Some Additional Considerations of Protocol Analyses

Bill Potter

California State University, Stanislaus

Protocol analyses are typically transcribed tape recordings of the vocal-verbal emissions a participant makes while engaged in a task. These recordings, generally used by cognitive psychologists, are used to shed light on how a participant solved a problem or completed a task. In the last issue of *The Analysis of Verbal Behavior*, several behavioral researchers examined the use of protocol analyses by behavior analysts. Austin and Delaney (1998) suggested that protocol analysis could supplement primary data, especially when tasks that have a limited set of solution paths are used (e.g., solving a math problem). These authors also offered several methods for increasing the validity of the obtained reports. Hayes, White, and Bissett (1998) examined the utility of protocol analysis in understanding the verbal behavior some have come to label *rules*. Critchfield and Epling (1998) examined issues related to protocol analysis and measurement systems, and, more important, the nature of the functional relations involved in covert and overt verbal behavior. Before behavior analysts rush to adopt protocol analysis, however, several issues should be clarified.

A-B-C Analysis of Transcripts

When conducting a protocol analysis, the researcher typically comes away with a transcript derived from the tape recordings made during a session.

Many thanks to Hank Schlinger for his "shaping" of this article.

Address correspondence to Bill Potter, Department of Psychology, California State University, Stanislaus, 801 W. Monte Vista Avenue, Turlock, California 95382 (E-mail: wpotter@toto.csustan.edu).

In isolation these transcripts are not very helpful from a behavioral perspective, especially if one considers them as part of a larger chain of responses, some verbal and some not, that are controlled by various aspects of the environment (e.g., the sound of the participant's own voice or the sight of a number after writing it down when solving a math problem). To provide a more complete analysis, the antecedents and consequences of the vocal emissions should also be considered. An example of such an analysis is shown in Figure 1. This research incorporated a matching-to-sample task in which mediating verbal behavior was examined (Potter, Huber, & Michael, 1997), with all experimental conditions arranged on a computer. Note that the transcript was analyzed along with both the sample stimulus (the flag-lake pattern) and the choice stimulus that the participant selected (the dots). Although this was not a complete re-creation of the stimulus conditions in effect during the actual trial, it was enough to allow the researchers to examine the supposed controlling variables of the verbal emission. To be more representative of the conditions in effect during the vocal emission, the researchers could have presented the entire screen (i.e., include all the remaining choice stimuli).

A complete transcript such as this can allow the participant to help to clarify the controlling variables of his or her verbal emission after the session is over. For example, in the Potter et al. (1997) study, the participants were asked to review their transcripts, similar to Figure 1, and then to try to identify what caused them to say what they

<p>"lets see a white computer chip in a small basket"</p>			<p>C</p>
---	---	--	----------

Fig. 1. A partial transcript, including transcribed verbal emission, sample stimulus, choice selected, and consequence. The last column indicates that the participant selected the correct choice stimulus, and thus heard the sound "correct" (Potter, Huber, & Michael, 1997).

said, that is, to identify the controlling variables of their responses. In most cases the participants "labeled" the stimuli. In the case of Figure 1, the flag-like pattern was the "computer chip," and then the participant connected the dots into a pattern easily recognized as a "small basket." Both cognitive and behavioral researchers point out that such retrospective reports can be faulty, especially if the participants are simply asked to explain what they were thinking when they emitted the verbal behavior indicated in the transcripts (Austin & Delaney, 1998; Ericsson & Simon, 1993). However the type of retrospective report described here differs from the traditional use of retrospective reports in two ways. First, the stimulus conditions under which the original response was emitted are re-created. This reduces the likelihood that current, nonrelevant stimuli influence the participant's responses. Second, such data are not used as primary data. The participant's reports are only used to verify the researcher's own scorings. In fact, interobserver agreements could have been calculated between the researchers' scoring of the transcripts and the participants' scoring. Potter et al. found a high degree of correspondence between their own scorings and the participants' scorings, although interobserver agreement was not calculated. These types of interobserver agreements may not be possible to calculate for all tasks. To the extent that they clarify the antecedents and consequences of a response, they may be useful to behavior analysts.

Evolution of Verbal Behavior Functions

A second issue of concern is the difference between well-established verbal behavior and verbal behavior that evolves to perform some function. In most protocol analyses, researchers examine verbal behavior that functions to resolve some problem. Seldom do they examine how those responses came to serve that particular function, that is, how the verbal responses change over time to solve the problem at hand. In the Potter et al. (1997) study, verbal responses tended to become smaller (i.e., shorter) over time. Wulfert, Dougher, and Greenway (1991) found similar results, noting,

Incidentally, all subjects changed their utterances over the course of the experiment. At the beginning, their verbalizations tended to be lengthy (e.g. "I guess these lines should go with the box, but then, I don't know, they could also go with the other one"). However, over multiple repetitions of the same stimulus configurations, all subjects essentially ended up using telegram style, simply naming sample and comparison stimuli ("cross-triangle") or occasionally only referring to the comparison stimulus ("triangle"). (p. 493)

In addition to the shortening of responses, it is possible that new functional relations emerge as the participant continues to be exposed to a task. Huber (1997) found that participants' vocalizations in a conditional discrimination task not only shortened over time but also appeared to be controlled by a different stimulus condition. Early in the sessions the participants appeared to emit tacts to both the sample

and choice stimuli, but later in the sessions the minimal responses seemed to be tacts of either the relation between the choice and sample stimuli or some kind of minimal descriptor (i.e., perhaps just a link in a now-established behavioral chain). It is interesting to note that some have hypothesized that relations in stimulus equivalence research "emerge" (or develop) during the testing conditions (for an interesting discussion of this issue, see McIlvane & Dube, 1990). This is analogous to evolving verbal behavior relations during the course of research, and to the extent that stimulus equivalence may be a function of verbal behavior (e.g., Horne & Lowe, 1996; Wulfert et al., 1991), it is possible that the evolution of verbal behavior during exposure to a task, and the emergence or stimulus equivalence relations, are actually the same issue.

"Thinking"

Many cognitive researchers conceptualize protocol analyses as revealing thinking or the processes related to thinking (Ericsson & Simon, 1993). Although this issue was partially addressed by Critchfield and Epting (1998), further analysis is warranted here. The present perspective provides no new analysis as compared to that offered by Skinner (1953, 1957) and others, but rather considers "thinking" in relation to Ericsson and Simon's (1993) account of protocol analysis. Thinking from a behavioral perspective is most likely composed of covert responses and resultant stimulation, which can be classified into two general categories. The first is covert verbal behavior. An example of this would be when we read silently and "hear" each word of our textual response. American sign language practitioners may covertly "sign" and "see" those signs. The second category consists of what have been called conditioned perceptions (Palmer, 1991; Skinner, 1953, 1957). These might include conditioned hearing (as described above) in

regard to covert textual responses and conditioned seeing in which we seem to "see" some stimulus that is not present. Most behavior analysts have few problems with covert verbal behavior, but conditioned perceptions are more difficult to explain and will not be addressed in this paper. The manner in which covert verbal behavior and conditioned perceptions may combine to create "thinking" is aptly described by Palmer (1991):

If asked our whereabouts last Wednesday, it helps to be able to say: "Wednesday . . . 9:00—lab; 1:00—learning seminar; selectmen's meeting in the evening." With such an intraverbal framework we could provide a plausible account of our day without further ado. However, such an account smacks more of inference than of memory. That is, the subjective impression that we are "reliving" the experience is missing if we merely recount what we must have been doing according to our invariant schedule. Key responses such as "selectmen's meeting" may control conditioned perceptions of a particular room in a particular town hall. These perceptual responses in turn control other responses. (p. 274)

Conceptualizing such covert responses and stimulation in this manner provides a parsimonious analysis of thinking according to behavioral principles. In addition, such an analysis provides an explanation of some of Ericsson and Simon's conclusions regarding protocol analysis. These researchers classify protocol analyses according to three different levels. Level 1 is simply talking aloud, whereas Level 2 consists of the description of thought content such as describing a conditioned perception. Level 3 includes the explanation of thought processes or thoughts, meaning not only describing any conditioned perceptions but also explaining how they may have influenced one's behavior. In all cases, participants are typically given instructions and training as to how they should respond.

In their review of research in this area, Ericsson and Simon (1993) found that participants' performances did not change as a function of talking aloud (Level 1) when doing a task, if the par-

ticipants were only asked to talk aloud. Research involving Level 2 protocol analyses showed some differences between participants asked to talk aloud and nontalking control groups. Level 3 protocol analyses produced clear differences between silent groups and Level 3 talk-aloud groups. Researchers examined a number of tasks in these different assays, including recall, problem solving, decision making, and so forth.

From a behavioral perspective, Ericsson and Simon's (1993) conclusions make sense. In Level 1 analyses, participants are simply asked, and trained, to talk aloud. In this case any covert verbal behavior is presumably emitted overtly. Critchfield and Epting (1998) pointed out that talking aloud is not a simple matter of amplifying covert verbal behavior, but that different functional relationships are involved. That is, some controlling variables must be different in order for typically covert verbal behavior to be emitted overtly (i.e., instructions, coupled with training and a nonpunishing audience). However, if one adopts the position that overt and covert "vocal" verbal responses are controlled by generally the same stimulus conditions and serve similar functions, then such responses would belong to the same response class (Catania, 1998; Donahoe & Palmer, 1994). In this case no special analysis is required, and the variations in response form have little or no impact on task performance. The assumption made here is that the different controlling variables affect only the topography of the response, not its function. For a thorough review of such effects, see Skinner's (1957) chapters on "The Audience" and "Supplementary Stimulation."

The Level 2 and Level 3 results reported by Ericsson and Simon (1993) also make sense from a behavioral perspective. If the participant is required to report ongoing conditioned perceptions (Level 2) and to report ongoing performance (Level 3), responses that would otherwise not be emitted in a

Level 1 analysis would now be emitted. In both cases additional relations are introduced into the task, and are likely to change the results. The extent of that change, then, is probably a function of the task's complexity, the participant's ability to stay on task, and a number of other variables. Such responses become even more disruptive when one considers that conditioned perceptions and covert verbal behavior may occur in parallel with other conditioned perceptions (e.g., seeing and hearing) but may control sequential acts by the participant.

Conclusions

Protocol analyses may be useful in providing a more thorough understanding of the variables that are responsible for behavior, particularly the behavior of humans with extensive verbal repertoires. But this usefulness is most likely limited to Level 1 analyses, that is, simply making covert verbal behavior overt. In the case of Level 2 and Level 3 analyses, the participant is likely to emit additional verbal responses that are not functionally equivalent to the covert responses. Behavior analysts who use protocol analyses should remain true to behavioral principles and methodologies. All protocol analyses should, however, be taken as sufficiency arguments only: It is possible that such performances would be obtained without the emission of verbal behavior. That is, covert verbal behavior could simply be collateral responding, serving no function relative to the performances under scrutiny.

Given the difficulty of research in this area, some researchers are examining the utility of computer models based on operant and respondent conditioning. Such an approach involves programming a computer to learn by experience, versus the traditional programming techniques of arranging for primarily if-then statements. Loosely modeled after the neuronal structure of the brain, such systems interact with the environment and produce conse-

quences, which modify the probability of particular responses being emitted. Donahoe and Palmer (1994) have explored the use of neural network models using a "biobehavioral" approach. In this approach, the computer is programmed as described above but also adheres to more molar aspects of the brain (i.e., neurons projecting to various structures in the brain, for example, sensory input being projected to the sensory association cortex, the motor association cortex then to the primary motor cortex). Hutchison (1998) has used a strictly behavioral neural network model to directly examine the acquisition of verbal behavior. In both cases the researchers are only seeking sufficiency arguments and interpretations of complex behavior. These computer tools, coupled with protocol analyses, appear to be the best methodology behavior analysts currently have to aid in understanding the complexities of verbal behavior.

REFERENCES

- Austin, J., & Delaney, P. (1998). Protocol analysis as a tool for behavior analysis. *The Analysis of Verbal Behavior*, 15, 41–56.
- Catania, A. C. (1998). *Learning*. Upper Saddle River, NJ: Prentice Hall.
- Critchfield, T. S., & Epting, L. K. (1998). The trouble with babies and the value of bathwater: Complexities in the use of verbal reports as data. *The Analysis of Verbal Behavior*, 15, 65–74.
- Donahoe, J. W., & Palmer, D. C. (1994). *Learning and complex behavior*. Needham Heights, MA: Allyn and Bacon.
- Ericsson, K. A., & Simon, H. A. (1993). *Protocol analysis: Verbal reports as data*. Cambridge, MA: MIT Press.
- Hayes, S. C., White, D., & Bissett, R. T. (1998). Protocol analysis and the "silent dog" method of analyzing the impact of self-generated rules. *The Analysis of Verbal Behavior*, 15, 57–64.
- Horne, P. J., & Lowe, C. F. (1996). On the origins of naming and other symbolic behavior. *Journal of the Experimental Analysis of Behavior*, 65, 185–243.
- Huber, S. (1997, May). *The use of protocol analyses in research*. Panel discussion conducted at the 23rd annual convention of the Association for Behavior Analysis, Chicago.
- Hutchison, W. R. (1998). Computer simulations of verbal behavior for research and persuasion. *The Analysis of Verbal Behavior*, 15, 117–120.
- McIlvane, W. J., & Dube, W. V. (1990). Do stimulus classes exist before they are tested? *The Analysis of Verbal Behavior*, 8, 13–18.
- Palmer, D. C. (1991). A behavioral interpretation of memory. In L. J. Hayes & P. N. Chase (Eds.), *Dialogues on verbal behavior* (pp. 259–288). Reno NV: Context Press.
- Potter, B., Huber, S., & Michael, J. (1997). The role of mediating behavior in selection-based responding. *The Analysis of Verbal Behavior*, 14, 41–56.
- Skinner, B. F. (1953). *Science and human behavior*. New York: Knopf.
- Skinner, B. F. (1957). *Verbal behavior*. Englewood Cliffs, NJ: Prentice Hall.
- Wulfert, E., Dougher, M. J., & Greenway, D. E. (1991). Protocol analysis of the correspondence of verbal behavior and equivalence class formation. *Journal of the Experimental Analysis of Behavior*, 56, 489–504.