THE USE OF CIPP IN THE EXTENSIVE LISTENING COURSE AT PSDKU PROGRAM

Renata C. G. V. Nikijuluw*
FKIP Universitas Pattimura, Indonesia
Corresponding E-mail: renataloppies@gmail.com

Abstract: This article includes similar readings on educational assessment. Specifically, it includes relevant reading papers on the essence of the assessment, its intent, and the various evaluation models and their advocates. It also provides similar readings on a separate assessment model, the CIPP (Context Input Process Product) model. Justifying the reasoning for the use of the CIPP model, this involves readings on the context, emphasis, strengths and shortcomings of the CIPP and its proponent. It also provides literature on the approaches used to gather soft data, such as questionnaires, interviews and the focus group debate.

Keyword: CIPP and educational evaluation

DISCUSSION

The Nature of Evaluation

Evaluation in many forms has been conducted to respond to the fast pacing changes and demands in education. Actions may have been different from the previous solutions since evaluation needs to answer to the particular needs and context of the program. As Davis (1980 cited by Hodgkinson 2005 p.1.2) writes, "most evaluations are situational or problem-specific that is each evaluation requires particular and appropriately responsive treatment". Moreover, the evaluation focuses on obtaining specific information relevant to a particular problem, program, or product (Worthern & Sanders 1973 p. 1.2). Assessment is related to evaluation since the former refers to the collection and organization of data to report on individual performance. Such is essential in improving program performance which is the concern of evaluation. Haywood (1977 cited in Hodgkinson 2005) points out that the educational process's improvement is through an adequate evaluation of learning with some form of assessment based on specific objectives and learning strategies. Evaluation only happens when the collected and reported organized data are used to decide a particular aspect of a program. Thus, Groundwater-Smith and Nicol (1980 cited in Hodgkinson 2005 p. 1.4) emphasize that "the ultimate goal of evaluation in education is to enable administrators or teachers to make informed decisions or choices between alternative strategies for action." Hence, judgment is a very important process in evaluation.

The Purpose of Evaluation

Evaluation is conducted to determine the worth of a program and/or the change made. The worth of an action is based on "determining the effectiveness of the various schemes in meeting a set of criteria" (Hodgkinson 2005 p. 1.5). Northern and Sanders (1973 cited in Hodgkinson 2005 p. 1.5) say that

"evaluation is the determination of the worth of a thing." Thus, essential information has to be collected to "judge the worth of a program, product, procedure or objective, or the potential utility of alternative approaches designed to attain specified objectives" (Worthern and Sanders 1973 cited in Hodgkinson 2005 p. 1.5). The evaluation also facilitates informed decision-making, therefore data that have to be gathered are valid, well-organized, and relevant (Hodgkinson 2005). Evaluation only occurs when these reported data are used to facilitate decision-making and to determine the worth of the program or an aspect of the program.

Models of Evaluation

Different educational programs serve different purposes, clients, and beneficiaries. To improve its services and strategies, it continuously evaluates its performance, impact, and effectiveness. These programs are also situated in unique contexts, with their own desired goals, a particular setting, and specific clients, such that, a particular type of evaluation would have to be carefully chosen to serve appropriately its purpose. The first view of evaluation is based on Provus' discrepancy model. The model involves the "blending of evaluation and management theory" (Groundwater-Smith & Nicol 1980 p. 4). Looking into the decision making, it considers the choices being made whether the program has to be improved, maintained, or terminated. The processes involved knowing the program goals and standards, determining the existence of discrepancy between some aspect of the program and standards, and using this discrepancy data to identify the weaknesses in the program (Groundwater-Smith & Nicol 1980).

Second is Stake's judgmental model whose emphasis is on the judges themselves. The evaluator is required to recognize the decision-makers and to see the purpose of "collecting and collating information as being expository" (Groundwater-Smith & Nicol 1980 p. 5). The task of the evaluator is to get information from the planning stages, to look for consistency between "what is being said to be done and what is being done" and to report the findings in such a way that the decision-makers may arrive at their conclusions and make their own decisions (Groundwater-Smith & Nicol 1980).

Third is Stufflebeam's decision-making model which is popularly known as the CIPP model. It is based on his definition of evaluation as "the process of delineating, obtaining and providing useful information for judging decision alternatives" (Stufflebeam 1969; 1971 cited in Groundwater-Smith & Nicol 1980 p. 3). For this evaluation of a course in an on-campus mode, this model was chosen by the evaluation team.

What is CIPP?
The CIPP model of the evaluation was developed by Daniel Stufflebeam and associates in the 1960s, based on their experience of evaluating education projects for the Ohio Public Schools District (Harold & McKee, 2003). CIPP is an acronym for **C**ontext, **I**nput, **P**rocess, and **P**roduct, requiring the evaluation of context, input, process, and product in judging a program's value (Harold & Mckee, 2003). The CIPP Model is a comprehensive framework for guiding formative and summative

evaluations of projects, programs, personnel, products, institutions, and systems. It is used in internal evaluations conducted by an organization's evaluators, self-evaluations conducted by project teams or individual service providers, and contracted or mandated external evaluations.

CIPP is a decision-focused approach to evaluation and emphasizes the systematic provision of information for program management and operation (Robinson, 2002). The information is considered valuable when it helps and informs program managers to make better decisions. It maximizes the effectiveness of critical decisions through the timely reporting of relevant information in a useful form to appropriate levels of decision making. Thus, it is considered as the combination of effective decisions based on timely and relevant information. The system focuses on four classes of decisions and is designed to yield four kinds of information to serve those decision situations. Evaluation activities should be planned to coordinate with the decision needs of program staff. Collection and reporting of data collection are also conducted for more effective program management. Along the process of program implementation, some changes occur and decision makers ' needs will change, thus evaluation activities have to adapt to meet these changing needs. It has to be ensured that there is a continuity of focus to trace development and performance over time (Harold and McKee, 2003). The CIPP framework links evaluation with program decision-making. It intends to offer an analytic and rational basis for program decision-making, based on a cycle of planning, structuring, implementing, reviewing, and revising decisions, each examined through a different aspect of evaluation –context, input, process, and product evaluation (Harold & McKee, 2003). It attempts to make evaluations directly relevant to the needs of decision-makers during the different phases and activities of a program. It also addresses questions asked by the key decision-makers, in ways and language that decision-makers will easily understand. It aims to increase the likelihood of the evaluation findings having relevance and being utilized.

The CIPP
CIPP is an acronym for a series of four evaluations: Context, Input, Process, and Product (Stufflebeam, 1960 cited by Harold & Mckee, 2003). Context Evaluation determines the identity of the target population, defining the parameters of the organization. What is the target population? What needs must be satisfied? It also assesses the environment where the evaluation takes place. It answers the question, "How does the organization's setting facilitate its mission?" It assesses needs, problems, assets, and opportunities to help decision-makers define goals and priorities and help the broader group of users judge goals, priorities, and outcomes. Input Evaluation is the assessment of resources and challenges, alternative approaches, competing for action plans, staffing plans, and budgets for their feasibility and potential cost-effectiveness to meet targeted needs and achieve goals for program success. This is an opportunity to assess potential problems to successful implementation (for example, political climate, market conditions). This is also the stage where we consider different options for the

implementation of problem-solving actions. Shinkfield and Stufflebeam (1995) propose dividing the sponsors into two normally competing teams and letting them vie to design improvements. The competitive nature of these teams might increase the involvement in the evaluation and the organization could choose the best components of each team's plan. It is important to build consensus among stakeholders for the changes considered for implementation; few sweeping changes are unanimously applauded. Decision-makers use input evaluations in choosing among competing plans, writing funding proposals, allocating resources, assigning staff, scheduling work, and ultimately in helping others judge an effort's plans and budget.

Process Evaluation deals with the possible need to restructure the program after the results of pilot testing and previous evaluations are in. It is an evaluation of the day-to-day operations of the organization. This is done after the program plan has been initially implemented. This might involve a management information system that tracks the daily happenings of the organization. What does each segment of the organization do that is facilitative to the other components? What do we do to hamper each other? What is the cost/benefit ratio of each of the most costly activities? A management information system could track such data before the implementation of program improvements for comparison after the implementation. Product Evaluation is an assessment of the effects of the program. Consumers would be assessed for their satisfaction with the products and/or services of the organization. Impact assessment can be done with a

fair degree of scientific rigor, typically employing qualitative methods. It is usually acceptable to assert some degree of causality between program and product when it appears there are no other reasonable causes (Hadley & Mitchell 1995). It identifies and assesses outcomes—intended and unintended, short term and long term—to help staff keep an enterprise focused on achieving important outcomes and ultimately to help the broader group of users gauge the effort's success in meeting targeted needs.

In the formative case—where evaluation helps guide an effort—context, input, process, and product evaluations respectively ask: What needs to be done? How should it be done? Is it being done? Is it succeeding? The evaluator submits interim reports addressing these questions to keep stakeholders informed about findings, help guide decision making, and strengthen staff work (Harold & Mckee 2003). If problems remain, or solutions need fine-tuning, it is important to reexamine the process (do another process evaluation). This will help determine whether the solutions are being implemented as intended, or whether there is a basic flaw in the program itself. The product evaluation could determine whether the program should be modified, fine-tuned, or terminated (Rossi & Freeman 1993).

There are many different definitions of evaluation, but one which reflects the CIPP approach is the following:

> 'Programme evaluation is the systematic collection of information about the activities, characteristics, and outcome of

programs for use by specific people to reduce uncertainties, improve effectiveness, and make decisions concerning what those programs are doing and affecting' (Patton 1986 p. 14).

The four aspects of CIPP evaluation (context, input, process, and outputs) assist a decision-maker to answer four basic questions (Harold and McKee 2003):

1. *What should we do?*

This involves collecting and analyzing needs assessment data to determine goals, priorities, and objectives. For example, a context evaluation of a literacy program might involve an analysis of its existing objectives, achievement test scores, staff concerns (general and particular), policies, and plans and community concerns, perceptions or attitudes, and needs.

2. *How should we do it*?

This involves the steps and resources needed to meet the new goals and objectives and might include identifying successful external programs and materials as well as gathering information

3. *Are we doing it as planned?*

This provides decision-makers with information about how well the program is being implemented. By continuously monitoring the program, decision-makers learn such things as how well it is following the plans and guidelines, conflicts arising, staff support and morale, strengths and weaknesses of materials, delivery, and budgeting problems.

4. *Did the program work*?

By measuring the actual outcomes and comparing them to the anticipated outcomes, decision-makers are better able to decide if the program should be continued, modified, or dropped altogether. This is the essence of product evaluation.

The CIPP model may combine quantitative and qualitative methods of inquiry from the social sciences such as questionnaires, interviews, focus group discussions, content analysis of documents and learning materials, analysis of records and databases, observation of sites and processes, literature search and analysis.

The Focus of the CIPP Model
The CIPP Model emphasizes that the evaluation's most important purpose is not to prove but to improve. Evaluation is therefore conceived primarily as a functional activity oriented in the long run to stimulating, aiding, and abetting efforts to strengthen and improve enterprises. However, the model also posits that some programs or other services will prove unworthy of attempts to improve them and should be terminated. By helping stop unneeded, corrupt, or hopelessly flawed efforts, evaluations serve an improvement function through assisting organizations to free resources and time for worthy enterprises (Harold & Mckee 2003).

Strengths and Limitations of CIPP
With its focus on decision-making, CIPP aims to ensure that its findings are used by the decision-makers. It takes a holistic approach to evaluation, aiming to draw a broad picture of understanding a project, its context, and the processes

at work. It helps shape improvements while the project is in the process (Harold & McKee 2003). However, some critics of CIPP have said that it holds an idealized notion of what the process should be rather than its actuality and is too top-down or managerial in approach, depending on an idea of rational management rather than recognizing its messy reality. In practice, the informative relationship between evaluation and decision-making has proved difficult to achieve and perhaps does not take into account sufficiently the politics of decision-making within and between organizations (Robinson 2002).

These argue that all stakeholders have a right to be consulted about concerns and issues and to receive reports which respond to their information needs, however in practice, it can be difficult to serve or prioritize the needs of a wide range of stakeholders. In stakeholder and participative approaches, evaluation is seen as a service to all involved in contrast to the administrative approach (such as CIPP), where the focus is on rational management and the linkage is between researchers and managers or decision-makers. In the stakeholder approach, decisions emerge through a process of accommodation (or democracy based on pluralism and the diffusion of power). So the shift in this type of approach is from decision-maker to audience. Cronbach (1982) argues that the evaluator's mission is to facilitate a democratic, pluralist process by enlightening all the participants'. However, some of the agencies receiving the reports from participative evaluations say that the reports are sometimes not helpful in decision-making, because they

lack clear indications for decision-making or conflicting conclusions.

Why the CIPP Model of Evaluation?
The CIPP evaluation model has general goals and standards. It includes investigations, questionnaires, visitations, interviews, and natural random sampling techniques. The objective is to determine the efficiency or effectiveness of the entire program or any of its components. According to Stufflebeam and Shinkield (1990), the four evaluation types of *context*, *input*, *process*, and *product* have their unique individual functions, and collectively form a complete evaluation model. Finedlay (cited in Kang 1994) thinks that the CIPP model was a good self-evaluation system thus, applied it to the evaluation of planning and organization for Ohio State University's Occupational and Technical Education Center. Kang (1994) utilized it to evaluate the accountability component of Taiwan Occupational Training as well. The CIPP model is more complete as it allows for formative, summative, and self-evaluation. Further, it is very straightforward, comprehensive, and easier to operate.

The CIPP model's epistemological orientation is objectivist rather than relativist. Objectivist evaluations are based on the theory that moral good is objective and independent of personal or human feelings (Harold & Mckee 2003). Such evaluations are firmly grounded on the following: ethical principles, strive to control bias, prejudice, and conflicts of interest in conducting assessments and reaching conclusions; invoke and justify appropriate and standards of technical merit; obtain and validate findings from multiple sources; search

for best answers, although these may be difficult to find; set forth and justify best available conclusions about the value and; report findings honestly, fairly, and as circumspectly as necessary to all right-to-know audiences; subject the evaluation process and findings to independent assessments against pertinent standards; and identify needs for further investigation. Fundamentally, objectivist evaluations are intended, over time, to lead to correct conclusions—not correct or incorrect relative to an evaluator's or other party's predilections, position, preferences, standing, or point of view. The CIPP model contends that when different objectivist evaluations are focused on the same object in a given set when they are keyed to fundamental principles of a free and just society and agreed-upon criteria when they meaningfully engage all stakeholder groups in the quest for answers, and when they conform to the evaluation field's standards, different and competent evaluators will arrive at fundamentally equivalent and defensible conclusions (Harold & Mckee 2003).

The CIPP model requires the engagement of multiple perspectives, the use of a wide range of qualitative and quantitative methods, and triangulation procedures to assess and interpret a multiplicity of information (Harold & Mckee 2003). The evaluator has to be resourceful in compiling a wide range of reasonably good information that in the aggregate tells a consistent and truthful story. The model advocates engaging multiple observers and informants with different perspectives; constructing "homemade" instruments as needed; mining and using extant pertinent information; addressing each evaluation question promptly; using multiple

procedures; cross-checking qualitative and quantitative findings; building a compelling case over time; and subjecting the evaluation to review by stakeholder groups and independent parties. In following this advice, evaluators are expected to search out and investigate ambiguities and convergence and contradictions in findings, listen to and weigh feedback from the program's stakeholders, and be appropriately circumspect in generating and reporting conclusions (Harold & Mckee 2003). Collectively the team should possess such competencies as knowledge of the pertinent subjective matter, planning, negotiation and contracting, leading groups, organizing and administering team efforts, using technology, interviewing, surveying, testing, quantitative and qualitative analysis, cost analysis, effective writing, and effective oral communication (Harold & Mckee 2003).

The CIPP model treats evaluation as an essential concomitant of improvement and accountability within a framework of appropriate values and a quest for clear and unambiguous answers. It responds to the reality that evaluations of innovative, evolving efforts typically cannot employ controlled, randomized experiments or work from published evaluation instruments—both of which yield far too little information anyway. Developers and service providers should validate the goals' consistency with sound values and responsiveness to beneficiaries' needs. They should plan effectively and invest their time and resources wisely. They should earn continued respect and support while responsibly carrying out their plans and producing beneficial results build on past experiences. Lastly, they should

convince consumers to buy or support their services and products by claiming valid and honestly reported services. Finally, Harold and Mckee (2003) disclose that the CIPP model employs multiple methods based on a wide range of applications keyed to professional standards for evaluations, supported by extensive literature and is buttressed by practical procedures.

DATA COLLECTION

In this case study, soft data were collected to determine the effectiveness of EDU5212 on-campus mode as perceived by the students. Soft data collection was preferred over hard data because the latter may not be sensitive to many of the issues which the evaluation team wishes or needs to explore (Andrews 2006). Furthermore, according to Andrews (2006), soft data collection allows the evaluator to determine what the respondents are thinking or about attitudes and values where the evaluator will not be able to predict the type of question which needs to be asked. The soft data collection techniques used in this case study were questionnaires, interviews, and focus group discussion. The literature on these techniques will be presented in this section.

Questionnaires

A 'questionnaire' has been defined as simply a list of questions contained in a form(s) that can be completed by respondents as individuals (Andrews 2006). The main advantage is the use of the questionnaire is the individual completion of the instrument even in the absence of the evaluator. Hence, it could facilitate the data gathering process. On the other hand, however, its major limitations relate to the inability of some closed questions (i.e., those which allow limited responses) to relate to the major issues (Andrews 2006). To eliminate the possibility of ambiguities, Andrews (2006) mentioned that care should be taken in its structuring. To ensure that no ambiguities are built into the questions asked as well as that the questions incorporated cover the area to be evaluated a 'Table of Specifications' should be employed. Moreover, it has also been considered appropriate to employ a limited number of question types in the questionnaire. Hence, it is considered more appropriate to limit the questions to one or two types.

The Likert Technique in the Preparation of Questionnaire. The Likert technique is a scheme developed by R.A. Likert to assess the individuals' affective states (Popham, 1993). Anderson (1981 in Popham 1993, p. 166) presented the eight steps in constructing a Likert scale. These eight steps are:

Step 1. *Write or select statements that are either favorable or unfavorable concerning underlying effective characteristics.*

Step 2. *Have several judges react to the statements. These judges should examine each statement and classify it as positive, negative, or neutral.*

Step 3. *Eliminate those statements that are not unanimously classified as positive or negative (since neutral statements are not acceptable for inclusion on a Likert scale).*

Step 4. *Decide on the number of choices to be offered for each statement. (Note: The original Likert scale had five alternatives: SD, D, NS, A, SA)*

Step 5. *Prepare the self-report instrument. Include directions. The directions should indicate that the respondents should indicate how they feel about each statement by marking SA if they strongly agree, A if they agree, NS if they are not sure, D if they disagree, and SD if they strongly disagree.*

Step 6. *Administer the scale to a sample of the audience from whom the instrument is intended.*

Step 7. *Compute the correlation between each statement response and the total scale score.*

Step 8. *Eliminate those statements whose correlation with the total scale is not statistically significant (Likert's Criterion of Internal Consistency).*

Furthermore, Popham (1993) revealed that the eight-step procedure above concludes with the application of a "criterion of internal consistency". To ensure that his scales possessed a degree of homogeneity (i.e., they were related statements), Likert requires that items on the scale must correlate significantly with total scale scores. The Likert approach to the construction of effective scales is considered to represent a useful weapon in the evaluator's assessment arsenal. Likert scales are relatively easy t create and are capable of yielding meaningful inferences about the respondents' affective dispositions.

Interview
The interview is a meaning-making process from the identification of a research topic, to the respondent selection, questioning, and answering to the interpretation of responses. The respondent, are not just repositories of information, but also "constructors of knowledge", who interprets and gives meaning to their experiences and endeavors (Holstein & Gubrium, cited in Weinberg, 2002, p. 113). Interview as an instrument in the collection of soft data can elicit oral responses from the subjects which are then recorded by the interviewer. The importance of this data collection method lies in its ability to produce very 'real' data especially on topics that were not previously investigated. Popham (1993) identified several advantages of the interview method, such as: eliciting candid responses from subjects as well as being able to follow up responses from the interviewee which is not possible with written questionnaires.

Interviews could be 'unstructured' or 'structured'. Unstructured interviews can open up whole areas and therefore useful in the initial stages of an evaluation in that 'it will often lead to the identification of issues or questions of which the evaluator was previously unaware. Thus, it often leads to the identification of issues or questions which can be investigated more formally by a questionnaire or test (Andrews 2006). The interview conversation may sometimes be considered a "potential source of bias, error, misunderstanding or misdirection". However, these problems can be controlled if the interviewer asks questions properly, and if the respondent gives out the correct and desired information (Holstein & Gubrium, cited in Weinberg, 2002, p. 112).

Focus Group
The focus group composes of about eight to twelve participants who do not know each other. Although the ideal

group comprises of total strangers there are situations in which this is practically not possible. When such situations arise the groups usually bend the focus group rules to cater to their situational needs (Popham 1993, p.195). This is obvious in our current situation. There are four of us in our group and we know each other because we are undertaking the same course. Furthermore, we are the only students enrolled in the on-campus course mode doing our master's studies in educational evaluation. Therefore it is inevitable to conform to the rules related to the focus group composition. Fusco and Sabo (1999,p.6) state that a focus group is a 'turning point in the evaluation and its value to the program. Focus group enables the evaluation team to form a better understanding of the program, what is happening, and the impact of the course on the students and staff. A Focus group is a tool that enables people involved to get together, socialize, and learn from others' experiences. Furthermore, the focus group is the impetus for increased communication between the director and the instructors. The director acts on the recommendation made by the focus group that directs future improvements add value to the program and guides the future activities of the focus group (Fusco & Sabo 1999, p.6)

Planning and conducting focus group
When a focus group is used in educational evaluations, either for data gathering procedure or in conjunction with other qualitative oriented purposes, the focus is usually on the decisions that need to be made by people involved in the educational program and those who decide on the continuity of that program. It is vital that evaluators who choose focus groups be 'particularly

clearheaded' about the mission of the focus group. Evaluators must have a clear understanding of their purpose of the group roles of participants and understand the needs of the decision-makers who supposed to benefit from the focus group data (Popham 1993) Focus group interviews data collecting method should apply to their particular situation and in which the advantages outweigh the disadvantages' (Popham 1993, p. 205).

The formation of focus groups in an educational evaluation study will require more than one group. The exact number of the group depends on the importance of the evaluation study, the complexity of the questions to be considered, and the number of different categories of an individual involved. In real practice, about three to five focus groups are involved in the evaluation study (Popham 1993). After forming the focus group the next task is to inform the participants. The following rules below specify the steps of group recruitment steps extracted from Popham's focus group rules.

Step one: Schedule a meeting time and venue.
Step two: Invite potential candidates of the group two weeks ahead through the letter or via email and any important information.
Step three: Establish appropriate incentives such as financial support if it is going to require a weekend meeting and so forth.
Step four: Send the following letter to each participant a week before the scheduled meeting. The letter must give full details of

the purpose of the meeting and the specific times of the meetings.

Step five: Contact each participant a day before the meeting reminding them again and stressing the importance of the evaluation study and their roles as participants (Popham 1993, p. 207).

Using a focus group in Educational Evaluation

There are many ways in which educational evaluation evaluators can benefit from using the focus group. Focus group data gathering procedure in an educational evaluation should directly contribute to improving decision-making. Focus groups can be used in collecting data in educational study commonly known as the 'qualitative oriented procedure' (Popham 1993,p.199) Moreover focus groups can provide a deeper understanding amongst various stakeholders about the school the purpose of the school. Data collected can be for formative, summative, and instrument refinement purposes (Popham 1993).

Formative Evaluation

Focus groups can offer insights into a program that students are undergoing. This enables valuable insights about the learners (students) in the educational program regarding the shortcomings that make have been overlooked by the course designer (Popham 1993).

Summative Evaluation

Focus group interviews can offer a better understanding of the results achieved by the educational program as well as the program's aspirations that were not achieved (Popham 1993).

Instrument Refinement

Focus group interviews can help prevail a new instrument's most egregious shortcoming. This means that new assessment instruments designed for a particular educational evaluation study outcomes can be tried out on a smaller group of students for whom the instrument was intended (Popham 1993).

Developing the interview guide for the focus group

The interview guide usually contains a dozen or fewer questions that the moderator will try to cover during the focus group meeting (Popham 1993, p. 208) The questions are sequenced in a logical order beginning with more general questions to more specific questions towards the end. 'Asking good questions take careful preparation and the preparation of the interview guide enables the evaluator to hone on the wording of the questions [so the questions] yield on the desired responses' (Popham 1993, p.208)

CONCLUSION

This theoretical explanation would transfer to the next one who is analyzing the Intensive Listening Course and supposed to get input on our course. The research will be able to include evidence to be provided at the Curriculum Review next year in the future.

REFERENCE

Brinkerhoff, RO, Brethower, DM, Hlucyj, T, & Nowakowski, JR 1983, 'Program evaluation: A practitioner's guide for trainers and educators', (Sourcebook, Casebook), Kluwer-Nijhoff Publishing, Boston.

Brinkerhoff, RO, Brethower, DM, Hlucyj, T, & Nowakowski, JR 1983, *Program evaluation: A practitioner's guide for trainers and educators*, Design Manual, Kluwer- Nijhoff Publishing, Boston.

Cronbach, LJ 1982, *Designing Evaluations of Educational and Social Programs*. Jossey-Bass, San Francisco.

Davis, E 1980, *Teachers as curriculum evaluators*, Allen and Unwin, Sydney.

Fetterman, DM 1988, 'Empowerment evaluation', *Evaluation Practice,* vol. 15, no.1, pp. 1-15.

Feuerstein, MT 1986, *Partners in evaluation: Evaluating development and community program with participants*, Macmillan Press Ltd., London.

Gillham, B 2000, *The research interview*, Continuum, London.

Gravestock, P & Gregor-Greenleaf, E 2008, *Students Course Evaluations: Research, Models and Trends*, The Higher Education Quality Council of Ontario, Canada

Groundwater-Smith, S & Nicoll, V 1980, *Evaluation in the primary school*, Novak, Sydney.

Haywood, J 1977, *Assessment in higher education*, Wiley, London.

Harold & Mckee, B 2003, "The CIPP model for evaluation", *A paper presented to the 2003 annual conference of the Oregon program evaluators Network (OPEN)*, Portland, Oregon.

Hodgkinson, W 2005, EDU5212 Educational Evaluation: *study book* 2006. University of Southern Queensland, Toowoomba.

Holstein, JA & Gubrium, JF 2002, 'Active Interviewing', in Weinberg, D (ed.), *Qualitative research methods*, Blackwell Publishers, Massachusetts.

Houtkoop-Steenstra, H 2000, Interaction and the standardized survey interview: the living questionnaire, Cambridge University Press, United Kingdom.

Kang, TL 1994, "A study of evaluation methods of vocational training effectiveness", *Department of Vocational Training of Executive Yuan*.

Lincoln, TS & Guba, EG 1985, Naturalistic *inquiry* Newbury Park, CA: Sage.

Madaus, GF, Scriven, M & Stufflebeam, DL (eds.) 1983, Evaluation models: Viewpoints on educational and human services evaluation. Kluwer-Nijhoff Publishing, Boston.

Norris, N 1990, Understanding educational evaluation. London: Kogan Page Ltd in association with CARE, School of Education, University of East Anglia.

Patton, M 1986, *Utilised focused evaluation*. Sage, Newbury Park, CA.

Popham, WJ 1993, Educational Evaluation, 3rd edn, Allyn and Bacon, Boston.

Popham, WJ 1988, *Educational evaluation*, 2nd edn, Prentice Hall, Englewood Cliffs, New Jersey.

Scriven, M. 2005, 'The logic and methodology of checklists', Western Michigan University, viewed 25 April 2006, < http://www.wmich.edu/evalctr/checklists/>

Shinkfield, AJ, & Stufflebeam, DL 1995, *Teacher evaluation: Guide to effective practice.* Kluwer, Boston.

Slavin, RE 1992, *Research methods in education*, 2nd edn, Allyn and Bacon, Sydney.

Stake, RE 1975, *Evaluating the arts in education: A responsive approach*, Charles E. Merrill, Columbus,OH.

Stufflebeam, DL 2002, 'CIPP Evaluation Model Checklist', *Evaluation Checklists Project,* viewed 25 April 2006, <http://www.wmich.edu/evalctr/checklists>

Stufflebeam, DL & Shinkfield, AJ 1990, *Systematic Evaluation*, Kluwer-Nijhoff.

Stufflebeam, D et al. 1971, Educational *Evaluation and Decision Making*. Peackock, Itasca, Illinois.

Wiersma, W & Jurs, SG 2005, *Research methods in education: An introduction*, 8th edn, Pearson Education, Inc., Sydney.

Worthen, BR & Sanders, JR 1973, *Educational evaluation: Theory and practice*, Wadsworth Publishing Company, Inc., California.

Worthen, BR & White, KR 1987, *Evaluating educational and social programs: Guidelines for proposal review, onsite evaluation, evaluation contracts and technical assistance*, Kluwer-Nijhoff Publishing, Boston.