

## **AN ANALYSIS OF THE CRITICAL INCIDENT TECHNIQUE (CIT) AS AN EFFECTIVE METHOD FOR INSTRUCTOR EVALUATIONS**

Joe Hanna, New Mexico State University, Department of Marketing and General Business, Box 30001, Department 5280, Las Cruces, New Mexico, 88003 (505-646-7168).

Elise Truly Sautter, New Mexico State University, Department of Marketing and General Business, Box 30001, Department 5280, Las Cruces, New Mexico, 88003 (505-646-3341).

### **ABSTRACT**

This paper focuses on the application of the Critical Incident Technique (CIT) as a method of service quality evaluation in teaching. The authors designed a study to examine if the CIT was an appropriate method to aid in instructor evaluation. The background and methodology of the technique will be examined. This will be followed by a five step description of the experiment, and then a detailed summary of the results and conclusions.

### **INTRODUCTION**

Universities are currently one of the premier service industries in the United States, with students coming from around the world to experience the unique services of our nation's higher education system. In order to maintain and improve current service levels in our education market, it will be necessary for educators of all disciplines to put more emphasis on the services required by students. With this service emphasis in mind, the CIT may be a very useful tool to aid educators in continued instructional improvement and development.

This technique was originally developed in 1954 by John Flanagan to assist in pilot training for the Aviation Psychology Program of the United States (Flanagan 1954). Since this time, the method has evolved into a popular technique often used in business practice to aid managers in determining employee effectiveness in such areas as work motivation (Herzberg, Manseur, & Snderman 1954), negotiation and bargaining power (Steers 1984; Luthas 1985; George 1989), and job performance (Fitts & Jones 1961). In each case mentioned above, the technique was used because of its ability to examine interpersonal interaction in the proper evaluation of professionals (Day & Bodur 1978; Quelch & Ash 1981). Since the CIT has been successfully

implemented as a method of evaluation for employees in diverse situations, it appears logical to consider this technique as an appropriate evaluation method for instructors in college classrooms.

Most would agree that a primary responsibility of college instructors is to aid students in the learning process. Unfortunately, understanding how an instructor can best assist a student in this process is not an easy task. In the current college environment there is limited opportunity for students to critique their instructors. Furthermore, if the opportunity does arise, it is usually at the conclusion of the class and the results are typically used more for administrative purposes than for developmental purposes. Instructor evaluations may be more beneficial to both the instructors and the students if they are used as a communication tool by which the quality of the service encounter can be improved. This two-way benefit is realized by increasing the human interaction component that is necessary before determination of a satisfactory or unsatisfactory service encounter can ultimately occur (Bitner, Booms, & Tetreault 1990).

The timely receipt of information could potentially aid the instructor in improving their overall ability to convey knowledge to their students. This would be far superior to the current reactive approach to instructor evaluations because evaluation information could be conveyed to the instructor in time for corrective action to be applied to that particular class, thus better accommodating current students. The CIT allows for this process by providing detailed descriptions about specific events and/or behaviors. This type of student/teacher interaction allows the student evaluator to offer suggestions for improvement, as well as, granting the instructor sufficient information with which to perform a self-evaluation aimed at teacher improvement.

## RESEARCH METHODOLOGY

The CIT is a qualitative technique consisting of five steps aimed at producing a complete and detailed picture of the subject under study. These five steps are: 1) determination of the general aim of the activity, 2) development of a plan for data collection, 3) collection of the data, 4) analysis of the data, and 5) interpretation of the results (Woolsey 1986).

This study employed the CIT to evaluate the quality of instruction provided in a principles of marketing class. The class was taught by an assistant professor of marketing. The sample size of the experiment included all registered students in the course, 42 subjects in total. Students served as the evaluators and were given two preprinted forms on which to record their observations. The evaluators were told to observe characteristics of the instructor and identify certain incidents throughout the class that they felt demonstrated particularly effective or ineffective teaching. In describing the incidents, students were told to include such things as actual lecture techniques, personal mannerisms, or even methods of class organization; anything which contributed to the effectiveness/ineffectiveness of the incident. After ten class sessions, each student evaluator was required to turn in the effective and ineffective evaluations with a cover sheet containing their name and social security number. The cover sheets were later detached from the evaluators for two reasons: 1) to assure evaluator anonymity, and 2) to allow the student evaluators to be given credit for completion of the assignment.

**Step One.** In this study, the general aim of the activity was to identify behaviors or characteristics which can be used in the training and development of university instructors. The goals of other studies could very well be different for each instructor and/or class, depending on which attributes of teaching the instructor wishes to examine. In this study the main goals were to obtain feedback from students that would help in the training of relatively new professors, or aid in the refinement of teaching techniques for more experienced professors. It was felt that if this goal was to be properly achieved, it would be necessary for the professor to obtain feedback on both their effective and ineffective teaching techniques.

**Step Two.** The Planning process for data collection requires the answers to several questions including who will be the evaluator(s)/observers and what will be included in the content of the evaluation(s)? As mentioned before, students were selected as the evaluators and the content of the evaluations was organized by using prepared evaluation forms (See Exhibit A). These forms were designed to aid the evaluator(s) in effectively describing a particular incident deemed critical by the evaluator(s). Two forms were designed and distributed, one specifically for the identification of an effective teaching technique, and one specifically for the identification of an ineffective teaching technique.

The effective evaluation form contained three questions in total. The first question attempted to focus the evaluator on one very specific incident and the related events surrounding the incident. The second question encouraged the evaluator to identify the actual teaching traits or behaviors that made the incident effective. Responses to this question were of primary interest to the independent judges charged with categorizing and sub-categorizing the evaluations in step four of the process. The third and final question attempted to get the evaluator to describe the actual result of the effective teaching trait or behavior so consistency could be established with the study's goals.

The ineffective evaluations were configured in a similar manner to the effective evaluations except that they contained an extra question. The additional question was added to increase the interaction between student and instructor by allowing the evaluator to offer their suggestions on how to improve upon an ineffective trait or behavior displayed by the instructor.

**Step Three.** The collection of the data was fairly simple. The professor conducting the class passed out the preprinted forms, gave students ten 75-minute lecture sessions to observe and evaluate the instructor, and then had a colleague pick up the anonymous evaluations. The instructor helped to guarantee anonymity by requiring each participating student to hand in a type written evaluation and a detachable cover sheet. The cover sheet was then separated from the evaluations by the independent colleague assisting in the collection process. Additionally,

while there is no specific time period requirement underlying the CIT, the time given to the evaluators should be sufficient for them to experience many diverse teaching attributes.

**Step Four.** The data analysis section of the experiment is often the most difficult part of the process (Woolsey 1986). This part of the process centers around separating the evaluations into primary and secondary categories. Since one of the goals of the research was to provide information for the development of relatively new professors (i.e. graduate student teachers), it was felt that the evaluations should be divided into two categories. One of these categories was labeled "teaching" and contained evaluations that identified techniques that could be taught to the instructor. The other category was labeled "teacher" and contained evaluations identifying attributes that were more innate and thus may be more difficult to teach the instructor (i.e. personality traits). Once the evaluations were divided into these two groups, they were further sub-categorized into more specific groups.

Once the researcher has completed the categorization and sub-categorization process, an independent judge is given the titles of the multiple groups and is asked to independently attempt their own categorization of the evaluations. To enhance the validity and reliability of the results, it is suggested that you obtain a 75% to 85% agreement rate with your primary categories ("teacher" and "teaching" in our study), and a 60% to 70% agreement rate with your sub-categories (seven segments for each of the effective and ineffective evaluations in our study) (Anderson & Nilsson 1964).

**Step five.** This step simply includes analyzing the results and drawing appropriate conclusions given the goals and objectives of your study.

## RESULTS OF STUDY

### Effective Evaluation Results

"teacher" category 84%		"teaching" category 77%	
Interaction	84%	Content	77%
Physical/Energy	83%	Lecture Examples	85%
Ability to deliver	78%	Organization	66%
		Explanations	94%

### Ineffective Evaluation Results

"teacher" category 92%		"teaching" category 81%	
Tone/Pace of Lecture	76%	Course requirements	83%
Energy Level	78%	Material covered	76%
Appearance	89%	Organization	88%
		Explanations	96%

## DISCUSSION

Examination of the effective evaluations identified the following sub-categories within the "teaching" category: 1) class content (i.e. "the topics of discussion are very interesting and informative"), 2) examples used in lecture to illustrate course material (i.e. "the professor uses lots of real world business examples to illustrate the theory of the course"), 3) overall organization of the course (i.e. "I like the way the professor provides the students with an outline of the lecture for that day"), and 4) explanation of material contained within the lecture (i.e. "the professor is good at explaining the theory of marketing"). The effective evaluation sub-categories contained within the "teacher" category included: 1) interaction with the students (i.e. "the professor encourages class participation by involving the students in the class"), 2) energy level or physical characteristics displayed during the lecture (i.e. "the professor is very enthusiastic and appears to enjoy lecturing to the students"), and 3) overall ability to deliver an effective lecture (i.e. "the professor is an excellent lecturer and is good at public speaking").

The effective evaluations categorized under "teacher" were dominated by one sub-category.

The most consistently identified sub-category was the use of "lecture examples" as an effective technique. The effective evaluations categorized under "teacher" also consistently displayed one sub-category; "physical/energy level."

Examination of the ineffective evaluations identified the following sub-categories within the "teaching" category: 1) assignments or requirements of the course (i.e. "I didn't like the way the professor required us to use the mainframe computer for class assignments"), 2) material covered within the course (i.e. "I felt that much of the material covered in the course will be of no benefit to me in the future"), 3) organization of the course (i.e. "I didn't like the way the instructor often deviated from the book during the lecture," and 4) explanations contained within the lecture (i.e. "I had trouble following some of the in-class explanations of marketing"). The ineffective evaluation sub-categories contained within the "teacher" category included: 1) tone and/or pace of the lecture (i.e. "our enthusiastic professor sometimes went too fast and I was unable to take good notes"), 2) energy level of the lecturer (i.e. "the professor was constantly moving around the room which made me nervous"), and 3) overall appearance of the lecturer (i.e. "sometimes the professor was dressed in such a way that I concentrated on her instead of what she was saying").

In reviewing the ineffective evaluations it was once again apparent that the majority of the student evaluators were consistently identifying the same sub-categories. The majority of the evaluations falling under "teaching" identified the sub-category of "course requirements" as the predominant ineffective characteristic. The majority of the evaluations falling under "teacher" identified the sub-category of "tone and/or pace" as the ineffective characteristic. At this point, the professor might be well advised to review question four of the ineffective evaluations to obtain potential solutions to the characteristics identified as being ineffective. It is interesting to note at this point that some of the sub-categories appear as both an effective and ineffective trait (i.e. energy level of the lecturer). This indicates that the professor needs to achieve a balance with this trait so that it is received by the students as being effective.

## LIMITATIONS & CONCLUSION

It is important to note that during the self-evaluation stage of this technique, the professor must interpret the results in a manner that allows for classroom changes while maintaining the overall integrity of the classroom situation. For instance in the ineffective results contained above it may be that many of the student evaluators simply do not like the workload required of the particular class. This would explain the ineffective evaluations under the sub-category of "course requirements," however, a change may not be the prudent response under the circumstances.

Perhaps one of the greatest benefits of this technique could be as a teaching instrument for graduate students relatively new to classroom teaching. Most PhD. programs in the United States utilize their PhD. students as instructors, researchers, or a combination of both. This makes sense in light of the fact that most PhD. students have a desire to continue in the academic arena where they will primarily be utilized as researchers and classroom instructors. Using the CIT as an aid in the improvement of graduate student teaching techniques could be very beneficial to current PhD. students since many of them will be the professors of tomorrow in our colleges and universities.

This instrument is by no means perfect, and its use should be confined to a self-improvement instructional tool. It is very likely that this type of method would yield minimum positive results when used as a method of administrative evaluation. However, since teaching is a vital component of the overall success experienced by a college professor, any evaluation technique that is aimed at overall teacher improvement needs to be considered. Using the Critical Incident Technique as a method of instructor evaluation could potentially increase the overall service quality of higher learning institutions here in the United States and abroad.

## EXHIBIT A

### Effective Evaluation Questions

- 1) Please describe a specific situation in which your instructor was particularly effective in his/her teaching. Include class, date, time, and events surrounding the incident.
- 2) Please describe the effective teaching trait, behavior, activity, etc. What did the instructor do?
- 3) Please describe the result of the effective teaching. Specifically, how did this teaching activity or trait affect your learning of the material? What happened as a result of the teaching technique that affected your learning experience?

### Ineffective Evaluation Questions

- 1) Please describe a specific situation in which your instructor was particularly ineffective in his/her teaching. Include class, date, time, and events surrounding the incident.
- 2) Please describe the ineffective teaching trait, behavior, activity, etc. What did the instructor do?
- 3) Please describe the result of the ineffective teaching. Specifically, how did this teaching activity or trait affect your learning of the material? What happened as a result of the teaching technique that affected your learning experience?
- 4) Since you have identified an ineffective teaching technique, please give a recommendation on how this problem can be improved.

### REFERENCES

Anderson, Bengt-Erik, and Stig Goran Nilsson (1964). Studies in the Reliability and Validity of the Critical Incident Technique. Journal of Applied Psychology, 48:398-403.

Bitner, Mary Jo, Bernard H. Booms, and Mary Stanfield Tetreault (1990). The Service Encounter: Diagnosing Favorable and Unfavorable Incidents. Journal of Marketing, 54:71-84.

Day, Ralph L. and Muzaffer Bodur (1978). Consumer Response to Dissatisfaction With Services and Intangibles. Advances in Consumer Research, 5:263-272.

Fitts, Paul M. and R.E. Jones (1961). Analysis of Factors Contributing to 460 Pilot-Error Experiences in Operating Aircraft Controls. Selected Papers on Human Factors in the Design and Use of Control Systems.

Flanagan, John C. (1954). The Critical Incident Technique. Psychological Bulletin, 51:327-358.  
George, R. Thomas (1989). Learning by Example: The Critical Incident Technique. Cornell Hotel and Restaurant Quarterly, 30:58-60.

Hertzberg, F., Mansueti, B. and Snyderman, B. (1959). The Motivation to Work, New York: Wiley.

Hoffman, K. Douglas and Scott W. Kelley (1991). The Education Service Encounter: The Satisfaction of Students. Journal of Marketing Education, 13:67-77.

Luthas, Fred (1985), Organizational Behavior. New York: Wiley.

Quech, John A. and Stephen B. Ash (1981). Consumer Satisfaction With Professional Services. Marketing of Services, Chicago: AMA, 82-85.

Steers, Richard M. (1984), Introduction to Organizational Behavior. Dallas: Scott, Foresman and Co.

Woolsey, Lorette K. (1986). The Critical Incident Technique: An Innovative Qualitative Method of Research. Canadian Journal of Counseling, 20:242-254.