DOCUMENT RESUME

ED 107 680 TM 004 494

AUTHOR	Paine, Whiton Stewart
TITLE	Impacts on an Educational Researcher.
PUB DATE	[Apr 74]
NOTE	18p.; Paper presented at the Annual Meeting of the American Educational Research Association (Chicago, Illinois, April 1974); For related documents, see TM 994 490-493; Not available in hard copy due to marginal legibility of original document
EDRS PRICE	MF-$0.76 HC Not Available from EDRS. PLUS POSTAGE
DESCRIPTORS	Bias; Consultants; *Course Evaluation; Educational Attitudes; Educational Researchers; Feedback; Formative Evaluation; Higher Education; *Interdisciplinary Approach; *Interprofessional Relationship; Participant Characteristics; Personal Growth; *Professional Personnel; *Program Evaluation; Role Conflict; Staff Role
IDENTIFIERS	*Center for International Studies; Cornell University

ABSTRACT

Finding himself in the midst of a stressful and ambiguous evaluation process, one evaluator reviewed how he played a variety of professional and interpersonal roles while evaluating problem-oriented, team-taught courses at the Center for International Studies, Cornell University. Attention was given to how various interpersonal factors facilitated and inhibited productive exchange in groups and individuals involved in the evaluation were described. One evaluator's professional growth was discussed. (Author/BJG)

14.14

ED107680

U S DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

IMPACTS ON AN EDUCATIONAL RESEARCHER

Whiton Stewart Paine
Cornell University

So far in this symposium we have described the basic evaluation of the Center for International Studies problem-oriented, team-taught, undergraduate program. Alison Brown and Stephen Brock have discussed how assessing this program has affected them as an Intellectual Historian and teacher and as an Educational Philosopher and Administrator. I was the Educational Researcher, or more exactly, Evaluation Specialist, called in to coordinate the study. For me, most of the impacts of this work have been in the interpersonal and professional areas. These classes of effects were particularly significant to me because of my background and because of the different roles I played within the evaluation.

Prior to coming to Cornell, I had done some work on program evaluation in medical and educational settings, and my thesis and dissertation involved the investigation of introductory courses. My formal academic training spans several areas of Psychology and Education and I also had training and experience in several models of interpersonal relations. But, despite a doctorate awarded in an interdisciplinary program, my experience with International Studies and with interdisciplinary team teaching was limited. One of the reasons I accepted the position in the Center for Improvement of Undergraduate Education was that it represented an opportunity to learn about this approach to college teaching.

BEST COPY AVAILABLE

TM 004 494

As was noted in the Introduction, the main planning for the evaluation took place in a two week period before the Rural Development course began in the Fall of 1972. This short time period was a source of initial worry but not of great concern because of four simplifying assumptions that were made about the task at hand. First, I assumed that the cordiality between the CIS and CIUE administrative personnel extended to the CIS faculty. Second, that the basic requirements of the evaluation had already been codified in the OE proposal. Third, that some of the techniques which had been used in previous studies could be easily adapted to this new situation. And fourth, that the course faculty had already specivied in detail the major goals and objectives of their offering. After all, this was Cornell, and the various individuals I had met with early in the Summer were able to clearly state what it was that they were going to do in the fall.

The first major impact of this study on me was the realization that these assumptions were all, to some degree, false. The result was considerable initial uncertainty enlivened by severe time pressure and by inadequate interpersonal relations both within the evaluation team and between that team and the first group of faculty to be evaluated.

Actually, the ambiguity which has surrounded most aspects of this evaluation have indirectly led to most of the major impacts of the evaluation on me. As a result of continuing uncertainty, I have played a variety of professional and interpersonal roles and each role has offered diverse opportunities for learning.

Background

At first, I was the outside "expert", the evaluation designer who, on occasion, became the defender of the faith against uncooperative faculty and unintelligible administrators. Within the CIS course on Rural Development I began as the nonparticipant observer, tolerated but not seen as useful. Over time this role changed as the faculty of this course and the one on Ethnicity began to use me as an educational consultant. Under these new conditions, I became more a participant in the planning and implementation, but not in the teaching, of these courses.

Simultaneously, I had become the principal investigator, acknowledged coordinator of a range of evaluation activities. These included the design and administration of different assessment techniques, and the selection, training, and supervision of interviewers, video-tape raters and other assistants. At the end of the data collection phase, my role shifted again and I became the digester, the one who looked at numbers and generated preliminary statements to be used by other members of the team as they framed larger generalizations. All of these roles were important sources of interpersonal and professional learning.

Interpersonal Impacts

My previous training and experience had provided me with an awareness of interpersonal factors which facilitate and inhibit productive interchanges in groups. Observing teams of faculty attempting to design and offer interdisciplinary courses and working within an evaluation team has sensitized me to some variables that I had previously tended to underrate. Above all, this experience has taught me a great deal about the importance of basic assumptions and different cognitive styles.

For example, on most of the dimensions used to delineate cognitive styles, Stephen Brock and I differ markedly. He tends to be a lumper; I, a splitter; he starts with an abstract totality; I want to look at concrete subunits and how they relate to each other. We are both convergent thinkers but in very different ways.

In and of themselves these differences were important sources of mutual incomprehension but communication was further complicated by the basic uncertainties surrounding the program to be evaluated. Steve linked the undergraduate program to a larger conception of interdisciplinary team teaching. I had real difficulties in justifying this linkage because the original, fairly explicit, program goals in this area had been superceded; because there was no central coordination of the program, and because I had little evidence that the teams teaching the courses had seriously analyzed the problem of interdisciplinarity as part of their design work for the offerings.

To me, Steve wanted to go on a snark hunt, and I was very unwilling because of my judgment that the snark might turn into a Boojim and disappear, leaving nothing behind but another busted evaluation. My competing strategy was to spread a series of nets and trap a variety of things, including, if we were lucky, a Boojim. This strategy won out at the price of continuing interpersonal difficulties that had to be dealt with later.

The problem here was not the usual one of personality conflict but rather of different world views which had few points of contact and, under stress and uncertainty, each of us stressed a cognitive style that the other found incomprehensible.

Actually, this problem has never been completely resolved. The writing of the final report is a case in point. My role in the writing team is to analyze raw data and generate preliminary conclusions. David Macklin translates these findings into a form which he and Stephen Brock can relate to the model they are building. Their synthesis then goes back to me to recheck the relationships they outline between specific conclusions and recommendations and the original data. This has proven to be an excellent way of using our differential competencies but is a somewhat unusual format for interpersonal relations.

A separate important impediment to effective communication in the evaluation team was caused by different conceptions of a central evaluation process. Since the program goals were not explicit and the available literature in this area is limited, the evaluation team had to develop its own criteria or standards of comparison for the interpretation of the results. Actually, we developed three competing standards. I tended to approach results in terms of an absolute criterion, to establish constraints, and then ask how well the available resources were utilized. Alison Brown, following her mandate to study the context of the program, analyze events in terms of the original program goals, goals which were later largely discarded by CIS.

Stephen Brock and Dave Macklin, drawing upon their familiarity with the general level of teaching at Cornell, applied a hypothetical "average" standard.

These different points of view led to continuing problems and a major breakdown in communication when Alison Brown presented her report on the history of the program to the other members of the evaluation team. The reaction she received was predominantly negative and, until the question of differing standards had been analyzed in some detail, collaboration was impossible.

Academically appropriate standards of comparison were also an inhibiting factor in the work of the teaching teams. Three of these teams were observed directly and follow-up interview data was available on the earlier groups of faculty. Our analysis indicates that disciplinary allegiances were not the main barrier to effective collaboration. Being trained in different disciplines had the minimal effect in this situation of initially limiting communication until some agreement was reached on terminology, and a productive dialogue began.

A more severe problem arose when two or more members of a teaching team had concluded that they both had expertise in the same area. At that point, intellectual discourse had a tendency to turn into interpersonal confrontation and collaboration stopped. It was: "My disciplinary insights are better than yours!"

For example, the course faculty in the first CIS course on Integration and Decentralization began their work by discussing the case country with which they all were familiar -- France. In a somewhat inexorable way, they went from the discovery that they could not agree on this crucial example, to an interdisciplinary effort that was offered to students as four, almost completely separate, disciplinary courses.

A second, and more corrosive example, occurred in the fifth course, on Ethnicity, Race, and Communalism. Of all the CIS courses, this one was characterized by the highest level of interpersonal conflict. The core faculty included three political scientiets, two of whom had considerable experience with Africa. In this course, conflict was immediate, severe, and continuing, with the result that at the end of the semester these two individuals had serious doubts about each other's academic competence.

This aspect of the ethnicity course highlighted an important pedagogical problem that also was present in other CIS courses. Conflict among the faculty did not seem to have a major impact when it occurred only within the staff meetings, but when the war was carried over to the lecture sessions, the student reactions were resoundingly negative. In this teaching situation, the faculty saw confrontation as an acceptable type of intellectual discourse, while students summarized such interchanges with comments like: "They present themselves in such a way that one wonders if they have seen each other since the last lecture", and, in a particularly cutting remark: "Educationally, they [the confrontations] did have a purpose -- to give us insight into the way the supposedly most elite of the world get along with each other". The latter was not a specified course or program goal! In a more serious note, there was some evidence that student detestation for faculty conflict generalized to a negative perception of other aspects of the course.

Finally, before leaving the interpersonal area, I would like to point out some more positive outcomes. Through this project, I have had an opportunity to interact with a variety of faculty and students. Some of these interactions have led to what I hope will be continuing relationships. Alison Brown and her husband, Jay, and one member of the CIS faculty have become personal friends and I expect that other, more professionally oriented relationships will also continue into the future.

Professional Impacts

In the last two years I have played four general professional roles -- program evaluator with this project and, outside this study, college instructor, educational researcher, and instructional development consultant. Working on the CIS evaluation has aided me in discharging many of the responsibilities of all these roles and should aid me further as I play them in the future.

First, this evaluation has been very useful to me as a teacher interested in the problems of teaching introductory courses in the social and behavioral sciences. In terms of the teaching technologies employed, these were fairly conventional introductory offerings. They all stressed a reading list, weekly lecture sessions, discussion sections, and a grading system. By observing how others use these course components and how students reacted to different procedures, I learned a great deal about the benefits and limitations of these standard procedures. In particular, this learning has been useful in my own teaching and in some instructional development activities.

Perhaps the most important conclusion was that the experiences in the CIS course generally validate the importance of the basic teaching skills stressed by McKeachie and others. Many of the difficulties experienced within the CIS courses were the result of rudimentary teaching errors. Yet, in an atmosphere characterized by initially high student and faculty expectations, even minor errors sometimes had major negative consequences.

Some of the things I have learned are being applied in a course I teach now, and other learning will affect future offerings. In addition, after watching others struggle with the problems of interdisciplinary team teaching, one result is a personal desire to participate in this type of educational experience. While I am more aware of the difficulties, evaluating this program has reinforced my conviction that this type of effort is crucial in training individuals to deal with larger problems.

Secondly, many of my activities at Cornell fall under the general rubric of "Instructional Development". The in-depth study of individual courses has been directly useful to me when dealing with faculty and graduate students interested in increasing their instructional skills. In particular, the CIS evaluation has been a fertile source of illustrative examples of effective and ineffective instrution.

Also, as my role changed within the evaluation, I became more aware of the importance of different types of feedback. Technical, summary, and process feedback were used extensively within the study. Technical feedback, generalizations drawn primarily from the literature was often effective, particularly in the design of examination systems. Summary feedback, generalizations and incidents drawn from earlier CIS courses, was useful when the course faculty was attempting to make decisions between competing options. Finally, during the semester a course was offered, some immediate process feedback was given. This last class of feedback was far less useful than expected. First, each course developed a sort of inertia, with the momentum of the original design making it difficult to institute mid-semester changes. Also, a common pattern in these courses was an interpersonal breakdown in the teaching team which also made it increasingly difficult to make decisions after the course began.

Testing the limits and values of these different types of feedback also was an aid in developing consultation skills. The Cornell faculty, as a group, are perhaps more difficult than average to work with because of the competitive nature of the institution and the strong individual commitment to personal competencies discussed earlier in this presentation. Under these conditions, the hardest lesson to learn has been when to speak and when to keep quiet.

I have also become quite impressed with the way in which the freedom to innovate can be eroded by political decisions. The CIS program was one child of the events of 1969, when universities smelled of tear gas and Cornell's Blacks sported guns to beat off attacks by irate fraternity men. Since that time, there has been a conservative counter-reformation among the faculty and this program has been one of the victims. One sign of this has been the accomodation CIS has made to the increasingly trivial demands of individual faculty members.

Two recent examples illustrate this process. While the course on Ethnicity, Race, and Communalism was being planned, a faculty member outside CIS complained that the topic infringed upon his course on racism in America. A comparison of the two courses reading lists showed that only one citation out of some 150 sources, was common to both. But, with reluctance, the course faculty decided to remove all American materials from the syllabus. Politically, this was a wise decision; pedagogically and intellectually it cannot be as easily defended. The students were, by and large, Americans, and if they had any familiarity with the course topic, it was within the American context. Also the bulk of the available research on the topics covered has been done by American writers dealing with the American situation.

More recently, there was a desire to offer a course on the world food crisis. This idea was dropped because the teacher of a graduate level agriculture course on this topic indicated that he was considering allowing some undergraduates into his offering. There are other examples but they raise the issue that perhaps extraordinary procedures to protect academic freedom are necessary to protect instructional innovations over a period of time.

As a side note here, one of the interesting limitations on technical feedback involved the specification of course goals and objectives. On their own, the course faculty were quite reluctant to specify, in detail, the expected impacts of the course on students, and the relationships between teaching techniques and those impacts. There was some flavor of "That's what they do [illegible] Education Department and it's not appropriate here". However, if the evaluator suggested a goal, the response was frequently surprised agreement and some relief that, indeed, this was what was occurring in the course. This clarification procedure was dropped because of my strong suspicion that I was creating a type of pseudo-reality out of nothingness. This would seem to be an important trap to be aware of in a curriculum evaluation with conventional faculty.

Thirdly, some aspects of the course were directly relevant to my interests as an educational researcher. These interests include relationships between discussion group dynamics and cognitive and affective learning. Because of political constraints and of the way in which the CIS courses were organized, the discussion sections provided one of the few opportunities for comparative studies within this evaluation.

The two courses analyzed in depth both had discussion sections, and the process observations indicated that the section leaders differed markedly in their styles, and the students differed in their reactions to the sections. Further validation of these observations was found on the evaluation forms. These differences also related to a question of interest to CIS, namely whether or not there was a significant difference between sections led by faculty and those led by graduate students.

In both courses, membership in a discussion section had a significant effect on ratings of the course as a whole. Additionally, in the Rural Development course, section membership had a significant effect on student performance on the course examinations designed by the faculty. A similar pattern occurred in the next course but high variability and low sample sizes resulted in a finding of no significant differences. However, in this latter course, section membership did affect some of the students' racial and ethnic attitudes.

Now, for those of you familiar with previous research in this area, the finding that discussion sections do make a difference on cognitive performance, perceptions, and attitudes comes as no surprise. However, it is worthwhile to demonstrate these linkages in instructional situations which are not set up for such a demonstration. Also, this demonstration was useful in an instructional development sense as well. It was illuminating to the faculty to see that discussion sections affected performance on their examinations and had an impact on the ratings of their course. In particular, the finding that sections led by graduate students were superior in both courses was something of a shock to the faculty. It may be that one of the primary services evaluators can offer is to help individuals rediscover the wheel within their own particular programs.

Finally, as an evaluator, the first major outcome of this work was the feeling of mixed wonder and satisfaction that was connected with the realization that the evaluation was over, that some of the outcomes hoped for had materialized and many of the disasters expected never occurred. Also, as in any professional activity, there were learnings about tools, about classroom observation techniques, interviewing, item construction, and the like. However, it is two broader impacts that I would like to talk about this afternoon. These are my present perceptions of some issues about professionalism and some thoughts on the CIPP model.

Throughout this paper there has been a stress on the ambiguity that characterized the evaluation. At the beginning, an additional important source of ambiguity was my perception of conflicting responsibilities. No one, including myself, was clear about the role of an evaluation specialist in this situation. Personally, I felt very much the focus of competing constituencies. Somehow I was to meet the expectations of CIUE who employed me; of the U.S. Office of Education, who paid me; of CIS, whose program it was and who also provided support; of the faculty in each course; of students; and, in some sense, of the larger evaluation community and Cornell University. Further complicating this situation were my developing relationships with the individuals at this table.

Now this is not an uncommon situation in evaluation work although this particular study may have been more complex than is usually the case. What concerns me now is my own retreat into a sort of "professionalism" under these conditions.

For example, in the stressful design period, I frequently overemphasized the original OE proposal and used it to ward off Steve

Brock's demand for an educational research study on the structure of knowledge. Similarly, it was a useful club when faculty cooperation was minimal. One of the turning points in the early stages of the evaluation occurred when I concluded that the requirements of the OE proposal could not be met under these conditions and proposed to the Associate Director of CIS that the money be refunded immediately. Working conditions improved markedly after the uproar subsided.

Levity aside, the issue here is central to the practice of evaluation. It is the question of what is the appropriate, the professional response when external constraints make a creditable job impossible. Admittedly, evaluators are frequently in the business of making lemonade out of lemons, but there are occasions when we must maintain credability by refusing to participate in the evaluation equivalent of a no-win game. One danger, and an obvious one in this evaluation, was giving up too early, but I still feel that if more cooperation had not occurred, either the proposal had to be renegotiated with OE, or the money returned. However, it is clear to me that this is an area where there is a need for guidelines, for a canon of professional practice to guide evaluators.

The most obvious area of learning concerns the CIPP (Context-Input-Process-Product) model. It was the perfect tool for organizing the collection of data in this situation. Because there was so much ambiguity within the program, this type of general model, with its flexibility and stress on decision-making seemed particularly appropriate. At least it was appropriate to me; Stephen Brock has already given you a different point of view. Also, at the last Christmas party, the staff gave me a very large, and old, and battered garbage pai labeled, GIGO, an obvious reference to CIPP.

Within the evaluation, the CIPP approach has been quite valuable as a strategy for defining the areas of the program to be assessed. In particular, Alison Brown's work is a direct outgrowth of this model's emphasis on the importance of Context information. CIPP accmmmodated a wide range of evaluation techniques and was a very useful tool for deciding on different data collection procedures and for grouping data from different sources. Using CIPP this way has reinforced my intellectual conviction that, in a complex situation, evaluators must look at many things in order to generate a holistic perception of the program studied. In addition, the concept of multiple operationalism, of comparing different weak sources of data on the same phenomena was productive in this instance and seems to be justified for this type of study.

Also, a surprising Process outcome from the use of the CIPP model was the one discussed by Steve. On the basis of my analysis of the data collected, we both got some of what we wanted. I was able to make specific recommendations on how the courses should be taught in the future to increase student and faculty satisfaction and learning. Steve, in turn, was able to use much of the data in an analysis of the interdisciplinary, team-teaching nature of the program.

However, it was at the Output end that CIPP did not meet our needs. The results of this evaluation must be communicated in a format which is appropriate to a wide range of interested readers, including OE personnel, CIS administrators, evaluators, Cornell faculty, and other teachers interested in International Studies. We first attempted to present the results in terms of the Context, Input, Process, and Product categories but rapidly discovered that the chapters were not communicating what we wished to communicate. Therefore, with much travail over the last six weeks, we were forced to develop a more integrated and intelligible format for our report. The next time I use CIPP, I expect to spend a great deal of time on this outline on order to avoid some major problems with the final report.

Finally, to summarize, it has been useful to me to reconceptualize the evaluation as a laboratory within which I could observe myself and others interact and work in a stressful and ambiguous situation. By looking at the last two years in this way, I can specify some of the interpersonal and professional impacts of the evaluation upon myself. Many of these impacts have been positive and in part this reflects the activities of our next speaker, David Macklin, who will discuss the evaluation from the perspective of Consultant and Social Psychologist.