Word count: 1208

Number of references: 9

# Protocol Analysis

Fernand Gobet

To appear in Tim Bayne, Axel Cleeremans, and Patrick Wilken (Eds.), The Oxford Companion to Consciousness. Oxford University Press

Send correspondence to:

Fernand Gobet, PhD
Centre for the Study of Expertise
School of Social Sciences
Brunel University
Uxbridge, Middlesex, UB8 3PH
United Kingdom

Phone: +44 (1895) 265484
Fax: +44 (1895) 237573

fernand.gobet@brunel.ac.uk
http://people.brunel.ac.uk/~hsstffg/

Protocol Analysis

Protocol analysis is a research method in which verbal reports are elicited from participants and transcribed for qualitative and/or quantitative analysis. It can be used in isolation or jointly with the collection of other process data such as eye movements or keystrokes on a keyboard. With concurrent protocols, participants are requested to think aloud when carrying out a task, for example when solving a problem in physics or in logic, with the explicit instruction to report everything that comes to their mind but not to describe or explain their thoughts. This technique provides useful information with tasks that have a certain duration, say at least a few dozens seconds, but is more limited with visuo-spatial tasks or tasks that are of short duration. In these cases researchers often use retrospective protocols, where participants are asked to report, after completing a task, the thoughts and/or actions that occurred when carrying out the task. For example, participants may be asked first to carry out an arithmetic problem, and then to report the processes that took place. Both concurrent and retrospective protocols must carefully be distinguished from introspection (see *Introspection*), where participants, often trained with this purpose in mind, are asked to describe their mental states and theorize about them.

Duncker (1935/1945) was one of the first psychologists to use concurrent protocols, but perhaps the first really influential use of this technique was provided by De Groot (1946/1965) in his analysis of chess players' thinking processes. De Groot identified an important weakness of protocol analysis, namely that protocols are often incomplete. There are several reasons for this: some aspects of the thought process do not reach the threshold of awareness; conscious thoughts may be too fast to be

verbalized; some thoughts cannot be immediately verbalized; and, finally, participants may suppress some steps of the thought process, for example to avoid the embarrassment of reporting mistakes in their thinking. In addition, the reporting of concurrent thoughts may slow down the thinking process itself. In spite of these deficiencies, De Groot concluded that protocols often provide a satisfactory outline of the macro-structure of the thought process, and that their degree of completeness may be assessed by asking participants to which extent the protocol faithfully reproduces this process.

Nisbett and Wilson (1977) argued that retrospective protocols suffer from a number of weaknesses, including the use of general knowledge to fill in gaps in memory, rationalization, and even lack of awareness of critical stimuli. It should be noted that the severity of these weaknesses is affected by the way reports are elicited, and increases with the time lag between the thoughts and their report.

Using Newell and Simon's (1972) information processing framework, Ericsson and Simon (1980; 1993) provided an extensive discussion of the advantages and disadvantages offered by verbal protocols. The originality of their approach is to validate their analysis with a theoretical model that clearly specifies when protocols will be reliable and when they will not. Their starting point is that people are conscious of only those things that are stored in short-term memory (STM). When this information is verbal in nature, the action of providing verbal protocols simply amplifies information that is already in short-term memory. When information is non-verbal, talking aloud adds a recoding stage, which slows the process down and introduces translation errors. A consequence of this theory is that retrospective

protocols suffer from the fact that information in not in STM any more, in particular when recall occurs a long time after completing the task. Ericsson and Simon's information processing model also postulates that encoding unfamiliar information into long-term memory (LTM) requires several seconds, which means that only a subset of the thoughts having taken place during the task will be encoded in LTM. Similarly, asking participants to provide explanations of their behaviour requires them to use information that was never in STM, and thus offers little reliability. Hence, retrospective protocols can be recommended only when they are taken a few seconds after the task and when it is plausible to assume that STM cues will enable access to information stored in LTM.

Based on their theoretical analysis, Ericsson and Simon offered practical advice as to how collect reliable protocols. Recommendations include giving warm-up tasks to the participants and sitting behind the participants to avoid contaminating the protocols with social interactions. The analysis of verbal protocols is notoriously complex and time consuming, and programs have been developed to facilitate this process. For example, Bhaskar and Simon (1977) describe SAPA, a semi-automatic protocol coding system used in their study of problem solving in thermodynamics.

Verbal protocols have sometimes been used in combination with other process data such as eye movements. For example, Carpenter, Just and Shell (1990) used concurrent protocols when participants solved the Raven's Progressive Matrices, a test of intelligence. This and similar studies suggest that the thought contents, as reported verbally in the concurrent protocols, correspond well with the eye-movement data. Similarly, De Groot and Gobet (1996) recorded eye movements in combination

with retrospective protocols. Chess players' eye movements were recorded when they looked at a position for five seconds; immediately after, the players attempted to recall where they had focused their attention during those five seconds. Results indicated a fair amount of individual differences, with some players having a remarkable memory of their eye movements, and others displaying poor recollection. An interesting result with respect to the study of consciousness was that players, after they had fixated the same square twice or more, tended to mention only the first fixation in their protocol. This is similar to the 'Ranschburg effect' identified in serial recall experiments, where people have difficulty in recalling items that are repeated in the list of stimuli to memorize.

Nowadays, mainstream cognitive psychology has accepted the use of concurrent protocols, but is still reticent about the use of retrospective protocols or introspection. Even critiques of the methodology (e.g., Wilson, 1994) agree that concurrent protocols offer a valid means for studying the contents of consciousness. Such protocols offer a high density of microdata and make it possible to capture the dynamics of sequences of mental states. It is also generally admitted that data gathered from concurrent protocols can be used as any other data to test theories, and some (symbolic) computational process models have in fact been able to reproduce the contents of verbal protocols in great detail (e.g., Newell & Simon, 1972).

However, there is also agreement that verbal protocols are ill-suited for studying nonconscious information. In addition, as anticipated by De Groot, there remains the issue of what is missing from the protocols—that is, to what extent nonconscious information plays a key role in the tasks usually studied with this methodology. There

is little doubt that such information affects behaviour, even in problem-solving tasks where information is often considered as fairly conscious, as shown for example by the substantial amount of experimental evidence that has been collected in the last decades on subliminal priming, automatization, and implicit learning.

Taking into account these strengths and weaknesses, the conclusion is that verbal protocols can offer very useful information for testing theories of human cognition, and that they are best used in conjunction with other methods (either experimental or computational) that ensure cross validation.

**References**

Bhaskar, R., & Simon, H. A. (1977). Problem solving in semantically rich domains: An example from engineering thermodynamics. *Cognitive Science, 1*, 193-215.

Carpenter, P. A., Just, M. A., & Shell, P. (1990). What one intelligence test measures: A theoretical account of the processing in the Raven progressive matrices test. *Psychological Review, 97*, 404-431.

De Groot, A. D. (1965). *Thought and choice in chess (first Dutch edition in 1946).* The Hague: Mouton Publishers.

De Groot, A. D., & Gobet, F. (1996). *Perception and memory in chess: Heuristics of the professional eye*. Assen: Van Gorcum.

Duncker, K. (1945). On problem solving (original German edition in 1935). *Psychological Monographs, 58*(5, Whole No. 270).

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

Ericsson, K. A., & Simon, H. A. (1993). *Protocol analysis: Verbal reports as data* (2nd ed.). Cambridge, MA: MIT Press.

Newell, A., & Simon, H. A. (1972). *Human problem solving*. Englewood Cliffs, NJ: Prentice-Hall.

Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we know: Verbal reports on mental processes. *Psychological Review, 84*, 231-259.

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