Evidence that 'thinking aloud' constitutes an externalization of inner speech¹

JOHN BENJAFIELD, *Institute for Cognitive Studies, Rutgers, The State University, Newark, N.J., 07102*

Thinking aloud as a method of studying problem solving has been seen as a variant of introspection. In the present study, introspective and thinking aloud protocols were compared. The thinking aloud protocols were more elliptical, present-oriented, and contained more indefinite referents. Reasons were given for interpreting these findings as indicating that thinking aloud makes overt the normally covert process of inner speech.

Having Ss think aloud has recently become the method of choice for students of the problem-solving process (Miller, Galanter, & Pribram, 1960, pp. 192-194). Critics of the method (e.g., Ray, 1967, pp. 88-89), including some who have used it themselves (e.g., Duncker, 1945, p. 11), have seen thinking aloud as a form of nonreflexive introspection subject to the same flaws as classical introspection. If this position is true, then the protocols obtained through the instruction "think aloud" should be the same as those obtained through introspection instructions, although the former should perhaps be more complete.

The preceding point of view is based on the assumption that thinking aloud involves S using language to express or describe his thought to E, and fails to consider the implications of the possibility, suggested by Vygotsky (1962) and Luria (1961), that Ss utilize language in the form of inner speech to plan and organize their thinking. It is possible that thinking aloud instructions made overt this normally covert process.

If thinking aloud does represent inner speech, then protocols obtained using the two different instructions should differ not only in content, but also formally, along the

dimensions suggested by Werner & Kaplan (1963, pp. 328-330). Thinking aloud protocols should differ from introspective protocols by being: (a) more condensed and elliptical, as measured by the number of fragmentary utterances; (b) a representation of an ongoing process rather than a description of a past process, as measured by the number of present vs past tense verbs; and (c) composed of more indefinite referents, as measured by the number of pronouns as opposed to nouns.

SUBJECTS AND PROCEDURE

Thirty-two Rutgers undergraduates served as Ss. Eight men and eight women were assigned to each of the two (thinking aloud and introspective) experimental conditions.

All Ss were given Katona's (1927) tetrahedron problem to solve. The problem requires S to correctly place the two identical halves of a tetrahedron together. In the present version of the problem, Ss were told: "Here are two identical objects. Each object is composed of two identical trapezoids, two identical triangles, and a square. I want you to think of a way of putting these two objects together so as to make an object with four surfaces. You are not allowed to touch the objects, but must rely on thought to imagine their correct placement."

Subjects in the thinking aloud condition were told in addition: "While you go about solving the problem, try to think aloud—that is, speak while you are thinking." Ss given the introspection condition were told: "While you go about solving the problem, think silently. From time to time, I will ask you what you have been thinking about."

Subjects in the thinking aloud condition were given 5 min to work on the problem. Ss in the introspection condition were allowed to think for 1 min, and then were interrupted and asked: "What have you been thinking about for the last little while?" This

procedure was repeated until S had spent a total of 5 min thinking. E wrote down everything all Ss said.

RESULTS

No S solved the problem in the time allotted. The protocols were analyzed into the following components: (a) number of words; (b) number of descriptive statements; (c) number of fragmentary utterances, including isolated phrases, single words, and incomplete sentences; (d) number of nouns; (e) number of pronouns; (f) number of present-tense verbs; (g) number of past-tense verbs. The results are given in Table 1.

Because of markedly skewed distributions and heterogeneity of variance, U tests were used to assess the statistical significance of the results. Ss given the thinking aloud task generated more words, more fragmentary utterances, more pronouns, and used the present tense more than Ss given the introspection task. Conversely, Ss in the introspection condition used more past-tense verbs than Ss in the thinking aloud condition.

DISCUSSION

The fact that Ss in the thinking aloud condition generated more words than those in the introspection condition jibes with the notion that thinking aloud protocols are more complete and contain more information than introspective protocols. In general, however, the picture that emerges is that thinking aloud protocols differ from introspective protocols by being more present-oriented, having a more elliptical form, and containing more indefinite referents.

An objection to interpreting these findings as indicating a close relationship between inner speech and thinking aloud might be that the instruction "think aloud" forces S to carry out two activities at once: thinking and describing his thinking. This, it might be argued, forces S's descriptions to be attenuated and indefinite. This argument is probably not correct for the following reasons. (1) Ss tend to use the present tense more in the thinking aloud condition than in the introspection condition. If both groups of Ss are describing what they have already thought, why do they not use the past tense equally? In fact, Ss in the introspection condition use the past tense to a greater extent. (2) A typical excerpt from an introspective protocol is: "I tried putting the triangular surfaces together, but that gave me eight sides." A typical excerpt from a thinking aloud protocol is: "Put them end

Table 1
Comparison of Thinking Aloud and Introspective Protocols

Condition	Thinking Aloud		Introspection		U	p
	Median	Range	Median	Range		
Number of words	96.5	60-160	88	60-170	70	<.05
Descriptive Statements	6	2-10	5	4-11	115.5	ns
Fragmentary Utterances	8	4-20	1	0-9	11	<.001
Nouns	3	2-12	4	0-14	121	ns
Pronouns	6	1-11	3	0-10	73	<.05
Present Tense Verbs	13	7-21	8	1-21	76	<.05*
Past Tense Verbs	1	0-5	4	0-8	50	<.002

* corrected for ties

to end . . . No." The former is clearly a description of previously performed cognitive operations. However, the latter takes the form of a command, or plan to conduct an operation ("Put them end to end"), followed by a pause (. . .) during which S is presumably imagining the planned operation, followed by an evaluation of the results of the operation ("No"). (It should be remembered that no S was permitted to touch the problem materials, and therefore each S had to imagine any operations he wished to carry out on the materials).

The structure of the thinking aloud excerpt is very much like the structure of Miller, Galanter, & Pribram's (1960) TOTE mechanism. In the present case, the planning and evaluative phases of the mechanism are carried out in inner speech. This conception fits well with the notion that inner speech

plays an executive role in the conduct of cognitive operations (Luria, 1961).

REFERENCES

- DUNCKER, K. On problem solving. *Psychological Monographs*, 1945, 58, (Whole No. 270).
 KATONA, G. Eine kleine anschauungsaufgabe. *Psychologische Forschung*, 1927, 9, 159-162.
 LURIA, A. R. *The role of speech in the regulation of normal and abnormal behavior*. New York: Liveright, 1961.
 MILLER, G., GALANTER, E., & PRIBRAM, K. *Plans and the structure of behavior*. New York: Holt, 1960.
 RAY, W. *The experimental psychology of original thinking*. New York: MacMillan, 1967.
 VYGOTSKY, L. S. *Thought and language*. Cambridge: M.I.T. Press, 1962.
 WERNER, H., & KAPLAN, B. *Symbol formation*. New York: Wiley, 1963.

NOTE

1. This research was conducted under the support of a postdoctoral fellowship from the National Research Council of Canada.

digm permits S to anticipate the response item and is thus essential in the repeated presentation method of paired-associate learning. However, the studies under consideration here are all based on a single learning trial and, therefore, do not require the stimulus-alone presentation. In the study described here, this slide was eliminated and the interstimulus interval was accordingly shortened.

Another change from the original studies was based on informal observations of skin resistance patterns. Ss appear to have a relatively restricted range of GSR values when they are not relaxed. Typically, some Ss have few or no scorable GSRs. To reduce this problem in the present study, every effort was made to encourage the S to relax.

SUBJECTS

One-hundred eighty-nine undergraduates enrolled in the introductory psychology courses served as Ss. Of these, 96 were females and 93 were males. The Ss were randomly assigned to one of four conditions.

STIMULI AND DESIGN

Six of the eight items used by Kleinsmith & Kaplan (1963) were used: KISS, RAPE, VOMIT, EXAM, MONEY, SWIM. The response items were single digits from 2 to 7.

A slide projector with an exposure time of 4 sec was used to present the stimuli which were typed in capitals. Two color slides, containing four of six colors used, were interpolated between paired-associate slides in training and between stimulus-item slides in recall. A stimulus-only slide was not used in training in this study, and hence the interstimulus interval was 12 sec (in contrast to 16 sec in the Kleinsmith and Kaplan study). The first slide in the training list consisted of the word READY followed by two color slides.

The order of items in training was controlled by the use of four lists designed such that any item was followed by any other given item in only one list. The serial position of each item was also controlled. Two sets of responses (single-digit numbers) were used for each paired associate; thus, for half the Ss in each condition a particular stimulus was paired with one digit, while for the other half a different digit was used. Consecutive items were not paired with consecutive digits.

The orders of the items in the recall lists were arranged so that no item appeared in the same serial position in training and recall. Further, the number of items that intervened between the presentation of a given item in training and its presentation in recall was controlled such that for half the items that distance was three or four and for the other half it was six or seven.

PROCEDURE

The Ss were given a single learning trial

The arousal-retention interval interaction revisited: The effects of some procedural changes¹

RACHEL KAPLAN and STEPHEN KAPLAN, *University of Michigan, Ann Arbor, Mich. 48104*

The arousal/time interaction obtained in the Kleinsmith and Kaplan studies apparently can be attenuated by set or motivational variables. A six-item paired-associate list was presented a single time to each of 189 Ss, who were then tested for recall at varying times (immediate, 6 and 18 min, or 2 days). The high arousal items showed significant reminiscence at 6 min, but the overall effects were not as strong as in previous studies. The role of motivational sets under incidental learning conditions is discussed.

The Kleinsmith & Kaplan (1963, 1964) studies of arousal and time interval in paired-associate learning appear to have played a role in the return of reminiscence to

verbal learning (Keppel, 1968). Since findings of this kind raise certain difficulties for most current theories of learning, establishing the generality of this phenomenon is of considerable importance. Levonian (1966) has studied arousal and reminiscence in the classroom with comparable results and his findings have been replicated by Kaplan, Kaplan, & Sampson (1968).

The present study concerns certain procedural modifications of the 1963 study. Since the overall recall level in that study was relatively low, the present study utilized only six items instead of eight. (The 1964 study where the stimulus items were nonsense syllables also employed six items.) Furthermore, since the method used in the early studies originates from the traditional paired-associate procedure, the stimulus element is presented alone prior to presentation of the paired associate. The stimulus-alone presentation in that para-