Critical Incident Technique in the Basket

Completed Research Paper

Janis Gogan
Bentley University
175 Forest Street, Waltham, MA, USA
jgogan@bentley.edu

Mark-David McLaughlin
Bentley University/Cisco Systems
175 Forest Street, Waltham, MA, USA
mclaugh_mark@bentley.edu

Dominic Thomas
Suffolk University
8 Ashburton Place, Boston, MA, USA
dthomas@suffolk.edu

Abstract

Different IS research methods focus on or measure different aspects of IS phenomena, and studies based on various methodologies thus contribute to the accumulation of strong evidence to support IS theories. We conducted a systematic and reproducible review of ten years of Critical Incident Technique (“CIT”) papers published in the IS Scholars Basket of eight journals. CIT is used extensively in industrial and organizational psychology, marketing and other business disciplines. In IS Basket journals 13 papers fully or partially applied CIT as a research method. We classified these studies per three key elements - incident elicitation, selection, and analysis—and seven criteria employed in other disciplines’ CIT studies. We found that few studies used “full” CIT and we conclude that CIT is underutilized as a research method in IS. Our paper contributes by offering CIT guidelines for authors and reviewers and suggestions for using CIT to build stronger theory in several IS sub-disciplines.

Keywords: IS research, research agenda, research methodologies, critical incident technique

Introduction

Information systems researchers need better and more varied methods in order to conduct studies that can build strong theory regarding how actors successfully and unsuccessfully use systems. In particular, studies that closely examine user behavior in specific contexts and under specific conditions can lead to stronger theory to guide effective systems design and implementation (Burton-Jones and Straub 2006). The Critical Incident Technique (CIT), which offers a robust method for closely observing key dynamics in the behaviors of systems actors, originated in industrial and organizational psychology and has been used in many marketing studies. However, as we argue here, CIT has been little-used in recent IS research despite extensive use in other disciplines, and IS researchers need guidance on its effective application.

Many (perhaps most) IS research methods originated in other business disciplines (marketing, strategy, organizational behavior, etc.) or in non-business disciplines such as “technology, psychology, economics, sociology, mathematics, linguistics, semiotics...” (Mingers 2001). Each method “has advantages and disadvantages; no strategy is more appropriate than all others for all research purposes” (Benbasat et al. 1987); see also (Rosemann and Vessey 2008). While our field benefits from having multiple methods in our toolkits, methodological diversity does give rise to two challenges: how to appropriately use borrowed methods (Goodhue 1998; Sidorova et al. 2008), and how to evaluate studies that used less-known methods. Although influential scholars support methodological diversity in IS research (Bryant 2008; Robey 1996), few papers have discussed how to effectively use and evaluate specific research methods (one excellent paper in this small stream is (Gefen et al. 2000). We address this need by first presenting evidence that CIT has been only lightly used in IS research and then presenting guidance on its effective use and evaluation.
We chose to focus on CIT for three reasons:

1. CIT’s ability to capture generalizable dynamics of actor behaviors fills a critical need in IS research for stronger theories about effective/ineffective systems usage (Burton-Jones and Straub Jr 2006; DeSanctis 2006) and participation (Markus and Mao 2004). Two recent papers (Cenfetelli and Schwarz 2011; Schultz and Orlikowski 2010) called for use of CIT in future IS studies.

2. In industrial and organizational psychology (a contributing discipline to IS), CIT is claimed to be the most important research method (Fivars and Fitzpatrick 2001). CIT has also been extensively used in marketing, another important reference discipline (Bernroider et al. 2013).

3. In our own reading, we noticed CIT was mentioned in some IS papers, but seemed to be used in varied ways, and some studies seemed to apply CIT differently from what is considered best practice for CIT studies in reference disciplines like organizational psychology and marketing.

Several prior reviews described how CIT studies are carried out and/or evaluated in marketing (Bitner et al. 1985; Butterfield et al. 2005; Gremler 2004; Schurr 2007). To the best of our knowledge, no review has described how CIT is used in IS, offered specific guidance on how to conduct or evaluate CIT studies in IS. We contribute to the IS literature by addressing these needs.

Our paper is structured as follows. First, we briefly describe how and why Flanagan (1954) designed and used CIT, and how it was adapted and extended in subsequent management and marketing studies. Drawing on that body of work, we list criteria that define “full” CIT (fully consistent with CIT principles and core elements as originally proposed by Flanagan and other CIT researchers). We define “partial” CIT as an adaptation of CIT which does not conform to all of Flanagan’s original (“full” CIT) elements and principles. We offer evidence that CIT is little-used in current mainstream IS publications by reporting on a systematic, reproducible literature review of recent IS studies published in the top journals in IS (the “IS Senior Scholars Basket” of eight journals) over a ten-year period. We took a close, systematic look at IS papers which claimed to use CIT—whether fully or partially, and whether as the sole research method or one of several methods in a mixed methods study. We discuss our findings, offer suggestions for the use of CIT in future IS studies and in training doctoral students, and conclude with a discussion of avenues for investigating the use of other lightly-exploited but potentially promising methods in future IS research.

CIT: Origins and Typical Use

CIT is a method pioneered by John C. Flanagan “for gathering certain important facts concerning behavior in defined situations” (Flanagan 1954). Critical Incident (CI) reports, properly analyzed, reveal useful information about individual behavior in specific task contexts (Andersson and Nilsson 1964). CIT captures behavioral observations either through direct observation or through eliciting and analyzing informants’ detailed stories about their experiences. Industrial and organizational psychologists recognize CIT as a premier method used in thousands of studies (Butterfield et al. 2005; Fivars and Fitzpatrick 2001). CIT has also been used extensively in marketing studies investigating causes of customer satisfaction and dissatisfaction, especially in service encounters (Bitner et al. 1990; Bogen et al. 2008; Edvardsson and Roos 2001). While CIT is most often used in multi-site studies, it can also be used within single case studies and cross-case comparison studies (Chell 2004).

Flanagan (1954) originated CIT when the US military hired him in the 1940s to conduct studies to reveal key skill requirements for pilots. He defined CIT as follows (Flanagan 1954):

“The critical incident technique consists of a set of procedures for collecting direct observations of human behavior ... [CIT] outlines procedures for collecting observed incidents having special significance and meeting systematically defined criteria.”

Flanagan (1954) defined an incident as:

“any observable human activity that is sufficiently complete in itself to permit inferences and predictions to be made about the person performing the act. To be critical the incident must occur in a situation where the purpose or intent of the act seems fairly clear to the observer and where its consequences are sufficiently definite to leave little doubt concerning its effects.”
Proponents claim that for some research purposes CIT generates different and sometimes more useful insights than attitude surveys and other quantitative methods, and can be more efficient than other qualitative methods such as in-depth interpretive ethnographies for identifying particularly effective and ineffective behaviors in given situations (Andersson and Nilsson 1964; Butterfield et al. 2005).

Next, we distinguish between three elements necessary for full application of CIT, and enumerate the principles supporting each element. The three elements are eliciting CI reports, selecting CIs for analysis, and analyzing CIs. All three activities are central to CIT per Flanagan (1954).

**Eliciting CIT Reports**

In Flanagan’s proposed CIT method, the unit of analysis is individual behavior. This is evident in his descriptions of techniques used to elicit incident descriptions, such as: “Describe the officer’s action. What did he do?” (p. 328) and “Think of some occasion during combat flying in which you personally experienced feelings of acute disorientation or strong vertigo.” (p. 329). Subsequent to Flanagan’s studies, individual behavior has been the unit of analysis in most—but not all—studies which claim to use CIT.

Flanagan proposed two “basic principles” for eliciting CI reports (Flanagan 1954):

- **a.** “reporting of facts regarding behavior is preferable to the collection of interpretations, ratings and opinions based on general impressions;”
- **b.** “reporting should be limited to those behaviors which, according to competent observers, make a significant contribution (whether positive or negative) to the activity.”

Regarding principle (a) Flanagan argued that surveys or interviews that ask subjects to report retrospectively on their typical behavior yield equivocal findings because of subject recall bias and other contaminating factors. Subsequent validation studies have confirmed that while it is difficult for respondents to confidently report on outcomes from typical behavior, recall is more accurate when they are asked to report on critical behaviors (Andersson and Nilsson 1964). So, in Flanagan’s CIT “typical” or “average” behavior was explicitly excluded from consideration.

Some CIT studies rely on direct observations: subjects perform particular tasks; trained researchers observe and code subject behavior per detailed coding schemes. Other CIT studies elicit unambiguous and complete critical incident (CI) reports, via interviews or questionnaires. To elicit CI reports respondents are prompted to describe specific episodes detailing either their own behavior or their recollection of someone else’s behavior. Thus, while some CIT studies rely on direct observations of subject behavior, many CIT studies rely on retrospective accounts of subjects’ own or others’ behavior, elicited through structured interviews or surveys. Per Flanagan (b), CI reports were to be elicited in relation to specific episodes of job performance in which respondents believed the behavior and outcomes of the behavior were particularly effective or ineffective (the “critical” aspect of each incident must be clear).


**Selecting CIT Reports**

The second element is selection, the task in which a researcher carefully considers a full set of potential CI reports and selects for analysis those that fulfill the selection principles. First, elicited behaviors described by respondents should have made a “significant contribution, either positively or negatively, to the general aim of the activity” (Flanagan 1954). As noted above, this principle derives from theories indicating that behaviors associated with less-typical outcomes (very negative or very positive) are easier to recall in accurate detail than typical behavior. CIT validation research compared direct observations to recall of critical behaviors and verified this principle (Andersson and Nilsson 1964). In Flanagan’s early job analysis studies subjects were asked to describe incidents involving very effective or very ineffective behavior. He explained (p. 329, quoting an earlier Air Force report):

“Well, the principal objective of job analysis procedures should be the determination of critical requirements. These requirements include those which have been demonstrated to have made the difference between success and failure in carrying out an important part of the job assigned... Too often, statements regarding job requirements are merely lists of all the desirable traits of human beings.”

Thus, regarding selection of incidents to analyze: Flanagan used deductive logic to compare incident reports against a priori criteria to decide whether each report was sufficiently clear in its: 1) specific outcome effectiveness or ineffectiveness and 2) tie between reported behaviors and outcome results. Essentially, the subject, after being prompted to describe an atypical incident, had to recount a clear story with a clearly identified protagonist and (if relevant) one or more other actors, a clear sequence of events, and definite closure (a clear outcome). Selected reports would be included in the CI analysis set.

**Analyzing Critical Incidents**

In industrial psychology and human resources management many CIT studies, similar to Flanagan, aimed to identify skill requirements for particular jobs or roles. Like Flanagan, some researchers collected 100 or more CI reports per study. Flanagan argued that, in order to spot important detailed skills at a level of granularity necessary for describing the tasks performed by a fighter pilot, many incidents needed to be examined before a researcher could feel confident of having reached theoretical saturation (defined as the point when no new information or themes are observed in additional interviews or incident reports). Thus, some subsequent CIT studies, seeking similar levels of granularity in identifying a comprehensive skill set for a job, also collected many CI reports. Other CIT skill requirements studies reported reaching theoretical saturation based on analysis of fewer incidents—as few as 40 or 50 when analyzing sub-roles within jobs or activities.

CIT studies need to describe how theoretical saturation was achieved and offer a clear justification that the number of CI reports collected was adequate for the research purpose. For example, when many CI reports are gathered, the researcher can prepare a chart showing which themes and sub-themes appeared in which CI reports (organized by themes, from most reports to least or organized by reports from most themes to fewest themes). Or, the researcher can make an argument based on data collection chronology. For example, in a set of 100 CI reports arranged chronologically, the researcher might be justified in arguing that theoretical saturation has been achieved if no new themes emerged in the last ten reports collected (based on multiple independent reviews of the primary data). In other CIT studies, the research objective may give rise to a different approach to theoretical saturation. For example, a study aiming to explore a heretofore unexplored topic—such as perhaps smart phone users’ very positive or very negative experiences with corporate apps (calling to mind the Apple-IBM partnership announced in summer 2014), may claim to have reached theoretical saturation when a given number of new themes have appeared in a collection of CI reports (i.e., when studying a new phenomenon, any one instance of a new theme may serve a revelatory purpose. Completeness in this context may be a secondary consideration, compared with uniqueness of observations).
The next key analysis concern for CIT is how to use CI reports to create a clear picture of the requirements for a job or role. Flanagan used inductive logic to classify behavior into categories. Although he was working prior to the development of the term “grounded theory,” his work was consistent with the discovery of grounded theory (Glaser and Strauss 2012). Subsequent CIT researchers (such as (Boyatzis 1998) have used both deductive (e.g., seeking evidence of previously-proposed behavior categories or themes) and inductive (open) coding (letting data reveal itself to the researcher).

Consistent with guidelines for a valid grounded theory analysis (Krippendorff 2013; Neuendorf 2002), the Findings sections of Flanagan’s papers reported on the number of incidents elicited and selected and usually included incident examples. Flanagan’s argument here was:

“Evidence regarding the accuracy of reporting is usually contained in the incidents themselves. If full and precise details are given, it can usually be assumed that this information is accurate. Vague reports suggest that the incident is not well remembered and that some of the data may be incorrect.”

By both counting incidents and including some or all incident reports in the Findings, a researcher makes the study procedures and data more transparent to readers, who can then verify whether their interpretation of the findings match those of the author/s.

We further note that the researcher should retain an archive of the complete set of critical incident reports, making it possible for other researchers to challenge the findings and offer their own interpretations on the basis of the same evidence. We do not include this last aspect as an essential criterion for CIT per se, since data retention for the purpose of re-examination by other researchers is a general norm for the advancement of science and not a unique requirement of any particular method.

Thus, a strong and complete CIT analysis will:

1) address theoretical saturation,
2) present and validate its coding methodology, and
3) clearly describe the selected and analyzed incident set and (at a minimum) provide some details or quotations directly from the CI reports.

**Full versus Partial Application of CIT**

Many prior studies in fields outside IS have validated the CIT elements described above as the method evolved and flourished over time. However, since CIT is not widely used in current IS research (as can be seen in our systematic literature review, below), we believe IS researchers need clear guidance on what would constitute a “full” CIT study as proposed by Flanagan and used by many subsequent researchers—that is, a study which includes the elements described above and summarized in Table 1. In particular, IS researchers who choose to study questions similar to those asked by Flanagan and other CIT researchers—such as, What are the key skill requirements of a particular job? or How does negative interaction between a service provider and a customer create distrust?—can make stronger claims if they conduct their studies in a manner that parallels those in the reference disciplines.

It is also possible to design a high quality study that utilizes a “partial” version of CIT. Consider, for example, a case study based on a semi-structured or loosely structured interview protocol. That study might have been designed to elicit information on specific topics, but not specifically designed to elicit CI reports. In the hypothetical loosely-structured interview, interviewees might be permitted to describe events that are meaningful to them, but not a central topic of the case study from the researcher’s perspective. The interviewer may allow an interviewee to continue down a narrative path to the point that the interview transcript will contain stories that are comparable to critical incident reports, even though eliciting CI reports was not a primary focus of the study. Some stories will contain sufficient detail that the researcher can apply CIT criteria to select them as CI reports, and analyze and discuss them using CIT analysis techniques. So, even if a study did not deliberately seek to elicit CI reports, if the paper’s analysis section presents and discusses CI reports, we would consider it to be an example of “Partial” CIT.
Alternately, a researcher might state that they used CIT and mention that critical incidents were elicited as part of their interview protocol. However, if the presentation and discussion of study findings does not describe individual behavior or include consideration of specific incidents, we would not consider this to be a CIT study (full or partial), even if authors described it as such.

To illustrate our approach to classifying CIT studies, we next discuss an IS study that was published before the ten year period under consideration in our current literature review: Kelly and Bostrom’s (1997) study of group support systems facilitation. They reported using CIT, and described the method as follows: from interviews with 26 experienced GSS facilitators, they collected and analyzed 93 critical incident reports of “effective or ineffective” interventions (their paper did not further specify selection criteria for including or excluding incident reports). Based on analysis of the CI reports, they proposed a model for managing socio-emotional issues in computer-mediated environments. The method description did not specify whether they asked facilitators to report on typical or atypical incidents. We observe that some quotations presented in their findings appeared to address typical situations and/or respondents’ general views on a topic. For example, on page 31 a GSS facilitator says “You need to use your gut as a barometer, but you need to check with [the] group, check [the] facilitator’s barometer.” Kelly and Bostrom applied both inductive and deductive logic to analyze and categorize findings. No mention was made of criteria for theoretical saturation, although 93 incident reports would be in the acceptable range.

We classify the Kelly and Bostrom study as Partial CIT rather than Full CIT, for the following reasons: they did not specify that they elicited accounts of atypical behavior (“particularly” or “very” effective or ineffective interventions would be atypical, but “effective or ineffective” could encompass all interventions) (Selection 3), and they did not report inclusion/exclusion criteria (Elicitation 1 and Selection 1). In classifying this study as Partial CIT, we are not speaking to the value of the study findings, since this paper underwent rigorous peer review before being accepted for publication in a top journal, and we respect that review process and the judgments of the reviewers. It might be that this research team actually did rigorously apply full CIT, but did not fully report all the pertinent method details in the published paper. A reader who is knowledgeable about CIT and expects to see conformance to all seven principles across the three elements of elicitation, selection and analysis, might not immediately trust the study findings, because of this missing information.

Other than (Kelly and Bostrom 1997), few other IS studies claimed to use CIT before 2004 (Clawson et al. 1993; Ellinger and Bostrom 2002).

We next examine studies in top IS journals that claimed to use CIT in the past 10 years, to determine to what extent CIT is represented in current IS research published in prominent journals.

<table>
<thead>
<tr>
<th>Element</th>
<th>Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Elicitation</strong></td>
<td></td>
</tr>
<tr>
<td>E1 Elicitation Clarity</td>
<td>Elicitation of incident reports guided by clear criteria, whether based on directly observed behavior or actors’ descriptions of their own recalled behavior (and consistent with Selection criteria below).</td>
</tr>
<tr>
<td><strong>Selection</strong></td>
<td></td>
</tr>
<tr>
<td>S1 Individual</td>
<td>Focus on specific behavior performed by specific individuals.</td>
</tr>
<tr>
<td>S2 Report Clarity</td>
<td>Unambiguous story: clear protagonist, sequence of events, outcome.</td>
</tr>
<tr>
<td>S3 Critical</td>
<td>Specific behavior tied to particularly effective or particularly ineffective outcome/s (NOT average behavior or typical outcomes)</td>
</tr>
<tr>
<td><strong>Analysis</strong></td>
<td></td>
</tr>
<tr>
<td>A1 Incident Details</td>
<td>CI reports or CI fragments enumerated or reported on in Findings</td>
</tr>
<tr>
<td>A2 Coding Clarity</td>
<td>Coding method/s clearly justified (whether inductive, deductive, or both).</td>
</tr>
<tr>
<td>A3 Saturation</td>
<td>Theoretical saturation clearly justified.</td>
</tr>
</tbody>
</table>

**Table 1. Criteria for Full (native) Critical Incident Technique**
A Systematic, Reproducible Literature Review

In order to ensure reproducibility of our review of current usage of CIT in IS research, we chose to systematically review papers claiming to use CIT for a recent ten year period (2004-2013) and in a subset of prominent IS publications—the eight journals comprising the AIS Senior Scholars “Basket” (ECIS, JAIS, JIT, JMIS, JSIS, ISJ, ISR, MISQ), which are widely seen as the top journals of our field. As discussed by (Webster and Watson 2002) and (Levy and Ellis 2006) it is important to specify a clear literature review process, but it is not necessary to claim that all extant studies have been found, as long as the selected literature subset is appropriate given the aims of the review.

A grounded literature review emphasizes maximally transparent review procedures and incorporates inductive reasoning to identify and classify content in the reviewed papers, enabling “key concepts to surface” (Wolfswinkel et al. 2013). A recently-introduced literature review method aims to “reach a thorough and theoretically relevant analysis of a topic” in five stages: define, search, select, analyze, present (Wolfswinkel et al. 2013). We present and discuss our review method per the first three of five stages in Table 2 (below), and we complete our review description in the text following Table 2.

1.0 Define
1.1 Define criteria for inclusion/exclusion
1.2 Identify fields of research
1.3 Determine appropriate sources
1.4 Decide on specific search terms

The Basket (ECIS, JAIS, JIT, JMIS, JSIS, ISJ, ISR, MISQ) consists of “prominent IS journals,” a proxy for acceptance by the IS community. Search limited to ten years in order to identify current use of CIT. We excluded essays, editorials, and reviews.

Regarding specific search terms: We did not know if all CIT studies would use the term “critical incident technique”. Might some use “critical incident method”, “critical event” technique, and so on? A forward-citation search on Google Scholar identified more than 6000 papers and books citing (Flanagan 1954). Reading about 40 of these papers, we observed some variation in use of the terms “technique” versus “method”, but consistent use of “critical incident” as a term.

2.0 Search
2.1 Search
An assistant, already hunting for other studies, read abstracts of all Basket papers in 2010, 2011 and 2012. Noting any paper with the phrase “critical incident” in the Abstract, she found two papers. She next conducted online searches for “critical” or “incident” in the text of Basket papers from 2008-2012, yielding 8 papers. While the phrase “critical incident” did not always indicate CIT as a research method, we noted that the phrase was always used when CIT was claimed as a research method.

In February 2014 we conducted a final online search on the eight Basket journals for the ten year period 2004-2013, for the phrase “critical incident”. This yielded 23 articles.

3.0 Select
3.1 Refine the sample
We eliminated non-empirical papers—essays, editorials, literature reviews, or teaching cases—and false positives (“critical” or “incident” used in ways that did not mean use of CIT as a research method; e.g., “incident” referring to an event that triggered an organizational change. Other false positives: discussions of critical theory in IS research, “critical” denoting something important or referring to specific incidents in the history of an organization or an industry).

We also checked if “critical incident” appeared in paper Abstract, if author/s claimed to use CIT, and if Flanagan (1954) was cited (manually examining each paper’s Abstract, Methods section and References). 12 papers did not claim to use CIT. 13 papers did claim to use CIT (of which, 10 cited Flanagan). As a final check, we ran a new online search in the Basket journals, for “Flanagan” and “1954” for the ten-year period. This did not reveal any additional papers claiming to use CIT.

Table 2. First three literature review stages (consistent with Wolfswinkel et al., 2013)

By the end of stage 3, we had identified 25 papers using the term “critical incident,” of which 13 studies claimed to use CIT. We sought to identify those of the 13 studies that used CIT in its full or native form (as described by Flanagan and formalized in Table 1), versus in partial or adapted form.

---

**Table 2. First three literature review stages** (consistent with Wolfswinkel et al., 2013)

By the end of stage 3, we had identified 25 papers using the term “critical incident,” of which 13 studies claimed to use CIT. We sought to identify those of the 13 studies that used CIT in its full or native form (as described by Flanagan and formalized in Table 1), versus in partial or adapted form.
We first checked whether CI reports were obtained through direct researcher observation or indirectly via surveys or interviews. No study in this ten-year sample was based on direct researcher observation or surveys; all the studies relied on interviews. We next checked whether CI reports were elicited per clear criteria consistent with selection criterion (E1), whether incidents focused on and described individual behavior (S1), selected incidents told a complete and clear story (S2), and extreme anchors (e.g. “very,” “particularly,” etc.) were used to elicit or select CI reports (not “typical” behavior and outcomes), (S3).

We also checked whether each paper’s Findings section noted how many incidents were analyzed and/or whether the Findings presented some incidents so that the reader could verify how the author/s operationalized their definition of “critical incident” (A1). We further coded for the other two CIT Analysis criteria: clear and defensible procedure for coding incident data (A2) and defensible argument for theoretical saturation (A3). This last aspect relied on two sub-criteria: (A3.1) theoretical saturation based on a high quantity of incidents selected for analysis, or evidence that no new codes emerged in the last x incidents, or (A3.2) saturation evident based on the nature of the study (e.g., an exploratory case study could presumably rely on a single revelatory critical incident to either refute prior theory or propose new theory). While Flanagan took a narrower stance to define acceptable levels of theoretical saturation based on quantitative evidence, our more inclusive criteria allowed for an exploratory or interpretive qualitative argument. We felt this was appropriate, given that the journal peer review process should have already judged the quality of the evidence supporting the study’s conclusions and contribution to theory.

During our review, we became aware of an aspect that we had not anticipated in advance but now coded: some mixed methods studies used CIT as one of two or more methods used (papers coded as single or mixed method). We also noted that few qualitative studies reported relying solely on CIT; some case studies or other qualitative research elicited critical incidents as part of a broader interview protocol.

Table 3 summarizes the 13 papers claiming to use CIT as a research method, based on searching for Basket papers containing the phrase “critical incident” from 2004-2013 and assessing them per the criteria described above. An Appendix summarizes the 12 papers that contained the phrase “critical incident” but were excluded from our selection set because they did not report on empirical research or did not use the term “critical incident” in the context of their research method.

To summarize, our systematic literature review aimed to learn: To what extent and how has CIT been employed in recent IS empirical research published in prominent IS journals?

Review Findings

Our search in the IS Scholars Basket during the period of interest (2004-2013) revealed two papers (Thomas and Bostrom 2010a; Thomas and Bostrom 2010b) that claimed to use critical incident technique as the sole research method. Five papers claimed to elicit critical incidents as part of a broader interview protocol in a qualitative study, and six papers claimed to use CIT as one of two or more methods in a mixed-methods study (study using multiple methods to capture and analyze data). After reviewing the Methods and Findings sections for these papers (and their appendices if relevant) we classified 11 papers as using partial CIT, based on the criteria laid out above in Table 1.

Two papers co-authored by Thomas and Bostrom (2010a; 2010b) reported on findings from one CIT study directed at understanding some leadership issues in virtual teams by conducting CIT interviews with 13 experienced team leaders. Their broader study (MISQ Mar 2010) yielded 52 incident reports, each describing leader interventions in one of 30 projects. Thomas and Bostrom defined a “critical incident of intervention ... as one in which the VT leader took action to improve team interaction by affecting the usage of one or more ICT and he/she deemed his/her action either particularly successful or unsuccessful and could point to specific team interaction results.” This definition is faithful to Flanagan in emphasizing atypical behavior (“particularly successful or unsuccessful”) and certainty of outcome. In total, that study and its two resulting papers are the only ones we can classify as full CIT, because all core elements and principles for CIT elicitation, selection, and analysis are present.
<table>
<thead>
<tr>
<th>Journal</th>
<th>Authors claim CIT?</th>
<th>Cite Flanagan (1954)?</th>
<th>CIT in Abstract?</th>
<th>CIT in Keywords?</th>
<th>E1 elicit incident clarity</th>
<th>S1 individual behavior clarity</th>
<th>S2 story clarity</th>
<th>S3 critical</th>
<th>A1 report details</th>
<th>A2 coding clarity</th>
<th>A3 saturation</th>
<th>% Elements Present</th>
<th>Full CIT?</th>
<th>Single or Mixed Method</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>EJIS 22(6) Nov 2013</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>43</td>
<td>N</td>
<td>S</td>
<td>Kutsch et al.</td>
</tr>
<tr>
<td>EJIS 22(4) Jul 2013</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>43</td>
<td>N</td>
<td>S</td>
<td>Schlichter and Rose*</td>
</tr>
<tr>
<td>EJIS 19(3) Jun 2010</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>29</td>
<td>N</td>
<td>M</td>
<td>Ngwenyama and Norbjerg</td>
</tr>
<tr>
<td>EJIS 19(2) Apr 2010</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>100</td>
<td>Y</td>
<td>S</td>
<td>Thomas and Bostrom*</td>
</tr>
<tr>
<td>EJIS 16(2) Apr 2007</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>71</td>
<td>N</td>
<td>M</td>
<td>Hackney et al.</td>
</tr>
<tr>
<td>ISJ 23(1): 2013</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>57</td>
<td>N</td>
<td>M</td>
<td>Rose and Schlichter*</td>
</tr>
<tr>
<td>ISJ 22(6) Nov 2012</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>43</td>
<td>N</td>
<td>S</td>
<td>Klendauer et al.</td>
</tr>
<tr>
<td>ISJ 17(4) Oct 2007</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>57</td>
<td>N</td>
<td>S</td>
<td>Yoo et al.</td>
</tr>
<tr>
<td>ISR 22(4) Dec 2011</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>43</td>
<td>N</td>
<td>M</td>
<td>Cenfetelli and Schwarz</td>
</tr>
<tr>
<td>ISR 15(4) Dec 04</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>43</td>
<td>N</td>
<td>M</td>
<td>Koh et al.</td>
</tr>
<tr>
<td>JAIS (no papers found)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>JIT 28(4) Dec 2013</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>43</td>
<td>N</td>
<td>M</td>
<td>Zolper et al.</td>
</tr>
<tr>
<td>JMIS 23(4) Spr 2007</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>57</td>
<td>N</td>
<td>S</td>
<td>ØSterlund et al.</td>
</tr>
<tr>
<td>JSIS (no papers found)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MISQ 34(1) Mar 2010</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>100</td>
<td>Y</td>
<td>S</td>
<td>Thomas and Bostrom*</td>
</tr>
<tr>
<td>Total “Y”s (Yes)</td>
<td>13</td>
<td>10</td>
<td>2</td>
<td>2</td>
<td>6</td>
<td>6</td>
<td>3</td>
<td>9</td>
<td>5</td>
<td>12</td>
<td>11</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Percent “Y”s</td>
<td>100</td>
<td>77</td>
<td>15</td>
<td>15</td>
<td>46</td>
<td>46</td>
<td>23</td>
<td>69</td>
<td>35</td>
<td>92</td>
<td>84</td>
<td>NA</td>
<td>15</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Schlichter & Rose and Rose & Schlichter reported different findings from one study in two papers, as did Thomas and Bostrom (EJIS 2010) and (MISQ 2010).

**Table 3. IS Senior Scholars Basket, 2004 – 2013 Empirical Papers Claiming to Use CIT**

Mapping the Table 3 findings to the CIT criteria (first listed in Table 1) to Table 4 (below), we first discuss the few elements that were used with a high degree of conformity to the CIT criteria.

Regarding coding scheme clarity: Table 4 shows that 92% of the 13 papers (all but one of the papers claiming to use CIT) provided clear coding details. Most qualitative research methods in IS—including case studies, discourse analysis, ethnographies, and other methods that rely on textual data—specify that the researcher is to clearly explain their stance (positivist, interpretive, critical) and their coding scheme. So, it is not surprising that most CIT studies conformed to this criterion.

A similar observation can be made regarding the defense of theoretical saturation, which is normally expected of qualitative studies (although not all case studies). Here we discuss a few of the papers from the 13 in our sample, in order to illustrate how we classified our findings, which show low conformance to the CIT criteria by many studies claiming CIT.
One of the two papers that did not report how saturation was achieved—a mixed-methods study by Koh et al. (2004)—was based on interviews with 15 respondents. This paper indicated that critical incidents were elicited, but did not report how many CI reports were obtained and selected for analysis. Without this data, it is not possible for the reader to judge whether saturation was achieved.

Another paper (Klendauer et al. 2012) mentions the term “theoretical saturation” but does not explain what procedure was used for ensuring saturation. We further observe that Klendauer et al. reportedly elicited incident reports by asking respondents to describe “effective or ineffective behavior [in] ... a typical and important work situation.” This elicitation technique would not fully conform to Selection criterion S3 (Critical) because of the word “typical” and because of lack of an adjective indicating particularly or very effective or ineffective. Furthermore, that paper did not indicate whether any of the Selection criteria (S1, S2, S3) were employed to further cull from the initial reports on 64 “situations” (or “projects”) a final set of critical incident reports for analysis.

<table>
<thead>
<tr>
<th>Elicitation</th>
<th>Selection</th>
<th>Analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>E1 Clarity: 46%</td>
<td>S1 Individual: 46%</td>
<td>A1 Incident Details: 46%</td>
</tr>
<tr>
<td>Elicitation of CI reports guided by clear criteria, directly observed behavior or actors’ descriptions of recalled behavior (and consistent with Selection criteria).</td>
<td>Focus on specific behavior performed by specific individuals.</td>
<td>Critical Incident reports or CI fragments enumerated or reported on in Findings</td>
</tr>
<tr>
<td>S2 Report Clarity: 23%</td>
<td>S3 Critical: 69%</td>
<td>A2 Coding Clarity: 92%</td>
</tr>
<tr>
<td>Unambiguous story: clear protagonist, sequence of events, outcome.</td>
<td>Specific behavior tied to particularly effective or particularly ineffective outcome/s</td>
<td>Coding method/s clearly justified (whether inductive, deductive, or both).</td>
</tr>
<tr>
<td>S3 Saturation: 84%</td>
<td>A3 Saturation: 84%</td>
<td>Theoretical saturation clearly justified.</td>
</tr>
</tbody>
</table>

**Table 4. 11 Scholars Basket Papers Claiming CIT: % Conforming to Full CIT Criteria**

Fewer than half of papers claiming to use CIT provided evidence of report clarity (S2), or indicated how many CI reports were collected, or included any CI reports (in whole or in part) in their presentation of findings (A1). Incident reports are the raw data for a CIT study. A paper based on a survey would explain how variables were operationalized and validated, report how many subjects responded, present some summary data regarding subjects’ demographics, and provide a summary of the findings. Similarly, a CIT paper—whether based on a positivist, interpretive or critical study—should explain how CI reports were initially elicited and subsequently selected for analysis (including criteria used to exclude some candidate reports from analysis). While a paper may not be able to present every CI report (especially if as many as 100 reports were elicited) it should include some examples so that the reader can verify that the researchers indeed operationalized “critical incident” consistent with best practice for this method.

We conclude that, although each of the 13 papers examined in our study made a unique and worthwhile contribution to the literature (a conclusion based on our respect for the peer review and editorial processes used by our field’s top journals), nevertheless the findings of our review suggest that either CIT was underutilized as a research method in most of these studies, or the final versions of the published papers do not accurately and completely report how this method was used.

**Discussion, Limitations, Suggestions**

Only two papers in our data set—both by Thomas and Bostrom—used CIT in its full (or native) form. This finding supports our assertion that CIT is a lightly-utilized research method in the IS discipline. To begin our discussion, we consider these two questions: What does CIT offer to IS researchers? Why isn’t it more often used?

**Why CIT? Why isn’t it used more?**

CIT offers a means to describe how people effectively or ineffectively use systems. The product of a well-designed and executed CIT study is a set of condensed, generalizable dimensions describing key behaviors and beliefs that lead either to success or failure in a job, role, or task. IS faces many challenges understanding how to design systems that properly support business processes in varied contexts. Close,
detailed examination of how individuals operate in those contexts and with those systems can lead to stronger theory to guide systems designs. CIT is an in-depth method designed for such close examination.

For example, imagine studying any large participatory IS today involving various user communities and IT personnel. Each has a role in the system. What adoption and routinization behaviors are associated with most effective use of a healthcare enterprise system? A study such as Khoumbati et al.’s could be usefully complemented with CIT analysis (Khoumbati et al. 2006). What are the most effective behaviors of IT leaders and team members when working with their business counterparts? Effective collaboration between IT and business employees contribute to IS success (Henderson and Venkatraman 1994), and personality traits alone do not explain variation (Kaiser and Bostrom 1982). Particularly effective or particularly ineffective behaviors of team members can be studied and dimensionalized using CIT. So doing, we might advance the IS field’s understanding of how usage and participation through specific roles lead to strategic alignment. Interesting behaviors—like how to conduct systems project recovery and how to lead iterative development effectively—would be more easily approached. We might also improve guidance to IS leaders regarding how to define various roles and provide effective training for people in those roles.

Of the 11 partial-CIT studies, we observed interesting examples of methodological experimentation and innovation. Some studies that we classified as Partial CIT apparently did not initially set out to elicit critical incidents, yet found opportunities to select CI reports from within their broader qualitative data sets and to incorporate these into their analysis. By providing guidance on criteria for incident selection (focus on critical individual behavior and a clear story and outcomes) and three analysis practices (apply clear coding criteria, achieve theoretical saturation, report incident details in the findings), we have contributed needed clarity for those qualitative researchers who might want to revisit their data to systematically select and closely examine critical incident reports embedded in their interview data. The guidelines presented in Table 1 should be helpful to researchers who choose to use CIT either as their sole method or as one method in a multi-method study.

Some of the 11 studies that we classified as partial CIT might have actually used the full CIT criteria or included several applicable procedures (such as justifying theoretical saturation or using a rubric to select complete and clear incident reports), yet these important method details were not included in the published paper. Possibly some of the studies summarized above actually conformed very well to CIT criteria, yet the published papers failed to provide the necessary evidence for the reader to verify conformance. While it would be reassuring to know that a given study actually did conform to all CIT criteria, it is nevertheless somewhat troubling to consider that authors or reviewers either failed to notice that the needed description was not provided, or did not believe this methodological information was necessary. We remain concerned that CIT, despite being a well-respected method in some reference disciplines, appears to be an under-appreciated and possibly a misunderstood research method in IS.

It is also possible that some authors used the term “CIT” loosely in describing their work, because they were not fully aware of the requirements of this research method. While their studies nevertheless made useful contributions to the IS literature, they might have made even stronger contributions had they conformed to the criteria laid out by Flanagan and subsequent CIT researchers in other fields, since these criteria aim to allow the researcher to focus closely on actual behavior in real situations, with definite (“critical”) positive or negative outcomes.

Based on a review of papers that did claim to use CIT, we cannot confidently comment on why CIT has been underutilized, misappropriated, or adapted. Our study findings, by demonstrating low conformance to CIT criteria in Basket papers, may point to issues in IS researchers’ and reviewers’ awareness of or appreciation for CIT. If so, the community of IS scholars may be missing out on fruitful opportunities to learn more about users’, customers’, partners’, IT professionals’, and managers’ behavior through eliciting and closely examining critical incident reports.

A disciplined approach to elicitation and selection of critical incident reports is a key to achieving strong findings (Andersson and Nilsson 1964), based on many prior studies that found that subjects recall atypical (“critical”) incidents better than typical behavior and that subjects are better able to establish a clear link between particularly effective behavior and particularly good outcomes and between particularly ineffective behavior and particularly bad outcomes, as compared with typical behavior and typical outcomes. Furthermore, Flanagan argued that a CI report needs to be unambiguous in terms of who the actor/s is/are and how each actor individually behaved; these reports provides a foundation upon which analysis
proceeds. A “critical incident” that refers to how a company or institution behaved would contain a great deal of ambiguity (Who, exactly was behaving: the sales rep? his manager? a customer service rep?). A broad assertion that “our customers appreciate that we are willing to go the extra mile for them” would be similarly ambiguous: (Which customers? What extra mile? Which of “our” employees are willing and which ones actually go that extra mile?)

As discussed above, incident elicitation and selection are unique elements of CIT, while analysis can be viewed through the lens of best practices for deductive, inductive or critical coding of qualitative data—topics about which (we hope) the community of IS researchers are sufficiently knowledgeable to carry out and our community of IS conference and journal reviewers are sufficiently knowledgeable to judge.

If the findings of our limited-scope literature review generalize beyond the Senior Scholars Basket (a big IF), then one might infer that the IS community’s strongest understanding of Critical Incident Technique relates to the analysis stage of a CIT study. We judged the principles of theoretical saturation (A3 – 84%) and coding clarity (A2 – 92%) to be evident in most of the papers sampled. As discussed above, we believe that this is because guidelines for the conduct of qualitative research in general address these two principles (Krippendorff 2013; Neuendorf 2002).

In our view, CIT is one of several valuable methods for investigating individual behavior in relation to information technology (e.g., use or non-use, detailed usage behavior, supervision of or participation in software development projects or IS planning and governance processes), from an interpretive, positivist, or critical perspective. One of the authors of this paper was the lead investigator on the two full CIT studies noted above (Thomas and Bostrom 2010a; Thomas and Bostrom 2010b). He is an ardent proponent of using full CIT to investigate a variety of questions in the IS discipline.

A researcher who chooses not to adhere to the full set of CIT elicitation, selection and analysis criteria laid out here faces several risks. One risk is that their study will fail to provide sufficiently strong evidence supporting their claims. A second risk is that readers who are knowledgeable about CIT may distrust the reported findings. Thus, we suggest that when a research team chooses to adapt CIT to suit their purposes, they have two obligations to fulfill:

1) Inform your readers as to the choices you made that represent a departure from full CIT criteria for elicitation, selection and analysis, and

2) Provide a convincing argument as to how your chosen adaptation affects the likely validity, reliability, and generalizability of your findings.

Study Limitations

We acknowledge a limitation of our study: we restricted our systematic review to the eight top IS journals comprising the IS Senior Scholars Basket, which we believe represents the best work of our field. Note our use of the word “represents.” We know that the Scholars Basket does not contain all of the best IS papers published, since we are aware of many much-cited papers not published in the Basket. However, since the top IS scholars have blessed these journals as representing our field’s best work, we assume that many IS scholars aim to publish their best work in these journals; they go to other journals when they fail to achieve acceptance in the Basket.

Many IS scholars regularly read Basket publications along with relevant specialty journals, and they also follow citation trails back to seminal works published in other outlets and attend IS conferences like AMCIS, ECIS, HICSS, ICIS and PACIS to learn about new work in the IS sub-fields that interest them. Scholars in our field also stay abreast of relevant work in reference disciplines such as computer science, management or sociology. Keeping up with papers published in the Basket is a widely shared professional norm, while in contrast each IS sub-discipline involves different sets of specialty journals and conferences.

Thus, we contend that limiting our review to the eight Basket publications fairly represents those papers most likely to broadly represent the IS research community. We acknowledge that a broader search, in another 5, 10 or 20 IS journals, would yield other studies claiming to use CIT and perhaps a different proportion of studies using full versus partial CIT. In the absence of a single ranked journal list representing the consensus of the IS community of scholars, we did not have a strong selection criterion to use to select additional journals beyond the Basket for inclusion in a broader literature review.
Suggestions: CIT in IS

The great promise of the critical incident technique is that it can be applied retrospectively through reports gathered from the individual performing a task or engaging in a transaction. Like Flanagan’s early studies and similar to numerous studies in industrial and organizational psychology, the two Full CIT papers in our sample sought to use CIT to identify key skills for successful performance in a particular role, as well as to identify specific behaviors associated with poor performance. Given that their study aims were similar to other studies aiming to identify key skills and particularly ineffective behaviors, it makes intuitive sense that Thomas and Bostrom would apply CIT in a manner consistent with prior studies aiming to do the same thing. Beyond intuition, prior research has already marshaled evidence to persuasively argue that each of the seven criteria provided in Table 1 will support the aims of such a study.

Thus, CIT is an excellent tool for IS researchers to have in their kits, particularly if they want to learn about phenomena similar to those studied by researchers in organizational psychology or marketing. Studies aiming to answer similar research questions—such as “What are the key skills necessary for effective leadership as a CIO?”—could make good use of CIT as one helpful research method. Consider two items that are high on the CIO agenda in 2014: Big Data and Security (“What are the key skills necessary for successful business analytics?” or “What are the key skills necessary for successful information security incident response?”). If one wished to carry out a study to identify key skills needed to be a data scientist or security expert, and also to identify the behaviors most likely to point to potential poor outcomes in these roles, then Full CIT would be a great research method.

Another stream of research that could benefit from CIT studies is the governance and management of outsourcing relationships. Prior work has established that cultural differences and other issues can impede successful outsourcing. CIT studies can contribute a deeper look at the behaviors associated with successful or unsuccessful coordination, collaboration, and control.

We noted above that CIT has been extensively used in marketing studies, particularly those aiming to understand successful and unsuccessful service encounters. One intriguing marketing study examined service failures that gave rise to incidents of “customer rage” (McColl-Kennedy et al. 2009). Given the many anecdotal accounts of the disconnect in communication between IT staff and non-IT employees, a series of CIT studies could closely examine IT customer rage incidents related to their interactions with help desks or the automated help features on software applications they use. In order to appropriately build on the substantial legacy of extant marketing research on service failures and successes, IS researchers will benefit from adhering to the CIT criteria that have long been used in marketing. The danger in partially applying CIT, without using all three elements and seven criteria is that the results of such studies may be seen as questionable or untrustworthy. Reviewers and editors who insist that these details be reported will contribute a valuable service to researchers and readers.

While CIT is well suited to some questions, it is less well suited to others. For example, if the context of use is the researcher’s primary concern, then one or more case studies would be the preferred qualitative method to gather relevant data, since case study method is designed to consider interactions among a large number of contextual and individual and other variables at once (that is the purpose of rich or “thick” description of a case). If the industry or organization is the unit of analysis, CIT would not be a useful research method. If the research focus is on the technology—such as when a study aims to prove that a new mechanism for inexpensive identification tags is superior to an older RFID technology — then CIT is not likely to play a role. A study of an organization’s IT investment planning that draws on economic theories (e.g., portfolio theory, options theory), is also not a likely candidate for CIT. However, if the primary focus is on an individual’s behavior in a job or role, and/or in relation to other individuals or entities, then CIT is an ideal method, particularly in contexts that are episodic in nature (such as help desk interactions, periodic IT planning exercises, purchasing, and other tasks or events that occur with some regularity). CIT allows many critical incident reports to be collected across multiple organizational and industry contexts, containing detailed descriptions of behavior. Neither of these two qualitative methods is better than the other in the abstract; each is suited to a different purpose.

CIT and Other Research Methods

We further argue, consistent with the IS scholars cited in our Introduction, that within a stream of research—such as studies of virtual collaboration or studies of social media use and impact or studies of mobile commerce—more will be learned if a variety of qualitative and quantitative methods (e.g., case
studies, ethnographies, CIT studies, surveys, experiments, simulations) are conducted over time, each focusing on the variables best suited to the method used.

Some papers in our data set reported applying partial CIT as one of several research methods. These researchers evidently saw value in applying some CIT elements in conjunction with one or more other research methods. This is an interesting, and potentially useful way to use mixed-methods for stronger study findings.

We turn now to a discussion of the implications of our work for doctoral education in IS.

**CIT and the IS PhD**

From study design to execution, submission for review and final publication, the IS research path contains many obstacles and challenges. PhD programs in Industry-Organizational Psychology programs teach how to apply CIT, whereas most IS PhD programs do not provide this training. Perhaps a lack of awareness about how to conduct CIT studies leads to sparse and possibly sub-optimal usage of CIT in IS research. So, one implication of our study is that CIT training should be made available to IS PhD students.

CIT is one among many useful research methods for IS. We believe CIT is particularly useful for an early-stage IS researcher. Unlike many other qualitative methods such as case studies and ethnographies, CIT is a highly structured method at the elicitation and selection phases. This makes it amenable to an apprenticeship model (first the apprentice takes notes while an experienced interviewer conducts several CIT interviews. Later, the expert observes the apprentice conducting CIT interviews, and eventually the apprentice can handle these on his/her own). Analysis can be segmented into the easier deductive coding followed by more demanding rounds of open and axial coding. At each of these stages the research team can come together to determine inter-rater reliability, achieve consensus on open codes and so on. Thus, a student or early-stage faculty member has an opportunity to progress through several stages of skill development, yielding, from a single study, good insights on how to conduct interviews, how to engage in deductive analysis, and how to let the data “do the talking” in inductive analysis.

CIT is but one or many useful methods for the IS researcher’s toolkit. An individual or team doctoral assignment to conduct a review similar to this (but focused on another IS research method) could help our students become aware of methodological options beyond Likert-scaled attitude surveys. Such a project could be done fairly early in one’s doctoral program, which might yield early conference papers similar to this one, to present at IS conferences.

**Conclusions**

We set out to assess the current state of critical incident technique (CIT) in IS research published in prominent IS journals. We found 13 papers that claimed to use CIT, although of those only two fully addressed the seven criteria traditionally employed in CIT studies reported in industrial and organizational psychology and marketing journals. The elicitation, selection and analysis criteria we identified will help ease the process of addressing this concern going forward.

One implication of our review of recent IS papers claiming to use CIT is that the findings of a study that fully conforms to the seven CIT criteria are likely to be more strongly accepted by those researchers within and outside the field IS who are already familiar with CIT. We have argued that researchers who use either full or partial CIT have an obligation to a) reveal study limitations (especially those adaptations that depart from full CIT criteria) and b) argue persuasively as to how their research method produced valid, reliable, generalizable results and/or pointed to intriguing suggestive results worthy of further study.

A recent paper, based on an analysis of IS research from 1995 to 2011, proposed the following concerning the importance of methodological diversity in IS research in light of IS’s important role as an interdisciplinary nexus (Bernroider et al. 2013): “Key to a healthy development of IS is the degree of permeability (fluidity, flexibility, and variety) allowing IS to take from or give knowledge to other disciplines.” Diverse methods enable cross-disciplinary trust. Since information systems touch every business and organizational function, the IS field in general benefits from inclusion of many methods in our research toolkits (Ågerfalk 2013; Venkatesh et al. 2013), and from greater methodological diversity across the IS field as a whole. By reviewing how CIT is used, and offering suggestions for how it can be used in future IS studies, we have contributed to the promotion of methodological diversity in the field.
A second contribution of our paper is that we illustrated a repeatable literature review process (adapted from Wolfswinkel, Furtmueller and Wilderom 2013) that could be applied to reviews of other IS research methods. Our review provided useful findings on current use of CIT in prominent IS publications, some challenges in its recent use, and guidelines to strengthen the IS community’s understanding of this important research method. Similar reviews of other research methods would offer similar value by reporting on current use of those methods and offering guidance for their effective use in IS research.

To the extent that the IS community thus learns about lightly utilized but useful research methods such as CIT, we will be able to better diversify the methods used in IS research and thus build stronger theories. Our initial investigation within the Basket opens the door for broader discussion of the merits of CIT and other lightly utilized methods for IS research, which we strongly encourage.

Acknowledgements

One of the authors gratefully acknowledges Bentley University, which provided a summer research grant to supervise the literature review and produce a first draft of the paper. All of the authors gratefully acknowledge Robert Bostrom who ably serves as the “grandfather” of CIT research in IS.

References


Appendix

12 Basket papers containing “critical incident” but did not use CIT

<table>
<thead>
<tr>
<th>IS Journal</th>
<th>Author</th>
<th>Why Eliminated</th>
</tr>
</thead>
<tbody>
<tr>
<td>EJIS 13(2) Mar 2004</td>
<td>Mustonen-Ollila et al</td>
<td>“critical incident” not used to mean method</td>
</tr>
<tr>
<td>ISR 20(3) Sep 2009</td>
<td>Harris, Collins, Hevner</td>
<td>“critical incident” not used to mean method</td>
</tr>
<tr>
<td>ISR 21(4) Dec 2010</td>
<td>Schultze and Orlikowski</td>
<td>Research commentary; advocates CIT as research method</td>
</tr>
<tr>
<td>JAIS 12(11) Nov 2011</td>
<td>Xu et al</td>
<td>“critical incident” not used to mean method</td>
</tr>
<tr>
<td>JAIS 9(3) Apr 2008</td>
<td>Chen et al</td>
<td>“critical incident” not used to mean method</td>
</tr>
<tr>
<td>JAIS 5(11/12) Dec 2004</td>
<td>Cenfetelli</td>
<td>Essay; advocates CIT as research method</td>
</tr>
<tr>
<td>JIT 22(4) Dec 2007</td>
<td>O’Callaghan</td>
<td>Teaching case; no mention of CIT as method</td>
</tr>
<tr>
<td>JIT 22(3), Sep 2007</td>
<td>Sauer</td>
<td>Editorial; “critical incident” not used to mean method</td>
</tr>
<tr>
<td>JSIS 16(2) June 2007</td>
<td>Vaast</td>
<td>Does not claim CIT as method; one of seven topics in semi-structured interview protocol: “what anecdote or incident that happened at Eastern Hospital comes to the respondent’s mind?”</td>
</tr>
<tr>
<td>MISQ 37(4), Dec 2013</td>
<td>Seidel, Recker, Vom Brocke</td>
<td>Interview protocol includes question: “Do you recall a critical incident or any other trigger that started that change?”</td>
</tr>
<tr>
<td>MISQ 36(3) Sep 2012</td>
<td>Carlo et al</td>
<td>“critical incident” not used to mean method</td>
</tr>
</tbody>
</table>