#904

How to Study Thinking in Everyday Life: Contrasting Think-Aloud Protocols With Descriptions and Explanations of Thinking

K. Anders Ericsson

The Florida State University

Herbert A. Simon

Carnegie-Mellon University

In his article Smagorinsky (this issue) describes how protocol analysis (Ericsson & Simon, 1980, 1984, 1993) can be applied to study thinking within the perspective of cultural–historical activity theory (CHAT). Smagorinsky's main claim is that verbalization of thinking as speech is a "process through which thinking reaches a new level of articulation" (p. 173) and he explicitly questions the empirical evidence reviewed by us that it is possible under *some* circumstances to have participants think aloud without altering the course of their thinking. A primary reason for his misunderstanding appears to be that Smagorinsky focused almost exclusively on our earlier publications that described our original theoretical proposals and thereby missed the numerous subsequent experiments explicitly testing them. In fact, Ericsson and Simon (1993) discussed over 30 additional studies that provide results consistent with our theoretical framework. In this commentary we briefly review the evidence that supports different types of verbalization activities and describe how our distinctions match the classic distinction between thinking as inner speech and as social speech and give explicit quotes by Vygotsky (1962), where he argued for very similar differences. Within our framework we can incorporate the circumstances where verbalization of thinking (thinking aloud) can be made without reactive effects and other circumstances where verbal descriptions and explanations of thinking serve "as a tool that potentially enables changes in consciousness" (Smagorinsky, this issue, p. 157).

We endorse Smagorinsky's efforts to integrate protocol analysis into the CHAT perspective as an alternative approach, for particular purposes, to the use of verbalizations as data, and we try to show that recent lines of research are currently applying protocol analysis in this spirit to studies of thinking in everyday life including socially situated activities. We have organized our comments on Smagorinsky's article as answers to the following questions: Is it possible to study objectively that form of thinking that occurs covertly in many types of typical tasks and activities in everyday life? If so, under which circumstances can adults think aloud without altering the structure and course of their spontaneous thinking? Are there other circumstances under which the request to

describe or explain one's thoughts can improve the coherence of thinking and thus facilitate individuals' learning and development? What are productive ways of studying the social components and contexts of human thinking?

IS IT POSSIBLE TO STUDY THE COVERT THINKING IN MANY EVERYDAY ACTIVITIES?

Everyone would probably agree that individuals' speech in everyday life reflects aspects of their thoughts and that they typically verbalize only a fraction of all their thoughts. In fact, the expression of many thoughts that might occur to them in social activities is actively suppressed. Virtually any type of social interaction would be dramatically changed if the parties verbalized all their thoughts, including those that people normally keep to themselves (e.g., their evaluations of other participants in the interaction).

Consider the potential effect of having individuals think aloud while engaging in the following three social activities. If a chess player thought aloud about what he or she was planning to do during a chess match, the opposing chess player would gain an advantage. Similarly, if medical doctors examining their patients verbalized all their thoughts about potential diagnoses, their patients would often be frightened or confused. It would at least be distracting for the customers at a restaurant if a waiter with exceptional memory verbalized all his thoughts while he encoded the dinner orders in memory. In all of these examples, thinking aloud would at least have an indirect effect on the verbalizing individuals by influencing the social reactions and behavior of the surrounding people who overheard the verbalized thoughts. The first step toward studying covert thinking requires that we find a nonreactive setting to reproduce this type of thinking under controlled conditions.

In his pioneering studies of chess, de Groot (1946/1978) showed that it is possible to reproduce the superior thinking of chess experts during chess matches in settings that do not require the presence of the opposing player. De Groot presented world-class and expert chess players with specific positions from unfamiliar chess games and asked them to select the best next move for each position while thinking aloud. Subsequent research has confirmed that the ability to find the best next move under such circumstances is closely related to individual differences in chess skill and performance during chess tournaments (Charness, 1991; Ericsson & Lehmann, 1996). De Groot's methodology has been generalized to identify naturally occurring tasks that are representative of a domain of activity and capture competence and performance in that domain (Ericsson & Smith, 1991). For example, written descriptions of the symptoms and background information of patients have been presented to medical doctors for diagnosis (Patel, Arocha, & Kaufman, 1994). Furthermore, the natural conditions under which some individuals display exceptional memory have been partially recreated in the laboratory to allow the reproduction of the exceptional performance under controlled circumstances (Ericsson & Polson, 1988a, 1988b). In the last decade, skilled performance observed in everyday life has been successfully reproduced with representative tasks and thus captured in the laboratory for a wide range of domains of expertise (Ericsson & Lehmann, 1996) including everyday skills, such as typing, memorization, and text comprehension. (For a discussion of how group behavior and team performance could be studied with this type of methodology and some current limitations for this research approach see Ericsson, 1996, and Okada & Simon, 1997). It is established that many types of everyday performance can be successfully reproduced outside their original reactive social setting, then the next step is to

study the processes that mediate it and, in particular, to analyze the participants' thinking-aloud protocols. To demonstrate that this can be done, it is necessary to review briefly the evidence that shows that, under *some* circumstances, thinking can be verbalized without altering its course.

EXTERNALIZING COVERT THINKING WITHOUT ALTERING IT

Philosophers since Aristotle have argued that thinking is essentially a sequential process, where one thought leads to another (Ericsson & Crutcher, 1991). Today it is relatively uncontroversial that thinking can be represented as a sequence of thoughts (relatively stable cognitive states) interspersed by periods of processing activity, as is illustrated in the top portion of Figure 1. The main methodological issues have been to determine how to gain information about the associated
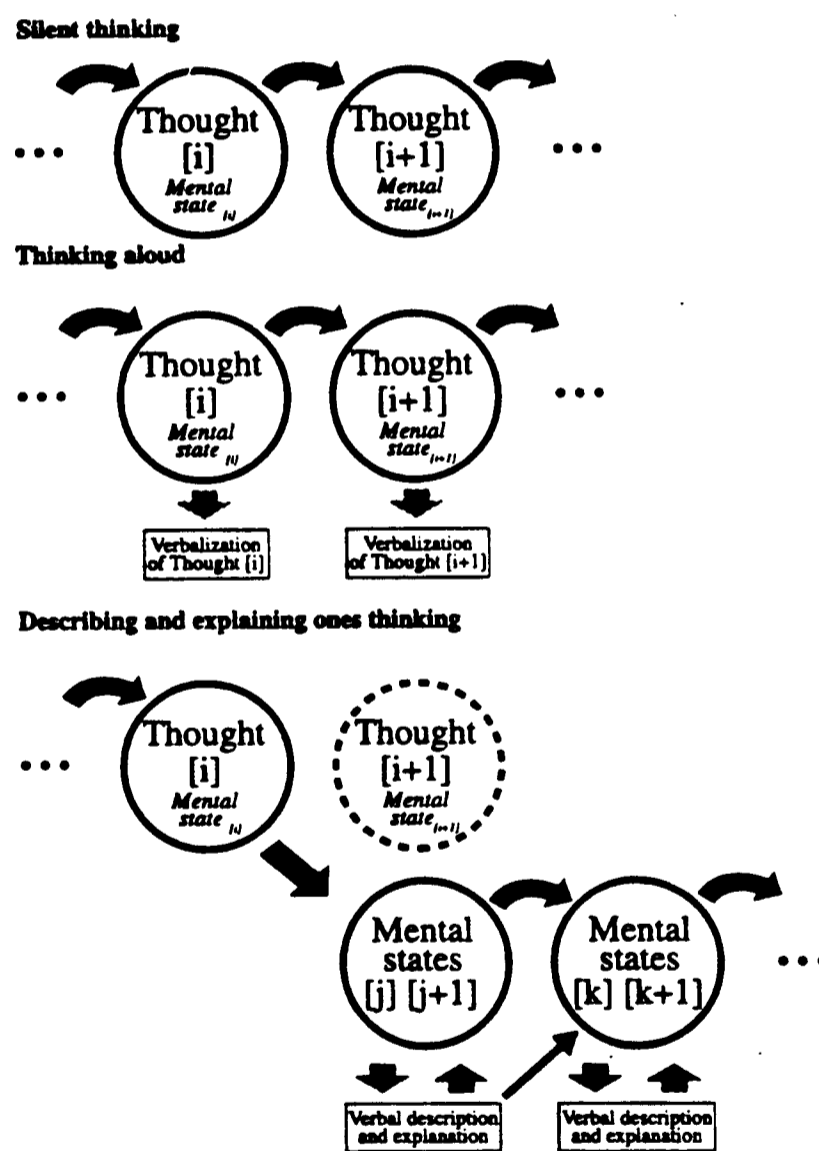

FIGURE 1   The relation between silent thinking and two different modes of verbalization of thoughts. The spontaneous sequence of thoughts associated with silent thinking is shown at the top. The middle panel shows how these thoughts can be verbalized during thinking aloud without changing their sequence. The lower panel shows how participants' efforts to describe and explain thinking can change the sequence of thoughts and lead to the intrusion of additional thoughts as well as lead participants to generate coherent descriptions for the sole purpose of allowing listeners to understand their thoughts.

thought states without altering the structure and course of the naturally occurring thought sequences. The history is full of attempts to gain information about the structure of thinking from self-observation, from skilled introspective analysis by expert observers, and most recently from complete explanations of the methods that experts claim to use when solving problems of a certain type. In our review of the reactive effects of those types of verbalization activities, (Ericsson & Simon, 1980, 1984, 1993) we showed that the participants have to go beyond merely verbalizing spontaneously generated thoughts to produce the thoughts that would contain descriptions and explanations of the kinds just mentioned (see the lower panel of Figure 1). In fact, we showed why verbal descriptions of one's thinking directed to another individual are likely to alter the course of thinking—consistent with Smagorinsky's claims in his article. Our distinction drew extensively on the classic distinction between inner speech and social speech, and to facilitate the mapping we make repeated references to Vygotsky's (1962) characterization of inner speech, although, as far as we know, Vygotsky never tried to elicit and study thinking (inner speech) in adults.

Nonreactive Verbalization of Thinking—Think-Aloud

The primary focus of our work has been to identify the circumstances where individuals could verbalize their thoughts without any, or at worst with minimal, reactive influences on their thinking. Perhaps the single most important precondition for successful direct expression of thinking is that the participants are allowed to maintain undisrupted focus on the completion of the presented tasks. Hence, participants are explicitly instructed to focus on the task while thinking aloud and merely to verbalize their thoughts (see the middle panel of Figure 1) rather than describe or explain them to anyone else. Smagorinsky (this issue) does a good job of summarizing most of the specific procedures for eliciting think-aloud protocols that minimize the distraction from social influences and diversion of the participants' attention.

However, because the expectation that participants should describe or explain their thoughts verbally to another person is often deeply ingrained, especially in students, we recommend that after being given the standard instructions to think aloud, the participants should be given a series of simple warm-up tasks (Ericsson & Simon, 1984, 1993), such as a mental multiplication (24 × 36), that provides opportunities to practice directing their full attention to the presented task while verbalizing their thoughts. Under these conditions the sequence of verbalizations consisting, for example, of "carry the 2," "fourteen," "one forty four," "let's see," and "seven twenty" do not constitute a coherent discourse but rather the verbalization of a sequence of intermediate calculations and mediating thoughts.

These types of verbalizations are consistent with Vygotsky's comment that "inner speech appears disconnected and incomplete" (p. 139). With think-aloud (see middle panel of Figure 1) the participants do not appear to monitor their overt verbalizations of thoughts, mostly generate incomplete sentences and phrases, and rarely correct their verbalizations including speech errors, or in Vygotsky's (1962) words "in inner speech words die as they bring forth thought. Inner speech is to a large extent thinking in pure meanings" (p. 149). Furthermore, the think-aloud verbalizations often provide a relatively incomplete record of all the knowledge and complex cognitive processes that compose successful task performance. Analyzing the thinking for a task calls for a task analysis to explicate all of the different thought sequences and methods that participants, given their knowledge of facts and methods, could have used to generate solutions for that task. Using methods

with high intercoder reliability makes it possible to assess many aspects of the participants' thought processes from the transcribed think-aloud verbalizations.

When we only consider think-aloud protocols elicited with the methods described previously, three major findings emerge. First, protocols provide detailed evidence on the sequences of thoughts in a wide range of tasks. The evidence shows that expert performers have not automatized their thought processes but verbalize thoughts involving planning, evaluation, and reasoning that contribute to their superior performance (Ericsson & Lehmann, 1996). Second, the verbalized thought sequences by participants performing a given task are consistent with the task analyses, and participants are at the same skill level display similar characteristics of thought.

Hence, at least for the tasks studied up to this point, which include representative tasks capturing everyday performance, Smagorinsky's (this issue) concern that the experimenter and participants may not converge on a shared representation of the task seems unfounded. By restricting the research to tasks with socially agreed performance characteristics, it appears that the structure of the task imposes the constraints for functional mental representations and successful thinking. Finally, when participants are properly instructed merely to verbalize their thoughts (think-aloud), no changes in the sequence of the thought processes have been found compared to participants completing the same tasks silently. Smagorinsky (this issue), in claiming a "lack of sufficient specific research on protocol methodologies" (p. 160) omits the post-1984 evidence. Since 1980, numerous researchers have collected a large body of evidence where many studies have been explicitly designed to assess the effects of different types of verbalization procedures. In our updated review (Ericsson & Simon, 1993) we discussed around 30 new studies, and since 1993 many additional ones have been published consistent with our conclusions.

The evidence identifies two modes of verbalization of thought that are qualitatively different (see Figure 1). When participants are thinking aloud, their sequences of thoughts have not been found to be systematically altered by verbalization. However, when participants are asked to describe and explain their thinking, their performance is often changed—mostly it is improved. In sum, when Smagorinsky concludes that "the processes of rendering thinking into speech" (p. 173) is a reactive process where "thinking reaches a new level of articulation," he is describing socially directed speech in perfect agreement with the effects we attribute to Level 3 verbalization (Ericsson & Simon, 1980, 1984, 1993). When Smagorinsky presents excerpts from protocols that purport to show altering effects of verbalization, these verbalizations do not meet the criteria for thinking aloud but include socially directed speech with descriptions and explanations. Hence a reactive influence from socially directed speech confirms the important distinctions between different types of verbalizations and indirectly supports the unobtrusive expression of thoughts while thinking aloud.

FROM CONCERNS ABOUT REACTIVE INFLUENCES ON NORMAL THINKING TO THE TRANSFORMING POWER OF REFLECTIVE THOUGHT

Thinking aloud has now gained acceptance as a central and indispensable method for studying thinking (Crutcher, 1994; Payne, 1994; Wilson, 1994), and it is time to start examining the mechanisms mediating alternative "reactive" modes of verbalization, such as giving verbal descriptions and explanations of one's thinking (Ericsson & Simon, 1993). The necessary transformation from a mere expression of one's thoughts to speech describing/explaining them to another individual is substantial as eloquently outlined by Vygotsky (1962): "It is a complex, dynamic process involving the transformation of the predicative, idiomatic structure of speech into

syntactically articulated speech intelligible to others" (p. 148). In consistency with Vygotsky's view, we proposed that participants, while describing and explaining their thinking, have to unpack complex thoughts into a understandable series of verbalized ideas and also need to develop a system of references to ideas and perceptually available objects (Ericsson & Simon, 1980, 1984, 1993). Under these circumstances, participants also monitor their speech to ensure that it is understandable, and they make corrections and further explications of their thoughts, whenever necessary. In particular, we argued that these requirements for verbalized explanations biased participants to adopt more orderly and rigorous strategies to the problems that were easier to communicate in a coherent fashion, but in turn altered the sequence of thoughts (see lower panel of Figure 1).

Although verbal descriptions and explanations may not reflect spontaneous thinking with complete accuracy, such verbalizations present a genuine educational opportunity to make students' reasoning more coherent and reflective. Recent studies collecting thinking-aloud protocols have shown that the more successful students studying completed examples engage more in active efforts to master the materials than less successful students, and the best students generate more self-explanations and monitor their learning better (Chi, Lewis, Reimann, & Claser, 1989; see Renkl, 1997, for a review). Most important, Chi, de Leeuw, Chiu, and LaVancher (1994) found that students prompted to generate self-explanations improved their recall and understanding compared to a control group.

There are now many investigators studying comprehension and problem solving who analyze the detailed structure of verbalized thought with the goal of identifying the critical thought processes that mediate more effective learning and improved transfer (Berardi-Coletta, Buyer, Dominowski, & Rellinger, 1995; Renkl, 1997; Trabasso & Suh, 1993). However, the best current evidence on the key attributes of superior thinking and performance comes from studies of expert performers thinking aloud which shows that their superior performance is primarily due to previously acquired mental representations that allow them to plan, evaluate, and reason about alternative courses of action (Ericsson & Kintsch, 1995; Ericsson & Lehmann, 1996; Gobet & Simon, 1996). There is even some evidence from expert performers supporting Vygotsky's (1962) claim that these benefits and transforming effects of social communication are enhanced when participants externalize their thoughts in writing. Many scientists point to writing as the most effective (as well as demanding) activity to improve and develop their thinking (Ericsson, Krampe, & Tesch-Römer, 1993). By externalizing the thoughts in an explicit form, the scientists can critique and refine them even before sharing them with their colleagues (cf. the method of repeated drafts described by Vygotsky, 1962, p. 144).

It is our hope and belief that the research described provides an important source of evidence and theoretical ideas that supports and elaborates Smagorinsky's (this issue) argument for the transforming powers of describing and explaining one's thoughts. We believe that most of the differences between the study of thinking with think-aloud methods and with socially directed descriptions/explanations reflect basic differences in the goals of these two kinds of investigation. With this difference in goals in mind, a collaborative pursuit of both directions of inquiry can contribute to our understanding of how to support individuals' development of independent thinking and skilled performance.

THE SOCIAL COMPONENTS AND CONTEXT OF HUMAN THINKING

In addition to his discussion of the effects of thinking aloud on thought processes Smagorinsky makes several contrasts between what he perceives as the individualistic bias of the thinking-aloud procedure and the social context of other forms of verbalization. The issues here are rather different

from the point just described (and about which we seem to be in agreement): Behavior is changed when students are asked not simply to think aloud but to analyze and explain their own task behavior. We believe that Smagorinsky's account seriously underestimates the social character of the behavior exhibited in thinking-aloud protocols; hence, it is in his characterization of the information processing (IP) perspective (quite inaccurate in our view) that we have our greatest differences with him.

As we have seen, the evidence is consistent that the course of the thought process can be inferred in considerable detail from thinking-aloud protocols. Obviously this process is very much shaped by what the person who has given the protocol knows and what skills he or she possesses—by all of the things that have been stored in the brain. But this store of knowledge and skills—including the problem-solving techniques that are employed—is to the highest degree a product of social interactions (with people and, yes, with books and artifacts) over the whole lifetime of the problem solver. The language in which the protocol is expressed is the language of the culture of the speaker—its vocabulary, its syntax, the ideas it expresses. The problem solver is unlikely to use calculus in solving a problem unless he or she is drawn from a culture that is familiar with it, or to paraphrase Shakespeare or the poetry of Jorge Borges, without previous social interactions that stored memories of them, including in those memories familiar patterns to evoke them when these patterns reappear in the current situation.

Although it has not always been expressed in these terms, the social character of expertise has been a central theme in the large body of research that has shown how expert performance depends on stored knowledge and the stored patterns that recognize when that knowledge is relevant and access it. The same thing appears when we look at the research on scientific creativity that has used, if not protocols, laboratory records and other detailed accounts of the scientist's verbal behavior to trace the processes of discovery (Kulkarni & Simon, 1988).

Against this background it seems quite strange and wholly misleading to characterize the cognitive models that are built from an IP perspective as "relatively static" and not "developmentally oriented," as Smagorinsky does. Given the large role played in these models by long-term memory, it can be said, contrary to Smagorinsky, that:

- They emphasize mediated action in a context.
- They insist on the importance of historical levels of analysis.
- They seek to ground their analysis in everyday life events (at least if experts can be said to have everyday lives).
- They recognize that mind emerges in the joint mediated activity of people.
- They assume that individuals are active agents, but do not act in settings entirely of their own choosing.

Of the contrasts that Smagorinsky proposes between the IP and CHAT accounts of cognition, all that seems to remain is that the IP accounts do indeed seek to emphasize explanations, including explanation in terms of cause and effect as well as methodologies that provide as much protection as possible against the subjectivism of the investigator. They have no biases of which we are aware against emphases on "the emergent nature of mind in activity," if we understand the intent of that phrase.

In short, we are disappointed in Smagorinsky's faulty characterization of the IP perspective and its associated research and, in particular, in his omission of the recent efforts to study everyday

life events. However, we are pleased to have this opportunity to correct some of the misconceptions and point out the many common goals shared by the two theoretical perspectives and the potential relevance of the recent research on expertise and expert performance for the CHAT perspective. More generally, we believe that protocol analysis will provide a particularly rich source of evidence that would allow both research traditions to collect mutually acceptable evidence relevant to the current and future theoretical claims and controversies.

REFERENCES

Berardi-Coletta, B., Buyer, L. S., Dominowski, R. L., & Rellinger, E. R. (1995). Metacognition and problem solving: A process-oriented approach. *Journal of Experimental Psychology: Learning, Memory and Cognition, 21*, 205–223.

Charness, N. (1991). Expertise in chess: The balance between knowledge and search. In K. A. Ericsson & J. Smith (Eds.), *Toward a general theory of expertise: Prospects and limits* (pp. 39–63). New York: Cambridge University Press.

Chi, M. T. H., de Leeuw, N., Chiu, M. H., & LaVancher, C. (1994). Eliciting self-explanations improves understanding. *Cognitive Science, 18*, 439–477.

Chi, M. T. H., Lewis, M. W., Reimann, P., & Glaser, R. (1989). Self-explanations: How students study and use examples in learning to solve problems. *Cognitive Science, 13*, 145–182.

Crutcher, R. J. (1994). Telling what we know: The use of verbal report methodologies in psychological research. *Psychological Science, 5*, 241–244.

de Groot, A. (1978). *Thought and choice and chess.* The Hague, The Netherlands: Mouton. (Original work published 1946)

Ericsson, K. A. (1996). The acquisition of expert performance: An introduction to some of the issues. In K. A. Ericsson (Ed.), *The road to excellence: The acquisition of expert performance in the arts and sciences, sports, and games* (pp. 1–50). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Ericsson, K. A., & Crutcher, R. J. (1991). Introspection and verbal reports on cognitive processes—two approaches to the study of thought processes: A response to Howe. *New Ideas in Psychology, 9*, 57–71.

Ericsson, K. A., & Kintsch, W. (1995). Long-term working memory. *Psychological Review, 102*, 211–245.

Ericsson, K. A., Krampe, R. Th., & Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review, 100*, 363–406.

Ericsson, K. A., & Lehmann, A. C. (1996). Expert and exceptional performance: Evidence on maximal adaptations on task constraints. *Annual Review of Psychology, 47*, 273–305.

Ericsson, K. A., & Polson, P. G. (1988a). An experimental analysis of the mechanisms of a memory skill. *Journal of Experimental Psychology: Learning, Memory and Cognition, 14*, 305–316.

Ericsson, K. A., & Polson, P. G. (1988b). Memory for restaurant orders. In M. Chi, R. Glaser, & M. Farr (Eds.), *The nature of expertise* (pp. 23–70). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Ericsson, K. A., & Simon, H. A. (1980). Verbal reports as data. *Psychological Review, 87*, 215–251.

Ericsson, K. A., & Simon, H. A. (1984). *Protocol analysis: Verbal reports as data.* Cambridge, MA: Bradford Books/ MIT Press.

Ericsson, K. A., & Simon, H. A. (1993). *Protocol analysis: Verbal reports as data* (Rev. ed.). Cambridge, MA: Bradford Books/ MIT Press.

Ericsson, K. A., & Smith, J. (1991). Prospects and limits in the empirical study of expertise: An introduction. In K. A. Ericsson & J. Smith (Eds.), *Toward a general theory of expertise: Prospects and limits* (pp. 1–38). Cambridge, England: Cambridge University Press.

Gobet, F., & Simon, H. A. (1996). Templates in chess memory: A mechanism for recalling several boards. *Cognitive Psychology, 31*, 1–40.

Kulkarni, D., & Simon, H. A. (1988). The processes of scientific discovery: The strategy of experimentation. *Cognitive Science, 12*, 139–175.

Okada, J., & Simon, H. A. (1997). Collaborative discovery in a scientific domain. *Cognitive Science, 21*, 109–146.

Patel, V. L., Arocha, J. F., & Kaufman, D. R. (1994). Diagnostic reasoning and medical expertise. In D. Medin (Ed.), *The psychology of learning and motivation* (Vol. 30, pp. 187–251). New York: Academic.

Payne, J. W. (1994). Thinking aloud: Insights into information processing. *Psychological Science, 5*, 241, 245–248.

Renkl, A. (1997). Learning from worked-out examples: A study on individual differences. *Cognitive Science, 21*, 1–29.

Trabasso, T., & Suh, S. (1993). Understanding text: Achieving explanatory coherence through on-line inferences and mental operations in working memory. *Discourse Processes, 16,* 3–34.

Vygotsky, L. S. (1962). *Thought and language.* Cambridge, MA: MIT Press.

Wilson, T. D. (1994). The proper protocol: Validity and completeness of verbal reports. *Psychological Science, 5,* 249–252.