DOCUMENT RESUME

ED 204 400 TM 810 412

AUTHOR Konold, Clifford E.; Well, Arnold D.
TITLE Analysis and Reporting of Interview Data.
INSTITUTION Massachusetts Univ., Amherst. Dept. of Psychology.
SPONS AGENCY National Science Foundation, Washington, D.C.
PUB DATE Apr 81
GRANT SED78-22043
NOTE 31p.; Paper presented at the Annual Meeting of the American Educational Research Association (65th, Los Angeles, CA, April 13-17, 1981).

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Data Analysis; *Hypothesis Testing; *Interviews; *Models; Problem Solving; Research Design
IDENTIFIERS *Protocol Analysis

ABSTRACT

Several aspects of interview research heretofore receiving little attention are discussed. A brief description of the different types of interview formats and levels of analysis is presented. Following a discussion of the problem of analyzing protocol data, some suggestions are offered about analysis procedures that derive from constructionist assumtions. A model is offered of the interview which describes its role in hypothesis formulation and hypothesis testing. Views on how the interview can be used in combination with other research methods to investigate problem solving are discussed. Finally, how interview research is currently being reported is examined, and recommendations concerning the types of information necessary for inclusion in such reports are offered. Suggestions are aimed at encouraging the researcher to remain skeptical of interpetations of protocol data, and to report the results of interview research fitting specific criticism from the research community. (Author/GK)

Analysis and Reporting of Interview Data

Clifford E. Konold and Arnold D. Well

Cognitive Development Project

Department of Psychology

University of Massachusetts, Amherst

Paper prepared for presentation at the Annual Meeting of the American Educational Research Association, Los Angeles, CA, 1981. This paper was supported in part by NSF Award No. SED78-22043.

Analysis and Reporting of Interview Data

For the major part of this century, American experimental psychologists interested in the study of learning and cognition have employed methodologies that require subjects to respond in limited ways to carefully structured situations. Inspired largely by the work of Piaget, however, there has recently been a reawakening to the advantages of considering as data the verbal reports of subjects engaged in a complex cognitive task. In problem-solving research, for example, verbal reports not only provide information about the end result of the problem solving process (the answer), but are also a rich source of data concerning the reasoning which led to the final answer. Having some insight into the processes involved in solving a problem not only is important in the formulation of models of human cognition but has important and obvious pedagogical implications.

In this paper we will discuss several aspects of interview research that have received little attention. After presenting a brief description of different types of interview formats and levels of analysis, we will discuss the problem of analyzing protocol data, and offer some suggestions about analysis procedures that derive from constructivist assumptions. Then we will present a model of the interview which describes its role not only in hypothesis formulation, but also in hypothesis testing, and discuss our views on how the interview can profitably be used in combination with other research methods to investigate problem solving. Finally, we will examine how interview research is currently being reported and offer some recommendations concerning the types of information that are important to include in such reports. A major point which will be emphasized throughout the paper is the importance of taking a critical orientation toward hypotheses derived from our research, that the

lack of such an orientation stands as the most serious threat to the acceptance of the interview as a legitimate method of scientific inquiry. Our suggestions are aimed at encouraging the researcher to remain skeptical of interpretations of protocol data, and to report the results of interview research in a way that will permit specific criticism from the research community.

Interview Terminology

There does not yet exist a widely accepted terminology to describe the various types of research interviews. At one extreme, the term "interview," as used in the social psychological literature (e.g., Cannell and Kahn, 1968), can refer to a verbal response to a single question such as "How old are you?" At the other extreme, the term can be used to describe an unstructured and complex exchange between a student and teacher about a difficult problem encountered in the classroom (Davis, 1975a, 1975b). Similarly, analysis of interview protocols can involve the coding of whether or not a correct solution was obtained, or an extensive treatment of every statement made by the subjects. Though there are numerous ways in which an interview could proceed and be analyzed, we will refer to three types of problem-solving interviews and three levels of protocol analysis. The terminology we employ is meant to capture only gross distinctions, but will serve the purposes of this paper.

Interview Formats

In terms of the format of the interview, we will refer to three general styles. These differ primarily in the types of follow-up questions (probes) that are permitted:

Thinking-aloud interview. In the thinking-aloud interview, subjects are given a problem and asked to verbalize their thoughts as they attempt to solve it. They are instructed not to engage in analysis or to reflect back on previous thoughts. Probing, if used at all, is generally restricted to encouraging subjects to vocalize more (cf., Ericsson & Simon, 1980).

Indepth interview. Also referred to as the Piagetian or clinical interview, this method also involves presenting subjects with a specific problem and then questioning them as they attempt to solve it. The probing technique, however, is much more flexible than with the thinking-aloud instructions. Subjects are often asked to reflect back on what they have just done and sometimes offered subtle challenges to their thinking. The interviewer, however, never purposefully gives evaluative responses, nor provides hints in the form of questions, statements, gestures, etc. (cf. Fredette, 1979).

Tutorial interview. In the tutorial interview, the interviewer is interested in eliciting a correct solution, but generally tries to provide no more help than is necessary. Probes are therefore permitted which lead subjects toward solution strategies that may never have occurred to them.

Levels of Analysis

We will refer to three general levels of protocol analysis:

Coded analysis. Coding involves identifying key elements of interest in the protocols and defining them in such a way that raters can go through the protocols noting the presence or absence of the elements. These might include the use of key words, phrases, equations, problem-solving strategies as well as gross to fine motor responses.

Descriptive analysis. With this method the researcher is usually interested in providing a clear restatement of what the subjects said and did during the interview. The attempt is made to describe the data as they are, making no inference about underlying structures that may account for the data. The focus is on the surface structure of subjects' verbalizations, on the meanings they are explicitly trying to communicate.

Interpretative Analysis. In interpretive analysis, inferences are made about the deep structures of subjects' reasoning processes. The researcher uses what the subject says and does to make statements about the process and knowledge structures (both explicit and tacit) the subjects are using to solve the problem.

In Table 1, we have categorized nine problem-solving studies with respect to the method of interviewing and analysis procedures employed. This should not be considered a representative sample. The articles were obtained either from our files or the Journal of Children's Mathematical Behavior. Though we will refer to these studies in more detail in the final section, we include them here to illustrate our terminology.

Insert Table 1 about here

Referring to these reports, one can appreciate the extent to which interviewing styles and analytical methods vary from study to study and, thus, the difficulty in categorizing them as we have. A single interview can proceed through different phases. It can begin with the interviewer using only facilitory probes, if needed, characteristic of the thinking-aloud interview. Once a solution has been arrived at, more indepth probing may be used, followed, if required, by probes designed to lead the subject to a correct solution. We would, however, classify this as a tutorial interview, though the investigators

may (and should) regard each phase of the interview differently (e.g., Rosnick and Clement, in press). Even more difficult to classify are the levels of analysis. Many studies which seem to be interested in descriptive or interpretive analysis also code parts of the interview, and it is often the case that entirely coded protocols are used to support descriptive or interpretive statements. If coded information was included in a report, yet the attempt was made to account for all or most of the subjects' verbalizations and behavior, it was categorized as a descriptive or interpretive analysis. These studies often include large segments of protocols along with their analysis (e.g., Rosnick and Clement, in press). If only selected portions of the interivew were coded, it was categorized as a coded analysis.

To illustrate the distinction between descriptive and interpretive analysis, we have included a segment and a brief analysis of a tutorial interview conducted by John Clement. (The transcript of the entire interview is available on request.) The student being interviewed (Mark) was asked to construct an equation which described the relation between velocity, acceleration and time for a ball dropped off a cliff. Each verbalization of the interviewer (I) and subject (S) was numbered consecutively from the beginning of the interview.

59 I) Suppose the velocity is . . turns out to be twice the time everywhere. As "t" goes from 1 to 2 to 3, velocity goes from 2 to 4 to 6.

60 S) Uh huh.

61 I) What would an equation be for that?

62 S) The velocity is always twice as large as the time.

63 I) Right.

64 S) So. . 2 "v" equals "t", I assume.

65 I) Ok. Write that down.

66 S) (Writes "2v=t") Ok. . ah...

67 I) Now check that out and see if it works.

78 S) Ok. "t" is 1, then "v" is 2. If "t" is 2, then "v" is 4, or you could actually think of this as 2 times "t" equals "t".

79 I) How does that work?

80 S) You set these equal. Since these two are equal, this is the same and you could replace this by "t", and you could say 2 times "t" equals "t".

In line 62, it seems as if Mark has comprehended the nature of the relation he is to express, in that he provides a correct restatement. But the equation he suggests in line 64 is the reverse of the correct equation, 2t=v. It is evident from the remainder of the interview that he did not simply make a careless mistake. It therefore seems clear, in a descriptive sense, what he has communicated. But why does he think that "2v=t" is the appropriate equation? This he does not explain--it just seemed to him to be correct.

Lines 78 and 80 demonstrate why a descriptive analysis is not always straightforward. As a result of plugging numbers into the equation "2v=t", Mark proposes that an equivalent equation is "2t=t". How he arrived at this conclusion is not obvious. But we think what he did was to replace first "t" and then "v" with 1, multiply the left-hand side (LHS) by 2 and thus obtain the solution 2=1. Since the LHS which has been associated with "v" is twice as large as the RHS which had been associated with "t", the relation, in his mind, has been accurately expressed. Replacing the variables with 2, then 4 also maintains the desired relation and supports his feeling that the equation, as

written, is correct. Mark then correctly reasons that since he is replacing both variables with the same values, he can use identical variable symbols on both sides. This descriptive analysis is an hypothesis about what the subject is trying to communicate. Given that the hypothesis is accurate, we still do not know why he is doing what he is. An interpretive analysis would hypothesize what knowledge Mark possesses that permits him to view equations as he does, that allows one side of an equation to be twice as large as the other. (We will suggest an interpretive hypothesis concerning Mark's solution later in the paper.)

Protocol Analysis

Having distinguished among types of interviews and analyses, we are now prepared to address what we consider the most difficult aspect of interview research--analyzing the data. The most straightforward analytic approach is to code selected parts of the interview, transforming them into more manageable, quantitative data. A methodology for coding is well established, and the resultant quantitative data can, if desired, be further subjected to standard, statistical tests. If descriptive or interpretive analyses are desired, however, researchers are virtually left to their own. Little has been written concerning a method for such analyses. In this section we will attempt, if not to establish more formal procedures for descriptive and interpretive analyses, then at least to encourage more dialogue about how a researcher might, "make up for the uncertainties in the method of interrogation by sharpening the subtleties of his interpretation." (Piaget, 1929, p. 9)

A Constructivist View

We begin by considering the problem of analyzing protocols within a constructivist framework and by looking at the interview from the point of view of the subject, the researcher, and the research community. The explication of constructivist assumptions is relevant because we, as most researchers using interviews to study problem solving, have been influenced by Piaget and offer constructivist interpretations of learning and problem solving. We feel that these same assumptions should inform our methods, and consequently our suggestions concerning protocol analysis derived from them.

According to a constructivist view, information or knowledge is not available in raw form to be picked up by passive human receptors. Rather, perception is a selective process insofar as certain features of a stimulus are noticed and others are ignored. It is a constructive process since even those aspects which are selected from the environment are not simply, in the process of selection, made available for human cognition. Rather, perceptions are constructions and, therefore, their nature is determined at least as much by the workings of the perceiver as it is by their presumed independent existence in the external environment.

When subjects are presented with a stimulus in the form, say, of a statistics problem, they attempt to assimilate or make sense of the problem in terms of existing knowledge. Existing cognitive structures permit the identification of the nature of the problem and the relevant information, and the specification of a plan to arrive at the solution. Such information is not inherent in the problem but is selected and simultaneously given particular meanings via existing knowledge.

Psychological research, by its very nature, is reflexive (Gadlin and Ingle, 1975). That is, if we hold particular views about how and why humans

behave as they do, and if we take these views seriously, they ought to be equally powerful in describing the behavior of psychologists. Recognition that as researchers of human phenomena we are both the subjects and objects of study brings us face to face with the contradiction between the objectivity required of a scientist and the subjective nature of human existence. As we attempt to analyze protocols, however, we feel it is critical that we recognize the limits of our objectivity-- that we take into account the subjective nature of our task. If we assume that in solving a problem, subjects are guided and limited by their existing knowledge, we must also be aware that our prior knowledge plays a similar role as we attempt to analyze the resultant data. The nature of our task as researchers is almost identical to that of our subjects, differing only with respect to the particular form of the problem we confront-- our subjects must make sense of the problem we give them; we must make sense of their verbal reports in the context of the problem. The same dynamic is operative when colleagues read an article in which we have reported the results of our research. Their prior knowledge serves as the means by which they come to an understanding of our interpretation of the subjects' understanding. Indeed, they can often learn more from a report about the way we as researchers think than they can about how subjects solved the problem.

To say that human perception is totally determined by prior knowledge, however, is incorrect. Information is never totally assimilated--some accommodation always takes place. Our objective as researchers is, in fact, to accommodate to the protocols we collect--to have them inform us rather than simply confirm prior expectations. The comments we make below are intended to encourage accommodation during the process of protocol analysis.

Goals, Assumptions and Suggestions

Research is undertaken with some general goals and assumptions that determine not only the questions asked and the methods used but also the way in which data, once collected, are analyzed. To understand the research and the conclusions derived from it, these assumptions and goals must be implicitly shared by the reader and the reseacher, or they must be explicitly stated as part of the research report. Our discussion of protocol analysis is based largely on our research experience, and we will attempt, therefore, to make explicit the major goals and assumptions which prescribe our methods of analysis.

Our primary interest has been to explicate student understandings of statistical concepts that we find difficult to teach in the classroom. If we can learn more about the types of prior knowledge students bring to the classroom, we should be able to alter our instructional approach in ways that will be beneficial to our students. With this goal in mind we tend toward interpretive as opposed to coded analyses. We wish to develop models of student problem solving that are powerful enough to capture important individual differences, yet not so specific that we have as many models as we have students. During analysis, we alternate frequently between indepth analysis of individual protocols in which we try to understand what a particular subject is thinking about 'x', to a more general analysis in which we ask what characteristics do all subjects, or a subgroup of subjects, share in thinking about 'x'. This strategy keeps us at a level of analysis that allows us to generalize our understandings beyond the individual, yet does not result in information that is too global to be of any value in instruction. It also provides us with an interactive framework for hypothesis formulation and testing. Hypotheses formulated on the basis of a single protocol are "tested" on other protocols; hypotheses about group performance are "tested" on individual protocols. In this way we hope to avoid

the extremes of either nomothetic or idiographic approaches.

In analyzing individual protocols, we attempt to construct a model that can account for statements by a subject in such a way that no statement seems contradictory or incompatible with another. The assumption that all of a subject's statements are logically consistent is certainly unfounded. If we do not make it, however, it is too easy to discount sections of a protocol that are inconsistent with seemingly understood sections by attributing them to unnoticed verbal slips or to subjects' inability to verbalize what they are actually doing. Piaget (1929) established criteria for differentially considering statements made by children. He felt, for example, that some statements are playfully uttered and are not intended to be taken literally. With college students, however, unless the subject seems flippant or rebellious, all statements should be considered as data to be accounted for.

The assumption that statements made by subjects are, from their point of view, logically consistent provides us with a criterion for analysis. In general, hypothesis A should be considered better than hypothesis B if it accounts for more of the subject's statements. This in turn suggests that attention should be focused on statements that seem somewhat confusing, that at first might be disregarded for the very reason that they do not fit with our understanding of the other statements. In attempting to take such data into account, different hypotheses will be considered. Frequently we have found it to be the case that new explanations which take into account previously ignored information require that different interpretations be given to those sections of the protocol that intially seemed to be best understood.

From a constructivist point of view, the major problem confronting the researcher who wishes to analyze protocol data is not that of generating explanations. We tend to notice in interview protocols those things that

make sense to us. Guided by hypotheses held prior to the interview, or by those derived from only a limited amount of data, explanations can be quickly advanced and data which support these explanations are easy to find. This problem underlies all scientific inquiry (Kuhn, 1962). Indeed, there are no methodologies, no rigorous criteria or formulae which will prevent scientists from usually seeing what they expect to see. As compared to positivistic experimental methodologies, an advantage of the interview technique and an accompanying constructivist epistemology, is that its adherents might be less inclined to view their methods as objectively-paved roads to truth. Carl Jung used to instruct his students to forget what they had learned about symbolism when they analyzed a dream. He believed that one could "never understand somebody else's dream well enough to interpret it correctly." He continually reminded himself of this "in order to check the flow of...associations and reactions, which might otherwise prevail over my patient's uncertainties and hesitations." (Jung, 1964, p. 56) A constructivist orientation can serve as a needed reminder that inferences drawn from data are always, to a greater or lesser degree, invalid.

Sullivan (1954), in reference to analysis of the psychiatric interview, cautioned that early impressions of the interviewer are "rough hypotheses, and, like all hypotheses in interpersonal work, they should be subjected to continuous, or recurrent, test and correction." (p. 121) This advice is equally valid for analyzing protocol data in problem-solving interviews. Care must be taken during the early stages of analysis to scrutinize initial explanations One way to do this is to attempt always to consider more than one hypothesis (cf. Rubin, 1975). This is a difficult task, however, and one must fight the tendency to construct straw-man alternative hypotheses which can serve to falsely confirm the superiority of the favored, initial hypothesis.

Something that we have found to be an effective means of generating

alternative hypotheses is to have the protocols analyzed by a group of investigators. While each individual might arrive too quickly at conclusions about what in the interviews is important and what understanding the subjects have of the problem, these hypotheses tend to be different for each member of the group. Much of our time in analysis is taken up by "negotiating" explanations. One member will advance an explanation and support it with data. Another member will refer to data that do not seem to support the explanation and might offer an alternative account. Thus, not only does a group help bring differing points of view together, it also increases the working store of data. Working together, we are forced to consider data that would likely have been neglected had each of us conducted the analysis on our own. Moreover, group discussions about the plausibility of the explanations thus generated often prompt us to review the data. Explanations are subsequently modified, and the process continues until the group is in general agreement about the adequacy of the explanations.

There are some characteristics of our group that lend themselves well to this process and that are vital to its success. Perhaps most important is that while our backgrounds are sufficiently similar that we can communicate well, our outlooks are different enough that we often disagree even about major goals of our research. Secondly, no member of the group dominates the others. If this were not the case, some members might not support or even advance alternative explanations or criticisms. Finally, the size of the analysis group (3-4 members) is sufficient to ensure diversity, but not so large as to result in endless discussion.

Once rather stable explanations have been developed in this smaller group, we present our conclusions to members of an extended group who are doing interview research in different content areas. The feedback we receive can send us back to the data and the negotiation process.

Challenging Hypotheses Derived from Protocols

As we have portrayed it, the analysis of protocols involves a cyclical process in which hypotheses are developed from data and then tested on the same data set. Protocols are a rich enough data source so that this cyclical process can continue for some time before stable explanations are developed. Once stabilization has occurred, hypotheses can and ought to be more formally tested by collecting more data. In this section we will argue that interviews can be used not only in formulating hypotheses, but in testing them as well. We will also advocate the use of interviews in combination with more traditional research methods.

In Figure 1, we have attempted to characterize the interview in relation to various types of hypotheses and research methods. Rather than viewing the interview solely as a means of generating hypotheses, we see it as a stage of research in which current hypotheses can also be tested and revised. We will elaborate this point in the process of describing Figure 1.

Speaking of the effective interviewer, Piaget (1929) noted that "at every moment he must have some working hypothesis, some theory, true or false, which he is seeking to check." (p. 9) We feel that one cannot avoid such working hypotheses, that they always precede and guide the search for information. In the case of the interview, these may be the very informal hypotheses characteristic of new areas of research, or they may be highly developed theories which have grown out of years of research. On the basis of these hypotheses, one or more problems are constructed. Depending on the specificity of the current hypotheses, these problems may be unstructured and may even be given spontaneously, or they may be carefully thought out, complete with anticipated probes designed to further elucidate subject thought processes.

Insert Figure 1 about here

The simplest form of the interview consists of a problem or question and the subject's response to it, with no interventions by the interviewer. A modification of this type of interview involves the interviewer as a facilitator who interjects probes to keep the subject focused on the problem and responding verbally. This modification is indicated in Figure 1 by dotted lines connecting Probe to Response. In the indepth interview, probes are additionally used to test hypotheses which derive from subject responses (as indicated by the dotted lines connecting Response to Hypothesis). It is this ability to revise and test hypotheses during the process of data collection which constitutes the greatest strength of the indepth interview. However, this fluid, ongoing exchange between hypotheses and data is the very characteristic of the indepth interview which can make it difficult to evaluate the validity of the research. So many hypotheses may have been entertained during the interview that it becomes difficult to specify ways in which the data have been influenced by the interviewer.

A critical analysis of the interview data should try to take into account any effects that hypotheses held by the interviewer may have had. In addition, it should repeatedly test interpretations of the data by determining the extent to which explanations account for the entirety of the statements made by the subjects. Arrows going both ways between Revised Hypothesis and the Interview represent this process of formulating, testing, and reformulating hypotheses.

Thus the testing of hypotheses occurs not only during the analysis phase, but also during the indepth interview. While these tests may permit the researchers to reject many hypotheses and may allow a degree of flexibility

unavailable with other research techniques, they are not, in general, the types of tests that are convincing to or open to criticism by those not directly involved in the research. We feel, however, that research in problem solving is not well served by the exclusion either of interview studies, on the grounds that they are subjective analyses of introspective data, or of more experimental methodologies because they are incapable of exploring thought processes at other than superficial levels. Rather, we prefer a more comprehensive research strategy in which interviews are used in conjunction with other techniques and in which neither flexibility nor rigor are sacrificed. Our goal can best be accomplished if, following protocol analysis, predictions are specified and tested on additional data. This type of follow-up research would provide additional opportunity for the reformulation of hypotheses and would also permit more educated evaluations by the research community. Such follow-up investigations are represented in Figure 1 as a "recycling" of the Revised Hypothesis through more traditional research studies and/or another series of interviews.

As an example of research that tests predictions derived from protocol analysis, we will describe a series of studies conducted by members of our research group who have been attempting to isolate specific difficulties that students have in solving algebra word problems. An example of the type of problem they have been using is given below.

> Write an equation using the variables S and P to represent the following statement: "There are six times as many students as professors at this university." Use S for the number of students and P for the number of professors.

This question was motivated by the interview segment cited earlier in the paper. An initial hypothesis was that Mark, and students like him, had difficulty writing equations when the variables involved were abstract, hard to

visualize entitities such as those encountered in a physics course. However, this proved not to be the case. If the above problem is administered to a random group of undergraduates, no more than 60% are likely to give the correct solution, 6P=S. The most frequently given response is the reversed equation, 6S=P.

Approximately 75 indepth interviews have been conducted in which subjects have been questioned and videotaped as they have attempted to solve this and related problems. The most compelling explanation that has emerged from interpretive analyses of these interviews is that subjects do not have an operative conceptualization of an algebraic equation: they do not view the equation 6P=S, for example, as involving an operation that is performed on the number of professors in order to obtain the number of students. Rather, for many, an equation is a passive description of a current state of affairs. According to this passive view, the equation 6S=P states that currently there are 6 students for every professor. The equal sign in this case does not imply strict equivalency of the expressions on either side, but simply an association between them.

Clement, Lochhead, and Soloway (1980) tested this hypothesis by making the following prediction: that the error rate on problems of the type above would be decreased if the problem were placed in a framework that would emphasize the operational nature of an equation. They felt that computer programming provided such a framework.

Three experiments were conducted in which students with some experience in writing computer programs were administered short written tests composed of problems similar to the one above. Students were asked to write or explain either (a) the appropriate algebraic equation, or (b) a program that would output the value of one variable when given the value of the other. In all

three experiments students made fewer errors using computer language than they did with algebra. On the basis of these experiments, Clement et al. went on to propose five specific hypotheses of why the programming context was successful in reducing error rates. They are currently exploring these hypotheses with more indepth interviews.

This research exemplifies how the interview can be used in combination with other research techniques to explore problem solving, at both the individual and group level. First, an interesting phenomenon was discovered in a informal, tutorial interview. Subsequently, more formal indepth interviews were used to revise hypotheses through a cycle of hypothesis generation and testing. Group experiments comparing performance on two types of problems served to demonstrate the robustness of the reversal error and to test the hypothesis that it resulted from a passive interpretation of an equation. With more interviews being conducted, hypotheses will be cycled through the process again. As hypotheses become more refined, related predictions will become more specific and thus more subject to disconfirmation.

In the study by Clement et al., the hypothesis derived from the protocols was tested in traditional group designs. A strategy which we are just beginning to adopt in our research involves not only deriving predictions which can be tested in such group designs, but also testing predictions about an individual's performance during an interview. Frequently, while we are analyzing a protocol, we generate hypotheses about what the subject is thinking which, in turn, suggest probes to test the validity of these hypotheses. In the past, these probes have been used in the next series of interviews, and their origin and purpose is rarely communicated in the report. We feel, however, that a probe so designed constitutes a test of a hypothesis in a traditional sense, and should be conveyed as such in the research report. This suggests the

need to distinguish between probes designed prior to the interview to test specific hypotheses and those that are spontaneously developed during the interview.

We are also beginning to interview subjects on repeated occasions. This allows us to challenge our interpretations of individual protocols by making predictions of individual responses to related problems or additional probes on the same problem.

When it has been convenient to do so we have asked subjects to comment on our analysis of their interview. We have regarded this as an informal aspect of our research but are beginning to view this as a method to test and develop alternative hypotheses. We are presently preparing to follow up on some interviews concerning student conceptions of probability in which we will have subjects return after we have analyzed an initial interview. At this time we will test our interpretation by first presenting them with subtly different solution strategies and ask which they think best characterizes their method of solving the problem. This latter technique is limited by the fact that we are frequently generating models that elucidate what we think are the tacit levels of a subject's thought processes. But at the very least we can obtain subjects' reactions to the descriptive parts of our analyses. Moreover, we are not sure that subjects would be unable to recognize tacit aspects of their reasoning if we could articulate and present them in an appropriate fashion. Even if they could not, we are excited at the prospect of discussing with subjects the nature of their thought processes. For while subjects may view themselves as being in an inferior role during an interview, we suspect that when we approach them with our explanations of their reasoning processes, they will feel that they have as much or more insight than we do. They can thus provide us with alternative hypotheses about how they are thinking which

we can test as we do any other hypotheses.

We have come to regard the interview not only as a method well suited to hypothesis generation and testing, but also as a common meeting ground, where researchers and participants can perhaps come to regard each other as "colleagues engaged in a search." (Raush, 1969, p. 125)

Reporting the Interview

There are two primary purposes of a scientific report. One is to permit other researchers to replicate and build on important work. The second (and related) purpose is to encourage the scientific community to evaluate the work and generate alternative explanations of the data which may lead to deeper understanding. Driver and Easley (1978) made the comment, in reference to Piaget (1974), that it is "a pity that the summary of the work in Understanding Causality is so brief and reports little of the actual dialogue with pupils, and that more detailed accounts of some of the experiments available have not been translated." (p. 76) If it is difficult to include all the information one would want in a book-length report of an interview study, it becomes impossible to adequately describe even a modest endeavor in a journal-length article. Our suggestions in this section, however, are inspired by the lack of attention devoted to what we regard as critical aspects of an interview, in cases where limited reporting space could not have been the problem.

To evaluate the adequacy of interview reporting we reviewed the nine studies included in Table 1, as well as a study by Karplus (1978). (The Karplus study was not included in Table 1 since no information concerning the format of the interviews was reported.) Our survey of the literature leaves us with the concern that frequently neither of the two purposes stated

above is well served by reports based on interview research. Certainly with interviews it is difficult to communicate clearly how interpretations were generated from the data. In many cases, however, investigators do not even report basic information regarding the way in which the interview was conducted. The following issues ought always to be addressed in the report of interview research:

Subject Caracteristics

Why and how were the subjects selected? How many were there? What were their ages, sex, relevant educational and socioeconomic characteristics? While number and age of subjects is generally reported, information is often lacking about how they were selected (e.g., Chi et al., 1980; Fredette and Clement, 1980; Ginsburg, 1977; Hebbeler, 1977; Karplus, 1978; Kennedy, 1977; Larkin, in press).

Interviewer Characteristics

Who did the interviewing and what is their experience in conducting such interviews? What, if any, relationship previously existed between the interviewer and the subjects? While the skill and experience of the interviewer is critical, especially for the indepth and tutorial interviews, we found information about the interviewer in a minority of the reports surveyed. In some instances the interviewer was identified (Davis, 1975; Fredette and Clement, 1980; Ginsburg, 1977) but the nature of the interviewer's prior experience and whether or not there had been any prior experience with the particular subjects interviewed was virtually never mentioned.

Materials and Instructions

What problems were given and did subjects read them aloud, silently, or were the problems read aloud by the interviewer? What were the subjects told about the purpose of the interview and how were they instructed to respond?

We found that most reports described the problems presented, although in some cases (e.g., Chi et al., 1980; Davis, 1975; Fredette and Clement, 1980; Karplus, 1978), no information was given about the manner in which the problems were presented. We were somewhat surprised to note that instructions to subjects were mentioned in only a small minority of reports.

Interview Situation and Characteristics

What was the interview format (thinking-aloud, indepth, tutorial)? What types of standardized and spontaneous probes were used? How long was the subject allowed to remain silent before a probe was given? (The method of probing is one of the most important aspects of interview research to communicate accurately, and unless large segments of the protocols are included in the report, examples of the types and sequencing of probes should be provided. How long did the interview last and were there problems presented other than the ones currently under discussion? If so, how far into the interview were the current problems presented? Were subjects given any training before the interview began? What were the physical characteristics (e.g., interview location, seating arrangements)? How was the interview recorded and was recording equipment in view of the subject?

The interview format was frequently unclear, but could be inferred. The method by which the interview was recorded was omitted in four cases. Each of the other characteristics mentioned in this section (practice, other problems, location of problems within interview, maximum silent period allowed, length and structure of sessions, physical characteristics) were not reported by the majority of the research reports we surveyed.

Analysis and Reporting of Data

Who analyzed the data, and was analysis made primarily from audio or video tapes, transcripts, or notes? If the data were coded, what was the interrater

reliability? If descriptive or interpretive analyses were made, how were they conducted? For example, was the analysis performed by a single individual, or were several involved and did they work separately or as a group? How were disagreements among individuals analyzing the data negotiated? If interpretations or conclusions are illustrated by excerpts from subject protocols, how representative are these excerpts of subjects in general? How many subjects seem to fit interpretations given, and what statements remain poorly accounted for? How did subjects react to the interview? Were they nervous, resistant, distracted, etc.?

While the level of analysis (coded, descriptive or interpretive) could generally be inferred, not all studies employing coded analysis reported any measure of interrater reliability (e.g., Hebbeler, 1977; Karplus, 1978) nor did they always indicate whether more than one rater was involved. Only two studies reviewed used an interpretive analysis, and one of these (Fredette and Clement, 1980) failed to indicate who had participated in the analysis. It was not always clear how representative reported interview segments were of the other interviews (e.g., Chi et al., 1980; Ginsburg, 1977).

If our suggestions seem simplistic, one need read only a few research reports to appreciate the laxity with which the interview study is often reported. Because of the nature of the interview, the research community is placed in the position of having access to reports with little of the original data available. (The only mention we came across concerning the availability of transcripts was a parenthetical comment in Davis (1972) which promised a complete transcript to "anyone who wants it badly enough," p. 38.) Given such limited access to data, it is even more important for investigators to report as clearly as possible the details of how the interview was conducted and how the data were analyzed. It would also be highly desirable to

make transcripts or tapes available upon request. This would facilitate the difficult task of evaluating interview research and should encourage follow-up work by other research groups.

Conclusions

We have characterized the interview not only as an informal source of hypotheses but as a general method for hypothesis generation and testing. We have implied that the interview may be a more powerful method of investigating certain aspects of cognition than more traditional, experimental methods. However, we feel that the effort to understand cognitive phenomena is ill-served by restricting the methodologies that are used, and have argued for a research approach in which the interview is used in concert with traditional experimental design.

Because of the flexibility of the indepth interview and the difficulty of conducting and reporting descriptive and interpretive analyses, we have argued that the research report should provide detailed information about how the interview proceeded and how the data were analyzed, and that interview transcripts should be made available to interested parties. We have emphasized the need to adopt a critical orientation toward explanations derived from interview data. While our suggestions are informal and simplistic, we are confident that further discussion of the issues to which they were directed will result in refinements of the interview methodology.

References

Cannell, C. F., & Kahn, R. L. Interviewing. In G. Lindzey & E. Aronson (Eds.), The handbook of social psychology, (2nd ed.), Vol. 2, Research Methods. Reading, MA: Addison-Wesley, 1968, pp. 526-595.

Chi, M. T. H., Feltovich, P. J., & Glaser, R. Representation of physics knowledge by experts and novices. Technical report, Learning Research and Development Center, University of Pittsburgh, 1980.

Clement, J., Lochhead, J., & Soloway, E. Positive effects of computer programming on the student's understanding of algebra. Proceedings of the annual Conference of the American Society for Computing Machinery, 1980.

Davis, R. B. Observing children's mathematical behavior as a foundation for curriculum planning. The Journal of Children's Mathematical Behavior, 1971-72, 1, 7-59.

Davis, R. B. Cognitive processes involved in solving simple algebraic equations. The Journal of Children's Mathematical Behavior, 1975, 1, 7-35. (a)

Davis, R. B. A second interview with Henry--including some suggested categories of mathematical behavior. The Journal of Children's Mathematical Behavior, 1975, 1, 36-62. (b)

Driver, R., & Easley, J. Pupils and paradigms: A review of literature related to concept development in adolescent science students. Studies in Science Education, 1978, 5, 61-84.

Ericsson, K. A., & Simon, H. A. Verbal reports as data. Psychological Review, 1980, 87, 215-251.

Fredette, N. H. The clinical interview as a tool for investigating cognitive processes. Unpublished manuscript, University of Massachusetts, Amherst, 1979.

Fredette, N. H., & Clement, J. J. Student misconceptions of an electric circuit: Do introductory physics experiences relate to the real world? Unpublished manuscript, University of Massachusetts, Amherst, 1980.

Gadlin, H., & Ingle, G. Through the one-way mirror: The limits of experimental self-reflection. American Psychologist, 1975, 30, 1003-1009.

Ginsburg, H. The psychology of arithmetic thinking. The Journal of Children's Mathematical Behavior, 1977, 1, 1-89.

Hebbeler, K. Young children's addition. The Journal of Children's Mathematical Behavior, 1977, 1, 108-121.

Jung, C. G. Man and his symbols. New York: Doubleday & Co., Inc., 1964.

Karplus, R. Intellectual development beyond elementary school IX: Functionality, a survey. Technical report, Group in Science and Mathematics Education, University of California, Berkeley, 1978.

Kennedy, M. L. Young children's use of symbolism to solve simple verbal addition and subtraction problems. The Journal of Children's Mathematical Behavior, 1977, 1, 122-134.

Kuhn, T. S. The structure of scientific revolutions. Chicago, IL: University of Chicago Press, 1962.

Larkin, J. H. The cognition of learning physics. American Journal of Physics, in press.

Piaget, J. The child's conception of the world. New Jersey: Littlefield, Adams, & Co., 1929.

Piaget, J. Understanding causality. New York: W. W. Norton & Co., Inc., 1974.

Raush, H. L. Naturalistic method and the clinical approach. In E. P. Willems & H. L. Raush (Eds.), Naturalistic viewpoints in psychological research. New York: Holt, Rinehart, and Winston, 1969.

Rosnick, P., & Clement, J. Learning without understanding: The effect of tutoring strategies on algebra misconceptions. The Journal of Mathematical Behavior, in press.

Rubin, A. Hypothesis formation and evaluation in medical diagnosis. Technical report, Artificial Intelligence Laboratory, Massachusetts Technological Institute, 1975.

Russell, R. L. Addition strategies of third grade children. The Journal of Children's Mathematical Behavior, 1977, 1, 149-160.

Sullivan, H. S. The psychiatric interview. New York: W. W. Norton & Co., Inc., 1954.

Levels of Analysis	Type of Interview: Thinking-aloud	Indepth	Tutorial
Coded	Chi et al., 1980 Larkin, in press	Hebbeler, 1977 Russell, 1977	
Descriptive			Davis, 1975a Kennedy, 1977 Rosnick and Clement, in press
Interpretive		Fredette and Clement, 1980 Ginsburg, 1977	

Table 1. Studies categorized with respect to type of interview and level of analysis.

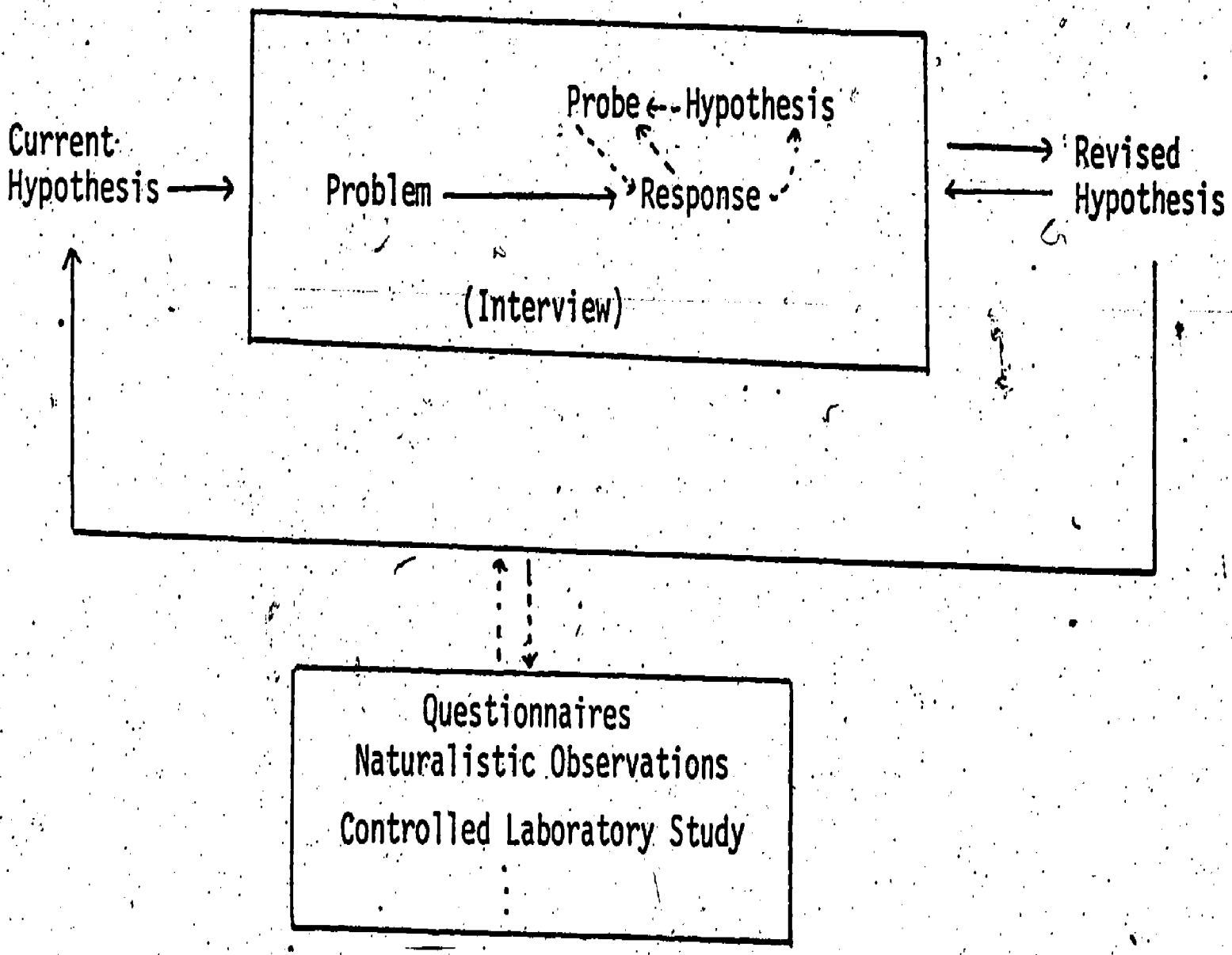

Figure 1. The interview as a component of a general methodology for formulating and testing hypotheses.