

Rhetoricians and researchers have vigorously debated the reliability and validity of protocol analysis findings. Social science scholars have contended the value of verbal data since their original use in the 1920s. This article reviews the history of verbal data in a variety of fields, places protocol analysis in its historical context, and examines more recent claims and criticisms regarding protocol analysis, concluding that protocol analysis, when conducted according to certain principles, can be an important addition to the repertoire of tools for researching the composing process.

The Reliability and Validity of Protocol Analysis

PETER SMAGORINSKY

Oak Park and River Forest High School, Illinois

In a recent issue of College Composition and Communication (February 1987), William F. Irmscher draws on his four decades as rhetorician and composition teacher to review the current state of writing research and make recommendations about how we should conduct it in the future. Among his beliefs is that "In studying the act of writing, investigators should . . . disrupt as little as possible the natural setting of writing with cameras, tapes, and talk-aloud protocols" (p. 85). His skepticism is part of a long history of doubt about the reliability and validity of verbal data. The purpose of this article is to examine the various uses and criticisms of verbal data over the years and to get a better perspective on the value of protocol analysis as a means of researching the composing process.

The first part of this review borrows heavily from documents produced by K. Anders Ericsson and Herbert A. Simon (1978, 1979, 1980), who are among the chief defenders of the use of verbal data. Verbal data, according to Ericsson and Simon (1979), can be collected in two ways: *retrospective* verbalization, in which a subject is asked to talk about cognitive processes that have occurred at an earlier time,

WRITTEN COMMUNICATION, Vol. 6 No. 4, October 1989 463-479
© 1989 Sage Publications, Inc.

and *concurrent* verbalization, which is uttered simultaneously with the performance of the task. Hayes and Flower (1983) further distinguish two types of concurrent verbalization: *directed reports*, in which the subject reports only specified behaviors, and *think-aloud protocols*, in which the subject utters every thought that comes to mind. Most verbal data collected on the writing process, such as the well-known studies by Linda Flower and John R. Hayes (1980a, 1980b, 1981a, 1981b, 1981c plus studies conducted with other researchers), use concurrent utterance, although scholars such as Rose (1980), Odell and Goswami (1982), Scardamalia and Paris (1985), and Langer (1986b) have used retrospective verbalization.

Though composition researchers have only recently begun to employ protocol analysis, the collection of verbal data has been a social science procedure for over half a century. Duncker (1926) and Claparede (1934) are generally credited with developing the idea of think-aloud data. They used their transcripts to study problem-solving processes. Their instructions to their subjects were similar to those given in modern protocol studies:

Try to think aloud. I guess you often do so when you are alone and working on a problem (Duncker, 1926).

Think, reason in a loud voice, tell me everything that passes through your head during the work searching for the solution to the problem (Claparede, 1934).

Data of this type, however, were discounted by their critics as mere "introspection," a most unscientific means of studying a process, even though at least as practiced by such structural psychologists as Wundt and Titchener, it was a highly developed and rigorous scientific method, flawed by its reactive nature and resultant potential for influencing the phenomena under study. Lashley (1923) maintained that "introspection may make the preliminary survey, but it must be followed by the chain and transit of objective measurement" (p. 352). This criticism was echoed decades later in a widely discussed review of literature on verbal data by Nisbett and Wilson (1977), who regarded it as something that might be useful in discovering psychological processes, but is worthless as a means of verification. It could, according to these critics, possibly generate hypotheses but was unacceptable as an indication of anything conclusive.

Defenders of verbal data have agreed with these scholars that when collected and analyzed nonsystematically, verbal data are indeed

inconclusive. They maintain that this dismissal of verbal data as a serious source of information has led to a careless attitude toward their collection, thus compounding their already dubious reputation. Ericsson and Simon (1980), however, posit that these data are useful when collected within the framework of a theory of how the verbal responses are produced. Such a theory would include information like what memory the response information has been stored in, what demands the response makes on short-term memory, and other areas of cognition consistent with the brain's information-processing capacities indicated by psychological research. They also maintain that what subjects say they do must be corroborated by their behavior: The report and outcomes of behavior must be consistent.

Herbert A. Simon (who, incidentally, won the Nobel Prize for Economics in 1978) had previously developed a systematic approach to collecting verbal data with Allen Newell, his colleague at Carnegie-Mellon University, which they called "protocol analysis." A protocol analysis is a procedure to identify psychological processes in problem-solving tasks (Newell and Simon, 1972). It is "a description of the activities, ordered in time, which a subject engages in while performing a task" (Hayes and Flower, 1980). In a protocol, our interest lies not just in the solution to a sequence of actions, but in the processes underlying the sequence itself. Protocol researchers seek to understand what mental processes take place as someone attempts to solve a problem.

Protocol analysis is a very useful addition to the repertoire of research tools for studying the composing process. The data from most other tools yield little about the internal structures of cognitive processes, particularly when the tasks are complex. Think-aloud protocols, in contrast, can yield significant information about the structures of the processes.

One possible concern about the validity of data collected from protocols is that the act of talking while performing a given task might alter the process from the way it would naturally occur. A wide variety of studies has indicated that verbalization does not interfere with cognitive processes. Ericsson and Simon (1978) identify three levels of verbalization:

Level I: Overt utterance of covert articulation or oral encodings, with no intermediate processes.

Level II: Explication of thought content, in which no new information is introduced, but the information is held in a compressed internal

format. Such utterance will cause the subject to take more time with the task, but will not change the structure of the process.

Level III: Utterance that requires the subject to explain thought processes. This involves not simply a recoding of information in short-term memory, but linking it to information in long-term memory.

Think-aloud protocols of the writing process, which include intermediate processes but do not require the subjects to explain their thoughts, consist of level II verbalizations. Several studies of general problem solving have shown that utterance of this type does not interfere with the cognitive processes involved in the task. Karpf (1972) had 40 subjects solve problems while thinking aloud, with 20 control subjects solving the same problems silently. Although the think-aloud group took 50% more time, there were no reliable differences in their ability to solve problems. Feldman (1959) had one subject predict the outcome of a binary choice problem thinking aloud, and had two large control groups solve the same problem quietly. There were no differences in the types of choices or the number of correct predictions. Similar results were found by Weisberg and Suls (1973), Montgomery (1977), and Tversky (1969).

Newell and Simon (1972) compared the number of solutions and detailed solution paths of seven subjects who thought aloud, and 64 who solved the problems silently. They found no differences in the specific proof steps taken by the two groups. Several studies (Bettman, 1970; Clarkson, 1962) have used think-aloud protocols to construct a model that would predict consumer behavior with 80-90% accuracy, suggesting that the patterns from think-aloud protocols reflect the normal course of thought activity. The cumulative results of these studies suggest that the internal structure of thought processes is not disturbed when subjects utter Level II verbalization.

The question we must ask ourselves as writing researchers is: Are these findings from other disciplines applicable to the study of the writing process? Thinking aloud to solve a problem is one thing, but when combined with the physical act of writing, does the act of thinking aloud disrupt the natural process enough to throw it off course? This is perhaps the gravest reservation about think-aloud protocols. The problem has been studied by researchers investigating the use of protocols to analyze the processes involved in problem solving in general, but we can only speculate at this point how much bearing their findings have on the use of protocols to study writing.

The studies of general problem solving have investigated whether the performance of a motor task affects verbalized cognitive processes. Ericsson and Simon (1979) report a number of studies involving verbalization of perceptual-motor processes, such as solving a puzzle (Thomas, 1974, Klinger, 1974) or performing a task requiring switch setting (Duncan, 1963, Shipstone, 1960, Ray, 1957). These studies indicate that subjects have trouble uttering and manipulating at the same time. In considering how these findings affect our acceptance of the use of protocols in studying the writing process, we must ask ourselves how much of a manipulation task writing is. The subjects in these perceptual-motor studies are trying to describe the process behind a decision on how to perform physically. The act of writing, however, is more a transmission of thought to paper, a formal expression of thought, than a manifestation of a decision. Unless one were to find the act of transcription to require cognitive attention (as in the attention a poor typist must give to the act of typing while composing at a typewriter keyboard), the physical act of writing should not require such great attention as to alter the composing process. If we accept the contention that the act of writing is comparable to the way in which people solve math problems with pencil and paper, since both involve thinking that triggers well-rehearsed sensory motor acts, then the studies of Feldman (1959) and others suggest that the act of writing should not interfere with cognitive processes.

Further research is needed in testing the extent to which the act of writing affects the cognitive processes that inspire composition. Witte (1987, p. 400), in studying the role of pre-text in the writing process, concludes that "If simultaneously writing and thinking aloud increase (or decrease) the frequency and extensiveness of pre-text, writing and thinking aloud simultaneously appear to do so inconsistently across writers and across writing tasks." Before we can confidently accept protocol findings, we need additional testing of this problem.

Some might argue, too, that not merely the motor process of writing, but the physical act of speaking could inhibit the natural process of composing. The research reported by Ericsson and Simon, however, does not indicate that the physical act of *utterance* affects the cognitive processes involved in problem solving.

Ericsson and Simon conclude that in problem-solving situations with a heavy cognitive load, subjects may disregard initial instruc-

tions unless the experimenter gives specific prompts to attend to them throughout the course of the session. Involvement in the high-load demands of certain writing tasks, then, may mitigate the effect of initial instructions to provide reasons for writing decisions.

To increase the utterance of thought processes, some experimenters have tried to constrain the manipulation by adding instructions to think about the task and give reasons for each stage of the problem-solving process. These studies would be "directed reports" in Hayes and Flower's (1983) parlance. Ericsson and Simon argue that instructions to verbalize motives and reasons for thinking in a task requiring such manipulation will change the course of processing, because in the studies that do not include in their instructions a request for motives and reasons, the protocols do not contain them. Gagne and Smith (1962), for instance, in asking subjects to solve the Tower of Hanoi problem, used different sets of instructions, some requiring subjects to state a reason for each move. This requirement improved the performance on some tasks, but not others. The researchers suggested that the extra instruction to verbalize reasons affected performance by forcing the subjects to think more about the processes involved in successfully completing the task. This again raises the problem found with the early collection of verbal data, in which the procedure helped shape the phenomena. Protocol researchers must be careful that conditions of the protocol and the task do not influence the thought patterns under study.

Another doubt about protocol research concerns the gaps in the transcripts. Protocols invariably include moments of silence in which the subject struggles with a thought, becomes inarticulate, or gets bogged down in the mechanics of writing. How do we account for the processes occurring during these lapses?

Ericsson and Simon (1979) point out that the information from protocols is incomplete, given the fact that subjects do not utter 100% of their thought processes. Indeed, even with prompting, the subjects do not talk 100% of the time, leaving gaps of 15-60 seconds in the transcripts. They hypothesize that if these gaps in information are consistent and lawful, then even they can provide information about the higher-level structure of the thought process. They point out two occasions on which subjects stop talking, with task-oriented processes taking priority over recoding and verbalization: (1) when the cognitive load is high (Durkin, 1937), and (2) when subjects attend to information that leads to direct recognition of appropriate action (Duncker, 1945). Thus, even the gaps in a protocol may reveal infor-

mation about the complexity of certain processes. Hayes and Flower note that "The psychologist's task in analyzing a protocol is to take the incomplete record that the protocol provides together with his knowledge of the nature of the task and of human capabilities and to infer from these a model of the underlying psychological processes by which the subject performs the task" (1980, p.9). From an aggregate of protocols, then, we may make inferences about the unstated processes of writers.

Cooper and Holzman (1983), however, argue that Flower and Hayes are trying to create a model for the composing process, which they then accept as a description of mental processes, even though this model is in part constructed from inferences made about what happens during gaps in the transcripts. They claim that writing is a stream-of-consciousness activity, rather than a problem-solving task, and therefore it does not represent the Level II tasks described by Ericsson and Simon; protocols, they assert, do not reveal the whole process that generates writing, because subjects will not utter all of their thoughts. Ironically, they criticize protocol analysis for its gaps, when protocols have far fewer gaps than data collected through any other method.

Dobrin (1986) maintains that writing usually takes places over time, with the composing process being interrupted and influenced by dinner, reading, playing tennis, vacations, and whatever else takes places in the normal course of events. A protocol, on the other hand, captures only what happens during one focused session.

Thus he believes that protocols do not in fact yield data that we can use to construct a model of the writing process. He objects to the Flower and Hayes studies and the model they create, saying that the "traces are contaminated" (p.720). He doubts that protocols provide absolute identification of the writing process. My question is, do they need to? Findings derived from protocols must necessarily be limited. Do we need to account for every mental process in order to derive helpful information about what we are studying from protocols?

Dobrin's response is that he is looking for "mechanisms." Different researchers, however, have different agendas. Most experimental researchers are concerned with reducing variables, constraining behavior in order to study some phenomenon in isolation. To claim that protocols are unnatural and therefore misleading because the conditions (prompts from a researcher, the "unnatural" setting, the time limit) affect performance would seem a dismissal of all experimental research. Researchers studying a specific problem are not concerned

with whether their subjects think about lunch, rather, they are concerned with studying particular operations. When Newell (1984) and Langer and Applebee (1987), for instance, compared the types of mental processes at work when students take notes, do short-answer study questions, and write essays, the protocol conditions helped the subjects and researchers focus on the types of reasoning operations used in the different activities. Structuring the research to eliminate extraneous thought is hardly a procedure of distortion. Dobrin's objection seems to be that protocols do not account for every single thought that passes through a subject's mind as he composes, whether the thought is irrelevant to the task or not. He seems to want a comprehensive research method, one that I doubt exists.

One researcher has tried, with limited success, to study the extent to which the laboratory conditions of a think-aloud protocol affect process. Uncomfortable with the artificiality of protocol conditions, Berkenkotter (1983) undertook a remarkable and ambitious study of a writer at work under *natural* conditions. She collected protocols from an established writer—Donald M. Murray—writing under three conditions:

- (1) Murray composed articles for professional journals at home in his den over a two-month period, turning on a tape recorder whenever he composed. These sessions yielded over 120 hours of tape. He also submitted to the researcher photocopies of all composing, including notes and drafts.
- (2) Murray wrote for one hour in a *laboratory* setting, following a specific topic ("Explain the concept of death to 10-12 year old readers of *Jack and Jill* magazine.>").
- (3) Murray wrote over a two-day period with the researcher present to record his utterance as he revised an article for a professional journal. The researcher also elicited retrospective data at intervals during the two-day session.

Berkenkotter's report on her segmenting and coding is vague ("classifying and then counting all audience-related activities . . . my coder and I tallied all planning, revising, and editing activities"), perhaps due to the extraordinary amount of data she was faced with. Her results are nonetheless interesting, perhaps due to Murray's unique style of composing at home: He made notes and drafts, and then dictated a rough draft to his wife, who would type it; Murray would then revise the typed draft.

The second writing condition, the customary protocol condition of composing in a room with a researcher and tape recorder, produced very different results from the successful articles Murray wrote at home. "I have rarely felt so completely trapped and inadequate," he said (Murray, 1983, p. 169). "To find equivalent feelings from my past I would have to go back to combat or to public school." He produced only 17 words of text, finding himself unable to relate to either his topic or his audience, seemingly inhibited by the situational variables of the protocol. I wonder, though, whether the topic—explaining the concept of death to preadolescents—might not have been a big part of his problem. Could he have written fluently about this under *any* circumstances? We must wonder why Berkenkotter used such a formidable task. Certainly, topic is a major variable in comparing writing conditions. A more logical approach would have been to give Murray the same task (i.e., write an article, write a personal narrative, state and defend an opinion) to perform under both conditions. Had a researcher world enough and time (and money), it would be more interesting and informative to compare the processes behind producing text on similar topics under *natural* conditions at home and the *laboratory* conditions of a protocol. Berkenkotter is to be commended for conceiving such a visionary study, and Murray deserves our thanks and respect for agreeing to serve as her subject. Her conclusion from Murray's difficulties under the second set of conditions that protocols do not elicit information that reflects a *real* process, however, is not justifiable from her comparison of these two conditions.

Some researchers use retrospective protocols, a procedure that usually involves an interview with the subject immediately after the composing session. Rose (1980), among others, feels that this technique elicits important information without interrupting composing. Rose wanted to compare competent writers with writer's block to competent writers who wrote with ease. His data-collecting procedure was "admittedly nonexperimental, certainly more clinical than scientific" (p.390), consisting of interviews in which the researcher used the writer's notes, drafts, and finished composition to direct his questions. Rose's technique is typical of retrospective protocol collection and does not require the type of rigorous segmenting and coding necessary for concurrent data analysis. Retrospective and concurrent protocols both have strengths and weaknesses; the weakness of a retrospective protocol is that we must rely on the subject's recall and impressions, a subjectivity that some might not trust as being reliable. Hayes and Flower (1983) identify two problems: Interference causes

loss of information from memory, and subjects tend to forget goals and subgoals once they have been accomplished (known as the "Zeigarnik effect"). A composite picture from both retrospective and concurrent protocols might yield the corroboration necessary to draw strong conclusions.

Indeed, Langer (1986b), in studying relationships between reading and writing processes, used both concurrent and retrospective protocols and found that "In general, while the think-alouds were longer than the retrospective reports, few significant mode differences were found" (p.178), the differences tending to be that the concurrent reports would include more remarks about surface or mechanical features of writing while attention to more substantive features of processes was very similar. Her data included 96 concurrent and 84 retrospective reports, certainly a very large sample for protocol research. The concern that concurrent protocols would interrupt and thus alter the composing process, then, is not supported by Langer's study.

Brand (1987) criticizes the cognitive process model developed from protocols for not taking language and personality into account. This, however, is something that we can do; it simply has not yet been a concern of protocol researchers. The segmenting and coding procedures can isolate and identify whatever the researcher wants to study, providing that the protocol elicits such information. If the writing task were altered to require an essay on a formative or emotional experience, the researcher could establish criteria to identify such language, and could then study emotion and personality.

Brand's concern raises the issue of how to go about analyzing data. Researchers have developed a variety of methods for analyzing protocol transcripts. The pioneering work of Emig (1971) and Mischel (1974) was of an exploratory nature, and the researchers did not attempt to treat the think-aloud transcripts in a systematic way, instead reporting impressionistic responses to the utterance of their subjects. Later, Perl (1979) and Hayes and Flower (1980) developed segmenting and coding systems to analyze features of a general "writing process"; Perl identified 27 categories, including planning, assessing, talking leading to writing, reading related to topic, and various types of editing, while Hayes and Flower identified such categories as metacomment, content statement, and interjections.

Since then, other researchers have used protocols to study more particular aspects of composing, and have developed their systems of

analysis to suit their research foci. The following studies are representative of ways in which researchers have used protocol analysis.

Flower and Hayes (1981c), studying what writers do when they pause as they write, collected protocols from one novice and three experts. They used independent raters to define "composing episodes" and thus identify the pauses. They do not explicitly describe their segmenting and coding procedures; they report identifying behaviors such as setting content goals, setting process goals, acting on goals, evaluation, review, metacomment, and others.

Berkenkotter (1981) developed a more rigorous coding system to compare the audience awareness of published professors of rhetoric and composition, and published professors of other disciplines. She asked the subjects to explain their careers to a high school audience. She identified 25 composing "activities" that fell into four categories: analyzing and/or constructing a hypothetical audience; setting goals and naming plans aimed at a specific audience; evaluating content and style (persona) with regard to anticipated audience response; and reviewing, editing, and revising for a specific audience.

Langer (1986b) used protocols to explore how children of different ages (8, 11, and 14) make meaning as they read and write. To study data, she developed a system which she called "Analysis of Meaning Construction," focusing on the relationships between the nature of the task and the knowledge and skills the children bring into play. Her system featured *Reasoning Operations*, such as posing questions, making hypotheses, and developing schemata; and *Monitoring Operations*, such as task or topic goals, genre/discourse structures, and lexical repertoire. Her use of these categories helped her examine the protocols of students engaged in both reading and writing tasks. She found that students at all three ages are concerned with the ideas they are creating, but engage in more questioning and hypothesizing at young ages of what they know or learn when they are older. Her *monitoring* research indicated that students are less attentive to meaning when they write than when they read, attending more to goals, subgoals, genre, mechanics, and lexical choices when composing. She used her Analysis of Meaning Construction to analyze data in two other studies (Langer, 1986a; Langer and Applebee, 1987; it was also adapted by Marshall, 1987) to study the effects of reading and writing on intellectual development.

From this review, I would draw the following conclusions about protocol analysis as a method for researching the composing process:

- (1) Although imposing *unnatural* writing conditions, the standard protocol method of recording a subject's utterance with prompts in a given time period can describe composing processes that parallel those that take place under *natural* conditions. The utterance, however, is more distilled, with the subject's attention more focused on the writing task. While protocol subjects take longer to complete their writing task, the processes are nonetheless similar. The protocol conditions resemble *exam* conditions under which many students produce much of their writing in school. We could hypothesize that writing done under "exam" conditions includes processes similar to those found in writing done over longer periods of time. A study like Berkenkotter's (1983), with better control for topic, could indicate the extent to which the processes are parallel. In general, protocol researchers would benefit from researching the method itself, as Berkenkotter tried to do, and as Langer and Applebee (1987) have done in comparing the results of concurrent and retrospective protocols. As noted in this review, there is considerable precedent for researching the strength of the method in other areas of the social sciences.
- (2) The method of protocol analysis should emerge from a theoretical framework underlying the problem the researcher is trying to study. Newell (1984) and Langer and Applebee (1987), for instance, hypothesized that essay writing would result in more complex thought processes than more common types of school writing such as note-taking and doing short-answer study questions. In developing their coding systems, then, they tried to identify thinking operations and arrange them in a hierarchy. Their studies have important implications for classroom teachers, suggesting that note-taking and short-answer evaluations allow for less noetic growth than do extended compositions. Similarly, Durst (1987) found that in analytic writing, students employ more varied and complex thinking than they do in summary writing, asking more complex questions, making higher level plans, and spending more time interpreting the readings and evaluating their own essays. These may seem like common sense conclusions, but other research (Applebee, 1981; Goodlad, 1984) indicates that students in our schools engage far more frequently in these low-level reasoning tasks, suggesting that many students' education is lacking in opportunities for developing intellectual skills. Without the view of thought processes provided by the protocols, Newell, Langer and Applebee, and Durst could not have arrived at these insights.
- (3) As has been frequently noted (i.e., Flower and Hayes, 1980; Hayes and Flower, 1983; Swarts, Flower and Hayes, 1980), the identification of processes must be substantiated by several forms of evidence. Thus, the protocol transcript should be compared with the written product. Independent judges should agree on the segmenting and coding of the

transcripts. There seems to be no accepted standard for agreement. Newell (1984), for instance, randomly selected five instances from each of his subcategories from 48 protocols and wrote each on a card, with a second rater coding a stack of 125 randomly ordered cards. Using this difficult system of reaching agreement using segments taken out of context, he reached the following rates of agreement: 79% for main categories, 60% for learning operations, 55% for elements. Since context often sheds light on how to classify a statement, we should not be surprised by these low rates of agreement.

Berkenkotter (1981) employed a very different system, using three independent coders to identify audience-related activities, documenting 25 activities falling into 4 categories. She reported only those activities on which all of her coders agreed. Anyone who has worked on a protocol study knows the difficulty of getting high rates of agreement; indeed, developing coding and segmenting procedures that independent judges can use and agree upon is among the most challenging aspects of conducting useful protocol research. The great disparity among rates of agreement in published studies, however, indicates the need for a better sense among protocol researchers of what constitutes acceptable agreement procedures and standards.

- (4) The researcher should identify variables that affect the subjects, and thus influence the content of the protocol transcripts. These variables include:
 - (a) The conditions of the protocol situation, including the specific instructions to the writer, the behavior of the researcher, and the time constraints.
 - (b) Inherent qualities of the subjects, such as skill, intellect and writing fluency.
 - (c) The subjects' acquired knowledge; i.e., what have their writing experiences been prior to the collection of the protocol? If we assume that these experiences (including methods of instruction) influence process, then we would benefit from knowing the experiences that have shaped our subjects' development. Our subjects are not a *tabula rasa*, but people who have been through years of education and other writing experiences. In examining protocols, we must distinguish between *natural* processes and learned processes, and try to determine means for identifying the sources of entry-level characteristics that could influence the thinking and composing processes that emerge in the protocols.

Protocol analysis is an expensive and meticulous research method that has had its share of growing pains. Much of the criticism directed at it in the journals has helped point out its flaws and limitations, and

helped researchers refine their methods. A number of imaginative and responsible researchers are developing new applications for protocol analysis that are expanding our knowledge about the complex processes behind composing. To discount protocol analysis as a method is to ignore the many contributions it has already made to our understanding of written communication, and to dismiss the knowledge it is bound to uncover about composing in the future.

REFERENCES

- Applebee, A. (1981). *Writing in the secondary schools*. NCTE Research Report No. 21. Urbana, IL: National Council of Teachers of English.
- Bereiter, B., & Bird, M. (1985). Use of thinking aloud in identification and teaching of reading comprehension strategies. *Cognition and Instruction*, 2(2), 131-156.
- Berkenkotter, C. (1981). Understanding a writer's awareness of audience. *College Composition and Communication*, 32, 388-399.
- Berkenkotter, C. (1983). Decisions and revisions: The planning strategies of a publishing writer. *College Composition and Communication*, 34(2), 156-169.
- Bettman, J.R. (1970). Information processing models of consumer behavior. *Journal of Marketing Research*, 7, 370-376.
- Brand, A. (1987). The why of cognition: Emotion and the writing process. *College Composition and Communication*, 38(4), 436-443.
- Claparede, E. (1934). Genèse de l'hypothèses (Genesis of the hypotheses). *Archives de Psychologie*, 24, 1-155.
- Clarkson, G. (1962). *Portfolio selection—A simulation of trust investment*. Englewood Cliffs, NJ: Prentice-Hall.
- Connors, R. (1983). Composition studies and science. *College English*, 45(1), 1-20.
- Cooper, M., & Holzman, M. (1983). Talking about protocols. *College Composition and Communication*, 34, 284-293.
- Davis, J. H., Carey, M. H., Foxman, P. N., & Tarr, D. B. (1968). Verbalization, experimenter presence, and problem solving. *Journal of Personality and Social Psychology*, 8, 299-302.
- Dobrin, D. (1986). Protocols once more. *College English*, 48(7), 713-726.
- Duncan, C. P. (1963). Effects of instructions and information on problem solving. *Journal of Experimental Psychology*, 65, 321-327.
- Duncker, K. (1926). A qualitative (experimental and theoretical) study of productive thinking (solving of comprehensible problems). *Pedagogical Seminar*, 33, 642-708.
- Duncker, K. (1945). On problem solving. *Psychological Monographs*, 58, no. 270.
- Durkin, H. E. (1937). Trial-and-error, gradual analysis and sudden reorganization: An experimental study of problem solving. *Archives of Psychology*, No. 210.
- Durst, R. K. (1987). Cognitive and linguistic demands of analytic writing. *Research in the Teaching of English*, 21(4), 347-376.
- Emig, J. (1971). *The Composing Processes of Twelfth Graders*. Urbana, IL: NCTE.

- Ericsson, K. A., & Simon, H. A. (1978). *Retrospective verbal reports as data*. C.I.P. Working Paper No. 388.
- Ericsson, K. A., & Simon, H. A. (1979). *Thinking-aloud protocols as data: Effects of verbalization*. C.I.P. Working Paper No. 397.
- Ericsson, K.A., & Simon, H. A. (1980). Verbal reports as data. *Psychological Review*, 87(3), 215-251.
- Ericsson, K. A., & Simon, H. A. (1984). *Protocol analysis: Verbal reports as data*. Cambridge, MA: MIT Press.
- Faigley, L. (1986). Competing theories of process: A critique and a proposal. *College English*, 48(6), 527-542.
- Feldman, J. (1959). *An analysis of predictive behavior in a two-choice situation*. Unpublished doctoral dissertation, Carnegie Institute of Technology.
- Flower, L. S. & Hayes, J. R. (1977). Problem-solving strategies and the writing process. *College English*, 39, 449-61.
- Flower, L., & Hayes, J. (1980a). The dynamics of composing: Making plans and juggling constraints. In L. Gregg and E. Steinberb (Eds.), *Cognitive processes in writing: An interdisciplinary approach*. Hillsdale, NJ: Erlbaum.
- Flower, L., & Hayes, J. (1980b). The cognition of discovery: Defining rhetorical problems. *College Composition and Communication*, 31, 21-32.
- Flower, L., & Hayes, J. (1981a). A cognitive process theory of writing. *College Composition and Communication*, 32, 365-387.
- Flower, L., & Hayes, J. (1981b). Plans that guide the composing process. In C. H. Frederiksen and J. F. Dominic (Eds.), *Writing: The nature, development, and teaching of written communication* (pp. 39-58). Hillsdale, NJ: Erlbaum.
- Flower, L., & Hayes, J. (1981c). The pregnant pause: An inquiry into the nature of planning. *Research in the Teaching of English*, 15(3), 229-243.
- Flower, L., & Hayes, J. (1984). The representation of meaning in writing. *Written Communication*, 1(1), 120-160.
- Gagne, R. H. & Smith, E. C. (1962). A study of the effects of verbalization on problem solving. *Journal of Experimental Psychology*, 63, 12-18.
- Goodlad, J. I. (1984). *A place called school: Prospects for the future*. New York: McGraw-Hill.
- Gorrell, R. (1983). How to make mulligan stew: Process and product again. *College Composition and Communication*, 34(3), 272-277.
- Hayes, J., & Flower, L. (1980). Identifying the organization of writing processes. In L. Gregg and E. Steinberg (Eds.), *Cognitive processes in writing: An interdisciplinary approach*. Hillsdale, NJ: Erlbaum.
- Hayes, J., & Flower, L. (1983). Uncovering cognitive processes in writing: An introduction to protocol analysis. In P. Mosenthal, L. Tamor, & S. Walmsley (Eds.), *Research in Writing: Principles and Methods*. New York: Longman.
- Irmscher, W. (1987, February). Finding a comfortable identity. *College Composition and Communication*, 81-87.
- Karpf, D. A. (1972). *Thinking aloud in human discrimination learning*. Unpublished doctoral dissertation, State University of New York at Stony Brook. (*Dissertation Abstracts International*, 33, 6111-B)
- Kaufert, D., Hayes, J. & Flower, L. (1986). Composing written sentences. *Research in the Teaching of English*, 20(2), 121-140.

- Klinger, E. (1974). Utterances to evaluate steps and control attention, distinguish operant from respondent thought while thinking out loud. *Bulletin of the Psychonomic Society*, 4, 44-45.
- Langer, J. (1986a). Learning through writing: Study skills in the content areas. *Journal of Reading*, 29(4), 400-406.
- Langer, J. (1986b). *Children reading and writing: Structures and strategies*. Norwood, NJ: Ablex.
- Langer, J., & Applebee, A. (1987). *How Writing Shapes Thinking: A Study of Teaching and Learning* (NCTE Research Report No. 22). Urbana, IL.: NTE.
- Lashley, K. S. (1923). The Behavioristic Interpretation of Consciousness II. *Psychological Review*, 30, 329-353.
- Marshall, J. D. (1987). The effects of writing on students' understanding of literary texts. *Research in the Teaching of English*, 21(1), 30-63.
- Mischel, T. (1974). A case study of a twelfth grade writer. *Research in the Teaching of English*, 8, 303-314.
- Montgomery, H. A. (1977). A study of intransitive preferences using a think aloud procedure. In H. Jungermann and G. de Zeeuw (Eds.), *Decision making and change in human affairs*. Dordrecht, Holland: D. Reidel Publishing Company.
- Murray, D. (1983). Response of a laboratory rat—or, being protocoled. *College Composition and Communication*, 34(2), 169-172.
- Newell, A., & Simon, H. A. (1972). *Human problem solving*. Englewood Cliffs, NJ: Prentice-Hall.
- Newell, G.E. (1984). Learning from writing in two content areas: A case study / protocol analysis. *Research in the Teaching of English*, 18(3), 265-287.
- Nisbett, R. E. & Wilson, W. T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, 84, 231-259.
- Odell, L., & Goswami, D. (1982). Writing in a non-academic setting. *Research in the Teaching of English*, 16(3), 201-223.
- Penner, B. C., and Voss J. (1983). *Problem solving skills in the social sciences: Methodological considerations*. Pittsburgh: University of Pittsburgh, Learning Research and Development Center.
- Perl, S. (1979). The composing processes of unskilled college writers. *Research in the Teaching of English*, 13, 317-336.
- Ray, W. S. (1957). Verbal compared with manipulative solution of an apparatus-problem. *American Journal of Psychology*, 77, 289-290.
- Rose, M. (1980). Rigid rules, inflexible plans, and the stifling of language: A cognitivist analysis of writer's block. *College Composition and Communication* 31, 389-400.
- Scardamalia, M., & Bereiter, C. (1983). The development of evaluative diagnostic, and remedial capabilities in children's composing. In M. Martlew (Ed.), *The psychology of written language: A developmental approach*. London: John Wiley & Sons.
- Scardamalia, M., & Paris, P. (1985). The function of explicit discourse knowledge in the development of text representations and composing strategies. *Cognition and Instruction*, 2(1), 1-39.
- Schweiger, D. (1983). Is the simultaneous verbal protocol a viable method for studying managerial problem solving and decision-making? *Academy of Management Journal*, 26, 185-192.
- Shipstone, E. I. (1960). Some variables affecting pattern conception. *Psychological Monographs*, 74, (Whole No. 504).

- Steinberg, E. (1986). Protocols, retrospective reports, and the stream of consciousness. *College English*, 48(7), 697-704.
- Swarts, H., Flower, L. & Hayes, J. (1980). Designing protocol studies of the writing process: An introduction. Paper presented at Research in Language Arts Workshop, NCTE convention.
- Thomas, J. C. (1974). An analysis of behaviour in the Hobbits-Orcs problem. *Cognitive Psychology*, 6, 257-269.
- Tversky, A. (1969). Intransitivity of preferences. *Psychological Review*, 76, 31-48.
- Weisberg, R., & Suls, J. M. (1973). An information-processing model of Duncker's candle problem. *Cognitive Psychology*, 4, 255-276.
- Witte, S. P. (1987). Pre-text and composing. *College Composition and Communication*, 38(4), 397-425.

Peter Smagorinsky teaches English at Oak Park and River Forest (Illinois) High School. He recently received his PhD in English Education from the University of Chicago, where his dissertation used protocol analysis to study the effects of different instructional foci on composing processes. His recent publications include articles in Written Communication, the English Journal, and the 1989 NCTE Classroom Practices edition, Students Connecting with Literature; co-authorship of the NCTE ERIC/TRIP monograph Explorations: Introductory Activities for Literature and Composition, 7-12; and co-authorship of Fostering the Reader's Response: Rethinking the Secondary Literature Curriculum from Dale Seymour Publications in Palo Alto, CA.