our theory can be extended to account for them. We will then turn to the issue of completeness of verbally reported information (Chapter 3) and review the new evidence in three particularly active task domains. In a final section we will mention recent contributions to the techniques for encoding protocols, raise general issues, and propose directions for future research. For other commentaries on recent developments in verbal protocol analysis see Ericsson and Simon 1985, 1987.

OVERVIEW

From the beginnings of psychology as a science, investigators, impelled by the difficulties of relying wholly on external observation in studying mental processes, have questioned subjects about their experiences, thought processes, and strategies. Claims for the validity of such verbal accounts were based primarily on the notion that individuals had privileged access to their experiences; as long as they were truthful, their reports could be trusted. With the emergence of behaviorism the untestable assumptions of trustworthiness and privileged access were rejected as unscientific, and verbal reports were treated like any other kind of observable behavior — as verbal behavior — but, beyond simple "yes" and "no" responses, were used relatively little in experimental research.

Later, when information-processing accounts were given of human performance in different tasks, these accounts essayed to reproduce all aspects of observable behavior. Verbal reports, in this view, were produced by the same cognitive processes that produce more traditional performance data, such as speed of button presses and sequences of eye fixations.

Normally verbal reports are elicited by asking a subject a specific question. To answer, the subject has to comprehend the question and transform it to retrieval cues that select the relevant information from the vast amount of information in long-term memory. In addition the subject has to put the retrieved information into a sequential form that allows the generation of a coherent series of verbalizations.

A critical problem with subjects' general verbal descriptions of their cognitive processes and experiences is that such reports do not relate clearly to any specific observable behavior. Even where subjects are asked to report on their cognitive processes used during many trials of an experiment, we cannot rule out the possibility that the information they retrieve at the time of the verbal report is different from the information they retrieved while actually performing the experimental task. To avoid this problem of accessing information at two different times — first during the actual cognitive processing and then at the time

of report — we proposed that, whenever possible, concurrent verbal reports should be collected, so that processing and verbal report would coincide in time (Ericsson & Simon, 1980; this book). Let us turn to a general description of how information is accessed and heeded during cognitive processes.

Concurrent Verbalization

Since the time of Aristotle, thinking has been viewed as a temporal sequence of mental events. A recent review of the history of the study of thinking shows that this assumption has never been seriously questioned (Ericsson & Crutcher, 1991). Thought processes can thus be described as a sequence of states, each state containing the end products of cognitive processes, such as information retrieved from long-term memory, information perceived and recognized, and information generated by inference. The *information* in a state is relatively stable and can thus be input to a verbalization process and reported orally. However, the retrieval and recognition *processes* delivering the end products, the information, to attention cannot be reported.

The standard method for getting subjects to verbalize their thoughts concurrently is to instruct them to "think aloud" (see Appendix for a complete instruction). With this instruction subjects verbalize new thoughts and generate intermediate products as these enter attention. For example, a subject given the task of mentally multiplying 24 by 36 while thinking aloud might verbalize: "36 times 24," "4 times 6," "24," "4," "carry the 2," "12," "14," "144," and so on. It is important to note that subjects verbalizing their thoughts while performing a task do *not* describe or explain what they are doing — they simply verbalize the information they attend to while generating the answer.

When subjects verbalize directly only the thoughts entering their attention as part of performing the task, the sequence of thoughts is not changed by the added instruction to think aloud. However, if subjects are also instructed to describe or explain their thoughts, additional thoughts and information have to be accessed to produce these auxiliary descriptions and explanations. As a result, the sequence of thoughts is changed, because the subjects must attend to information not normally needed to perform the task.

There is a dramatic increase in the amount of behavior that can be observed when a subject is performing a task while thinking aloud compared to the same subject working under silent conditions. A brief instruction to think aloud usually suffices to bring about this major change in observable behavior. In light of the fact that subjects do not need to practice before being able to "think aloud" one infers that this verbal reporting is consistent with the structure of their normal

cognitive processes and their general skills for verbalizing needed information. Although spontaneous thinking aloud is rare in the everyday life of normal adults, adults normally engage in many other forms of verbalization relevant to thinking. For example, in many educational settings a student or teacher will explain verbally how to solve a problem or task in front of observers. Adults are often asked to describe how to do something. Decisions and judgments are often discussed publicly and challenged, thereby requiring individuals to justify and even rationalize their choice and evaluations.

Explanations, descriptions, justifications, and rationalizations are socially motivated verbalizations generated to communicate to one or more listeners. Successful communication with listeners requires additional processes to attain coherence and take account of differences in background knowledge. Hence social verbalizations may be quite different from the sequences of thoughts generated by subjects themselves while solving problems, performing actions, and making evaluations and decisions.

When an investigator instructs a subject to think aloud, some subjects may misunderstand the instruction and produce instead the more common social communication, explaining or describing the process to the experimenter. The scientists who independently discovered "talk aloud," "think aloud," and "reflection parlé" early in this century employed instructions (for more details see Chapter 1) that did not address this danger of confusing different modes of verbal reporting. In our book we outline a number of practices that reduce this problem and may often eliminate it. First the experimental situation is arranged to make clear that social interaction is not intended, and the experimenter is seated behind the subject and hence is not visible. The "think aloud" instruction explicitly warns the subjects against explanation and verbal description.

Second, after the instruction is presented, the subjects are given practice problems in which it is easy to verbalize concurrently and from which they attain familiarity with the normal content of think-aloud verbalizations. Subjects can then be given the instruction again and may further practice problems until they are acquainted with the procedure and verbalize their thoughts concurrently. Third, the social interaction between subject and experimenter is minimized. To remind subjects to think aloud, the experimenter tells subjects to "keep talking" instead of making the social request, "tell me what you are thinking." Finally, subjects are always told to focus on completing the presented task, the instruction to think aloud being secondary. Only when the thinking-aloud subjects remain completely focussed on the task can we expect the same sequences of thoughts as in the silent condition. A prerequisite for subjects being able to focus on a task is that the presented task has a clear focus. Garner (1987) discusses an application with concurrent verbalization where the task was left unspecified and vague.

Under such circumstances it appears that subjects generate a better-defined task and that this process of defining a more specific task is influenced by extraneous social factors. However, the characterization of clearly defined tasks is prerequisite for studying reproducible performance whether concurrent reports are requested or not (Ericsson & Oliver, 1988).

There are some temptations for experimenters to use additional procedures that may compromise pure concurrent verbalization. Concurrent verbalizations, unlike social verbalizations, often lack coherence and contain frequent disjointed sequences without explicit relations between the thoughts. Think-aloud protocols leave unanswered how the solution was generated in detail and why a given method was adopted among many possible methods. In contrast, a verbal account that explains and describes each step will satisfy the social demands for coherence and reasonable completeness. However, to guarantee a close correspondence between the verbal protocol and the actual processes used to perform the task, this urge toward coherence and completeness must be resisted.

If we are committed to recording only the actual concurrent sequences of thoughts, then in order to infer the underlying cognitive processes, we need methods for encoding and analyzing the sequences. For a given task only a limited number of possible sequences of thoughts will generate the correct answer efficiently. By task analysis, using the constraints defined by an individual's pre-existing skills and knowledge, one can enumerate the logically possible thought sequences that can produce the correct answer. For example, the mental multiplication task mentioned earlier can be solved by many short-cut methods that differ from the common right-to-left paper-and-pencil method. Recognizing that 24 times 36 equals (30-6) (30+6) would allow a subject to derive $30^2 - 6^2 = 900-36 = 864$, or that 24 times 36 equals (2x12) (3x12) would yield $6x12^2 = 6x144$, which is an easier product to calculate mentally. For these two short-cut methods as well as many others, one can specify the sequence of intermediate products that could be verbalized concurrently as the answer is generated. The recorded verbalizations can then be encoded in categories defined *a priori* by the task analysis, and can be matched against the alternative possible sequences of thoughts. Often all but one hypothesis will be rejected due to mismatches between observed (verbalized) thoughts and hypothesized thoughts.

The mere fact that one can derive predictions *a priori* from a task analysis that match the information actually verbalized is an important validation of the entire framework. The task analysis supports additional methods of validation. Other types of observation, such as solution times and eye fixations, can be recorded for the solutions generated with concurrent verbalizations. The hypothesized alternative solutions yield predictions for these other types of data and, as we show in Chapter 4, provide validation for the verbalizations from these other types of data.