relevant consumption situations and sample
To identify appropriate consumption situations for the experimental treatments, college students were asked to list challenging spending-related situations. From these responses, it was determined that a high involvement topic was the use of credit cards. Credit card usage had the highest mention, and the financial consequences of credit card debt or abuse of credit cards (such as exceeding the allowable limits) are daunting for college students who live on tight budgets. On the other hand, eating fast food was identified as a low involvement consumption situation, as it had low mention incidence with respect to being a significant change in college students' lifestyles, but it was associated with students' health (diet) concerns.

Further justification for these two topics is suggested by their practical significance. Use of credit cards by college students is a multimillion-dollar business, plus it is students' first encounter with credit management. Reports indicate that perhaps as many as 85 percent of all college students own credit cards (Murro and Hirt, 1998), and studies have documented college student credit card balances in the thousands of dollars and students holding multiple credit cards (Souccar, 1998). Fast food consumption by young people is also a multimillion-dollar industry and diet, health, and weight are salient concerns of college students (US Department of Health and Human Services, 1989). College students are especially vulnerable to poor nutrition as they are removed from parental control (Baker, 1991). These preliminary research phases justified research on these topics as both are managerially significant, but neither is well understood. That is, both are suitable to CIT qualitative investigation by marketing researchers.

Research design and administration
The formal research design was a 2 x 2 x 2 x 2 repeated measures design with random assignment of subjects. The experimental variables were:

1. Informant gender (male versus female),
2. Target/relationship type (friend versus researcher),
3. Informant-target gender match (same gender versus different gender), and
4. Topic involvement (high versus low).

Topic was the repeated measure. The experimental procedure was administered to college students in a group setting.

For the topic treatment, two CIT situations were prepared. With the credit card, informants were placed in a scenario event where while shopping with friends they encounter a good bargain in a department store, and they are unsure how close their credit card balance is to its limit, but they use the credit card to buy the item. For the fast food scenario, the event concerns noontime with friends when everyone is feeling hungry, and they end up eating lunch at Burger King even though not everyone agrees on going there. The situations were ordered randomly within informants. To accomplish the relationship type and informant-target gender match treatments, respondents were told that the person to whom they were describing their thoughts/feelings was either a best male/female friend or a researcher named Joseph/Josephine Bradley. Bradley was described as someone working for a respected research organization conducting confidential interviews with individuals selected at random.

To encourage those receiving the best friend treatment to visualize a conversation with that friend, respondents were asked to write the friend's name and to record at least five things that made this person a good
friend. For the researcher treatment, subjects were instructed to write five things about Bradley without looking back at the description of Bradley. Cover sheet instructions directed informants to treat the described event as though it actually happened to them, to act as though they were actually talking to the person who is asking, and to mentally picture the person asking for the information. To ensure that the treatment manipulations correctly impacted the respondents (Perdue and Summers, 1986), these incidents were pretested with a convenience sample of students to verify that the situations were reasonable and realistic.

Both the friend and the researcher requested that the informant write down “... how you felt when you last used your credit card/ate at a fast food restaurant. What kind of feelings and emotions did you experience...?” Informants were provided a blank page to write their descriptions. The next page contained a follow-up question asking the informant to “... tell me what negative consequences you have experienced because of your use of credit card(s)/fast food(s) in general.” Other questions asked about demographics, credit card ownership, fast food use, and two sets of Likert-scale items relating to the fast food and credit card descriptions were included for manipulation checks and confirmation purposes. A total of 249 informants took part in the study.

**Measurement of self-disclosure**

Analysis of the narrative texts provided by the informants was accomplished through the use of a qualitative analysis computer program called Atlas.ti (see http://www.atlasti.de/index.html). This program greatly facilitates the coding and counting functions essential to the analysis of qualitative data (Spiggle, 1994). Atlas.ti was used to determine counts of the number of mentions of words or phrases pertaining to each of several categories established by the researchers. It is a quick and efficient way of performing qualitative text analysis, and its use also eliminates concern for interjudge reliability issues (see Ronan and Latham, 1974; or Kolbe and Burnett, 1991) as Atlas.ti achieves objective measurement of narrative text that is perfectly reproducible.

Four measures of self-disclosure were used. These measures and associated descriptive statistics are discussed in the following sections.

**Total words**

A gross measure of self-disclosure is a simple word count of each informant’s text by topic. For the credit card situation, respondents used an average of 102.3 words in their descriptions. Similarly, respondents averaged 95.8 words to describe their fast food situation.

**Self-references**

A second measure of self-disclosure is the number of self-references, where it is assumed that greater use of self-references is indicative of more disclosure on the part of the informant. Atlas.ti counted all instances of references to the self (I, me, my, etc.) supplied by each informant. The mean (standard deviation) for credit card was 11.1 (5.9), while for fast food it was found to be 8.6 (5.4).

**Emotions: guilt and total emotional disclosures**

A revealing type of self-disclosure concerns sentiments or emotions. Through an iterative method, informants’ thoughts and words were classified into eight mutually exclusive categories (guilt, pleasant feelings, frustration, anxiety, need, self-discipline, reward, and recklessness). As only guilt disclosures were found for both the fast food and the credit card topics, total guilt self-references constituted the third measure of self-disclosure. The fourth measure of self-disclosure was complicated by the fact that the remaining seven emotional self-disclosure types each appeared for only one topic. Consequently, a total emotional disclosures measure was created by summing all eight emotional disclosure types. Guilt disclosures and emotional disclosures were then treated as correlated dependent measures (correlation of 0.63, p<0.001) to take into consideration their shared variance. In fact, all dependent constructs were analyzed as correlated variables.

**Data transformation**

Inspection of the distributions of the various types of disclosure revealed that most failed normality. Consequently, square root transformations were applied to bring the skewness statistics to acceptable levels (Hair et al., 1995, p. 70). While subsequent analyses were performed on the transformed data,
means are reported in non-transformed units to facilitate interpretation.

Findings

Informant appropriateness, manipulation check, and hypothesis guessing

Only those informants who owned at least one major credit card (VISA, MasterCard, etc.) were included in the analysis. Approximately 46 percent owned two or more major credit cards. Almost three-quarters of the informants indicated that they ate at a fast food restaurant at least one time per week, and slightly more than one-quarter indicated that they ate fast food four or more times per week. Thus, all informants were qualified to comment on their experiences with these purchase situations.

A manipulation check was performed with the item, “The person I was writing for was someone I know.” With the friend condition, any informants who indicated disagreement were excluded, while with the researcher condition, informants who indicated agreement were excluded from subsequent analysis. Hypothesis guessing was tested, and no informant indicated a guess about any facet of the experiment.

Hypothesis tests

To test the first three hypotheses, the self-disclosure measures were treated as correlated dependent variables and subjected to MANOVA with informant gender, target, and informant-target gender match as independent variables.

Hypothesis 1: informant gender

Informant gender was found to have a significant main effect (F-ratio = 4.850; p ≤ 0.001, effect size = 0.08). Female informants divulged more than did male informants across all measures of self-disclosure (words: females = 218.0, males = 179.0, self-references: females = 21.2, males = 18.2; guilt: females = 0.94, males = 0.64; and emotions: females = 2.2, males = 1.8). Thus, H1 was supported.

Hypothesis 2: relationship type

Similarly, the target was found to have a significant main effect (F-ratio = 2.447; p ≤ 0.047, effect size = 0.04). Informants used more total words and more self-references when talking to a friend than to a researcher (words: friend = 207.7, stranger = 187.3; self-references: friend = 21.2, stranger = 17.9). No significant differences were found for disclosure of guilt or emotions to a friend compared to a researcher. Thus, H2 was partially supported.

Hypothesis 3: gender match

No significant main effect was found for gender match (F-ratio = 1.969; p ≤ 0.10). Further inspection of the univariate tests, however, revealed two significant differences. For self-references and emotions, the different gender condition yielded greater self-disclosure than did the same gender condition (self-references: same-sex pair = 18.5, different-sex pair = 20.9; emotions: same-sex pair = 1.8, different-sex pair = 2.2). This finding is a reversal of the prediction of H3.

A significant interaction for target x gender match (F-ratio = 2.501; p ≤ 0.043) was found. Inspection of the means revealed that different-sex friends received more guilt (1.0 versus 0.6) and emotional disclosures (2.6 versus 1.6) than did same-sex friends, while there were no differences in guilt and emotion disclosures for strangers regardless of the sex match. This finding further contradicts the similarity hypothesis (H3).

Hypothesis 4: topic involvement

A repeated measures MANOVA was required here. It was found that the product type significantly impacted the self-disclosure patterns (F-ratio = 23.428, p ≤ 0.000, effect size = 0.28), and inspection revealed the credit card situation yielded more words (102.3 versus 95.8), self-references (11.1 versus 8.6), guilt disclosures (0.51 versus 0.27), and emotions (1.4 versus 0.6). Thus, H4 was supported.

Confirmation of findings

As noted in the method description, informants were also asked to answer five-point Likert-scaled questions concerning the nature of their disclosures for each product. That is, the informants provided self-report information on their self-disclosures. These items included statements such as “The fast food descriptions I wrote were accurate,” and “I told private things about myself.” These 12 items were subjected to principal components factor analysis with varimax rotation resulting in four dimensions of self-reported self-
disclosure. The factors were named, and their respective items were analyzed for Cronbach’s alpha reliability with the following results: easy to tell this person (four items, alpha 0.89), felt uncomfortable describing (four items, alpha 0.69), told everything I could (two items, 0.71), and told private things about myself (two items, 0.75).

Admittedly, these are not rigorously developed scales. However, the factor analysis eliminated redundancy, and the alpha levels suggest a modest degree of internal consistency.

Sum scores were created for each self-report self-disclosure dimension, and MANOVA was applied as statistically significant correlations were observed among these measures. Significant main effects were found for the informant’s gender and the target of the disclosure (F-ratio = 4.280; p ≤ 0.002 effect size = 0.07, and F-ratio = 97.483; p ≤ 0.000 effect size = 0.62, respectively). Inspection of the means determined that female informants agreed more with statements to the effect that they told everything about themselves and they told private things about themselves. With respect to the target, friends garnered more self-report agreement that informants told everything about themselves and that it was easy to tell this person about themselves (effect size = 0.59). The gender match hypothesis garnered no support in the confirmation tests (F-ratio = 1.189, p ≤ 0.316), and interactions were not found.

The test for topic differences was more problematic as it required a repeated measures approach and the separation of items into ones that pertained to fast food and those that pertained to credit card usage. Using paired sample t-tests, a statistically significant difference was found with “told private things” (t = -2.492, df = 248, p ≤ 0.013) with more reported self-disclosure for credit card (3.24) than for fast food (3.07).

The self-report measures and the qualitative self-disclosure measures were not identical, as one is perceptual while the other is behavioral. But, the self-report measure results are a form of convergent validity test of the qualitative analysis results. As noted in the above paragraph, the self-report measures confirmed the gender and target hypotheses, failed to support the gender-match hypothesis, and provided some confirmation for the topic involvement hypothesis.

Discussion

Three factors, each with a sound theoretical basis and much empirical support from non-marketing studies, are claimed to affect an individual’s self-disclosure. Specifically, women are socialized to be more expressive than are men, and self-disclosure study findings generally verify that females divulge more than males. At the same time, reciprocity operates in self-disclosures, and typically more intimate self-disclosures are imparted to friends than to strangers. Additionally, there is some evidence that informant-target similarity brings about more disclosure, and same-gender self-disclosures are sometimes found to be more revealing than are different-gender self-disclosures.

Finally, a fourth factor from consumer behavior notions holds that high involvement product experiences should generate more self-disclosure, because they are information-laden than are experiences with low involvement products. While topic sensitivity differences have been found in self-disclosure research, involvement is a novel explanatory construct contributed by the present research.

An experiment substantiated that gender and target effects exist in the narrative text obtained by the use of CIT with marketing topics. Follow-up self-report measures confirmed these findings. Female informants divulged more than males, and disclosures to friends were greater than were those to a stranger. In addition, a high involvement topic generated more self-disclosure than did a low involvement topic, and there was some self-report confirmation of this finding.

However, the gender-match hypothesis was not supported. In fact, some evidence was found contrary to the gender-match hypothesis in that different-gender pairs generated more words and more guilt admissions than did same-gender pairs.

Taken at face value, these findings suggest recommendations to researchers who wish to utilize CIT to generate narrative text as a means of understanding marketing phenomena. These researchers should take into consideration the three determinants of narrative text quality:

1. Informant factors;
(2) target factors; and
(3) topic factors (in this order, based on effect sizes).

First, to the extent possible, researchers in this qualitative phase of their research programs should take measures to insure that the informant group has a balanced representation of females-to-males as women are in fact more expressive than are men. Unless there is some compelling reason to do so, using males exclusively or using a disproportionate number of males may decrease the richness of the narrative text generated by these informants. Next, if it is in the power of the researcher, it will be advantageous to couch the CIT setting such that the informant is disclosing inner thoughts to a friend rather than to a stranger or some ambiguous researcher as more self-disclosure is accorded to a friend than to a stranger. The gender-match aspect of the CIT elicitation context remains problematic: it may or may not have an impact on the quality of the narrative text issued by informants. This target factor requires more research in order to understand its nuances; consequently, as a conservative approach, it is perhaps judicious to distribute target gender-match and mismatch evenly across the informant pool. Finally, since high involvement topics generate more disclosure than do low involvement topics, it is in the researcher’s best interest to select high involvement products or to otherwise use CIT situations or occasions concerning the topic that somehow appropriately tap into an informant’s high involvement circumstances.

References


