

When can we introspect accurately about mental processes?

RONALD T. KELLOGG
University of Missouri, Rolla, Missouri 65401

Some theorists have suggested that the cognitive processes determining a person's performance in a given task are unconscious, making introspection a poor research tool for studying these processes. Others have argued that the relevant processes usually are consciously controlled and can be detailed by asking the person to introspect. Here, a synthesis of these two positions, a dual-factor approach, is proposed. Some of the processes involved in achieving a cognitive goal, such as learning a new concept, are viewed as unconscious and automatic; however, other processes are intentionally allocated conscious attention, in certain tasks, to accomplish the goal. To illustrate this dual-factor position, evidence is presented in support of the view that when concept learning occurs solely by automatic frequency processing, introspective reports are inaccurate, but when the nature of the task prompts intentional hypothesis testing, introspective reports are accurate, revealing clues that subjects engage in a conscious hypothesis-testing strategy.

The old, yet, once again timely, debate over the validity of introspection has been recently revived in two important articles. Ericsson and Simon (1980) argued that introspective responses can yield important insights into the nature of cognitive processes underlying task performance. Their article was, in part, a reaction to an influential article by Nisbett and Wilson (1977), who argued that the relevant mental processes are unconscious and, therefore, never truly open to introspection. In the present article, I suggest that introspection can be either a valid or an invalid research tool, depending on the demands of an experimental task.

By way of empirical example, I argue that, depending on the demands of a task, a person may accomplish some cognitive goal either by conscious attentional processes or by unconscious automatic processes. To illustrate this dual-factor view of cognition, I briefly present two sets of experiments on concept learning. Several investigators have reported evidence indicating that both conscious, or explicit, processes (hypothesis testing) and unconscious, implicit processes (feature frequency or strength processing) are important in concept learning (Reber, Kassin, Lewis, & Cantor, 1980; Anderson, Kline, & Beasley, Note 1; Bourne, Robbins, & Kellogg, Note 2; Elio & Anderson, Note 3). Although the distinction between conscious and unconscious processing has been examined in numerous areas of research (e.g., Mandler, 1975; Posner & Snyder, 1975; Shevrin & Dickman, 1980; Shiffrin & Schneider, 1977;

Wilson, Hull, & Johnson, 1981), here I focus on concept learning.

In the case of concept learning, feature-frequency processing presumably always occurs whenever a person perceptually encodes the stimulus features of a concept exemplar. Frequency processing isolates the defining features of a concept, if there are any, by passively cumulating the relative number of times that the various stimulus features occur among the exemplars of a particular category (Bourne, Ekstrand, Lovallo, Kellogg, Hiew & Yaroush, 1976). Frequency processing seems to be an automatic, unconscious learning operation (Hasher & Chromiak, 1977). But the resulting feature-frequency distribution for a particular category can be used to generate consciously (1) reasonable hypotheses about defining features (Kellogg, 1980a), (2) feature-frequency estimates (Kellogg, 1981), or (3) typicality judgments (Chumbley, Sala, & Bourne, 1978; Goldman & Homa, 1976).¹

In addition to frequency processing, a person may at times actively attend to sampling, testing, and storing in memory hypotheses about the defining features of the concept (Levine, 1975). Hypothesis testing refers to a conscious strategy (Kintsch, 1970, p. 359). But even when the nature of the task does not encourage hypothesis testing, concept learning can still succeed by means of frequency processing (Bourne et al., Note 2). This happens whenever the person's attention is allocated to processes other than hypothesis testing, such as perceiving and memorizing concept exemplars (Reber et al., 1980).

According to this dual-factor position, when concept learning relies solely on automatic frequency processing, introspection should be inaccurate. But when it draws on both frequency and hypothesis processing, intro-

Preparation of this article was supported by a special grant to the author from the University of Missouri. I thank Lyle E. Bourne, Jr., Cathy Riordan, and James Shanteau for their comments on an earlier version of the manuscript. Request for reprints should be sent to Ronald T. Kellogg, Psychology Department, University of Missouri, Rolla, Missouri 65401.

spection should yield information about hypothesis testing.

Introspection can be operationally defined in multiple ways (Ericsson & Simon, 1980). Any procedure that asks the subject to think about his or her mental processes is a form of introspection. The verbal protocol procedure (thinking aloud) is a standard form of introspection, but this method may alter the natural processes employed in a given task. To avoid the problem of impingement, relatively unobtrusive methods of introspection were used in the studies to be discussed here. In one set of experiments, introspective ratings about stimulus processing were solicited after the experimental task was completed. A positive correlation between the introspective measures of how much subjects claimed to have attended consciously to the stimuli and measures of how much they actually processed the stimuli (e.g., recognition performance) indicates that they were consciously aware of processing the stimuli. Lack of a correlation between these measures suggests the subjects processed the stimuli automatically and unconsciously. Correlations between retrospective ratings and behavioral performance measures have been successfully used as an operational definition of introspection in memory (Kellogg, 1980b) and attribution research (Nisbett & Wilson, 1977; Wilson et al., 1981).

In a second set of experiments, introspection was operationally defined in terms of a recognition probe for a trial event. This is a form of directed introspection (Ericsson & Simon, 1980). If the event was consciously attended to on Trial *n*, then it should be accurately stored in short-term memory on Trial *n* + 1. To mitigate the effects of the probes altering the natural processes employed in the task, probes were presented randomly on only a subset of the trials.

CONCEPT LEARNING ON AN UNATTENDED CHANNEL

In the first set of relevant experiments, task demands were manipulated to allow attended processing of concept exemplars in single-task conditions and to prevent it in dual-task conditions. The reader should refer to the original report for details of the method and results (Kellogg, 1980b). Difficult mental multiplication was aurally presented as a primary task to subjects assigned to the dual-task conditions; while subjects focused on the mental multiplication, exemplars of a category of faces were presented on a secondary channel. Only the faces were presented to subjects in single-task face conditions. And, only multiplication problems were presented in single-task multiplication conditions. These conditions provided a baseline of multiplication accuracy to assess whether the faces interfered with primary task performance in the dual-task conditions.

Intention to process the faces in dual-task conditions

was manipulated by instructions. Those in the single-task conditions were instructed to devote full attention to either the faces or the multiplication. After viewing the exemplars, all subjects who viewed faces were asked to introspect about how much conscious attention they had paid to the faces. Finally, a surprise recognition test, designed to measure the degree of learning of the concept relating the faces, was given.

Recognition performance in the single-task face conditions appeared to be determined by conscious attentional processes, such as hypothesis testing, as well as by automatic frequency processes. In contrast, performance in the dual-task conditions was totally mediated by automatic mechanisms. Three findings supported this conclusion. First, processing the faces failed to interfere with the primary task. Second, all dual-task conditions, which varied with regard to intention to process the faces, exhibited statistically equivalent degrees of conceptual memory,² as would be expected if the learning processes employed in dual-task conditions were automatic. Finally, all dual-task conditions (automatic processes only) showed significantly less conceptual memory than did the single-task face conditions (automatic plus conscious processes).

The introspective ratings of how much attention the subjects gave to faces were significantly positively correlated (Experiment 4, $r = .55$; Experiment 5, $r = .42$) with their degree of conceptual memory in the single-task conditions only. The more these subjects claimed to have attended, the more they learned and used the concept. In contrast, the introspective ratings of subjects in the dual-task conditions bore no relation to their conceptual recognition scores. When subjects did not attend to the stimuli, their attention ratings were essentially random guesses about how much they processed the stimuli. Thus, introspective reports reliably indexed degree of stimulus processing when conscious attentional processes were involved, but not when only automatic processes were implicated, as predicted by the dual-factor view.

It is important to note that the amount of training given on the faces was equivalent for the single- and dual-task conditions. As many investigators have pointed out (e.g., Shiffrin & Schneider, 1977), with extensive practice on a task, the relevant processes can become unconscious and automatic. The above results are important because they show that, with practice held constant, both conscious and unconscious learning processes can be observed by manipulating the demands of the task.

One might object to the above conclusions on the grounds that the introspection ratings were collected several minutes following presentation of the last face. With such a long delay before introspection, information concerning conscious processes might have been forgotten. Although this criticism is plausible, it is unclear why introspection succeeded for subjects in the single-

task conditions and yet failed for those in the dual-task conditions. The delay for both conditions was equivalent.

PROBING SHORT-TERM MEMORY DURING CONCEPT LEARNING

Still, a strong case for the dual-factor position on introspection requires that short-term memory be probed immediately after the concept learning processes occur. A recent series of experiments addressed this issue. Again, for details, the original articles should be consulted (Kellogg, Robbins, & Bourne, 1978; Bourne et al., Note 2). We examined whether a person attends sufficiently to testing a hypothesis on Trial n to result in the hypothesis being stored in short-term memory on Trial $n + 1$ in a classification learning task. Short-term memory was probed for stimulus, response, and feedback information, as well as for hypotheses. But here, hypothesis retention is of key interest. On a certain percentage of the trials, each subject was required to select (indicate by manually interacting with a computer terminal) his or her current hypothesis(es). Then, on a small number of randomly selected trials, the subject was presented a two-alternative forced-choice recognition probe of a hypothesis selected on the preceding trial. The task demands varied in terms of the percentage of trials that required subjects to indicate their current hypothesis.

We predicted that subjects could be encouraged to allocate conscious attention to hypothesis testing when hypotheses are consistently selected. If so, a hypothesis recognition probe should reveal an accurate record of the hypothesis selected on the preceding trial. In contrast, under task demands that do not call for hypothesis selections frequently, hypothesis testing should not be attended to very much, if at all, leading to the prediction that hypothesis recognition performance should be poor.

As expected, we obtained accurate hypothesis recognition in one of the four situations investigated (Bourne et al., Note 2). The data showed that accurate hypothesis recognition (mean proportion of correct responses = .90) was obtained only when the task heavily emphasized hypotheses by requiring the subject to select his or her current hypothesis on every trial. When hypothesis selection occurred on 25%, 50%, and 75% of the trials, the overall mean for hypothesis recognition was .68, .73, and .72, respectively.³ Strikingly, in these three conditions hypothesis recognition was usually no better than chance (.50) on negative classification trials. On these trials, the stimulus failed to contain the hypothesized defining feature of the concept, leading to the judgment that the stimulus was not an instance of the concept. Despite the differences in hypothesis memory across conditions, concept learning, as measured by accurate classification performance, was observed in all four situations. This constellation of

finding is consistent with a dual-factor view of concept learning, but it is hard to explain on a straightforward hypothesis-testing view.

The degree of practice in the classification task was the same for all hypothesis-selection conditions. It would make little sense, therefore, to argue that hypothesis recognition was poor because the hypothesis-testing process had become automatized over time. Also, such an argument would predict that hypothesis recognition should be poorest when hypothesis selection occurred on 100% of the trials. The results indicated the opposite.

We concluded that if a task places heavy emphasis on hypotheses, then subjects allocate conscious attention to testing hypotheses, leaving an accurate record of this activity available for introspective recognition probes. If, however, hypotheses are deemphasized by the experimenter's asking for them on less than 100% of the trials, then most subjects learn the concept primarily by an automatic feature-frequency process. In these cases, introspective probes reveal little, if any, record of conscious hypothesis testing.

CONCLUSION

The two sets of experiments discussed here indicate that introspection may or may not reveal a subject's awareness of concept learning processes, depending on the constraints of the task. If the task places emphasis on allocating conscious attention to learning processes, then introspection should reveal the details of this process. But if the task does not encourage the subject to allocate conscious attention to such processes, then performance may be mediated entirely by automatic learning processes, which leave no record available for introspective reports. In short, a general principle of cognition might be that the processes underlying task performance may be either conscious or unconscious, depending on the nature of the task.

It appears, therefore, that there are elements of truth in both of the seemingly contradictory positions outlined by Ericsson and Simon (1980) and by Nisbett and Wilson (1977). Instead of trying to prove that introspection is accurate, if properly probed, in all but the highly practiced task (Ericsson & Simon, 1980) or inaccurate in all tasks (Nisbett & Wilson, 1977), it may be more fruitful to consider the influence of task demands on the question of whether introspection can be a useful research tool.

REFERENCE NOTES

1. Anderson, J. R., Kline, P. J., & Beasley, C. M. *A general learning theory and its application to schema abstraction* (Tech. Rep. 78-2). University of Pittsburgh, Office of Naval Research, 1978.
2. Bourne, L. E., Jr., Robbins, D. W., & Kellogg, R. T. *Recognition failure for hypotheses in concept learning*. Manuscript submitted for publication, 1981.
3. Elio, R., & Anderson, J. R. *The effect of information representativeness and order of acquisition on schema formation*. Paper

presented at the meeting of the Psychonomic Society, St. Louis, 1980.

REFERENCES

- BOURNE, L. E., JR., EKSTRAND, B. R., LOVALLO, W. R., KELLOGG, R. T., HIEW, C. C., & YAROUSH, R. A. Frequency analysis of attribute identification. *Journal of Experimental Psychology: General*, 1976, **105**, 294-312.
- CHUMBLEY, J. I., SALA, L. S., & BOURNE, L. E., JR. Bases of acceptability ratings in quasinaturlistic concept tasks. *Memory & Cognition*, 1978, **6**, 217-226.
- ERICSSON, K. A., & SIMON, H. A. Verbal reports as data. *Psychological Review*, 1980, **87**, 215-251.
- GOLDMAN, D., & HOMA, D. Integrative and metric properties of abstracted information as a function of category discriminability, instance variability, and experience. *Journal of Experimental Psychology: Human Learning and Memory*, 1977, **3**, 375-385.
- HASHER, L., & CHROMIAK, W. The processing of frequency information: An automatic mechanism? *Journal of Verbal Learning and Verbal Behavior*, 1977, **16**, 173-184.
- KELLOGG, R. T. Feature frequency and hypothesis testing in the acquisition of rule-governed concepts. *Memory & Cognition*, 1980, **8**, 297-303. (a)
- KELLOGG, R. T. Is conscious attention necessary for long-term storage? *Journal of Experimental Psychology: Human Learning and Memory*, 1980, **6**, 379-390. (b)
- KELLOGG, R. T. Feature frequency in concept learning: What is counted? *Memory & Cognition*, 1981, **9**, 157-163.
- KELLOGG, R. T., ROBBINS, D. W., BOURNE, L. E., JR. Memory for intratrial events in feature identification. *Journal of Experimental Psychology: Human Learning and Memory*, 1978, **4**, 256-265.
- KINTSCH, W. *Learning, memory, and conceptual processes*. New York: Wiley, 1970.
- LEVINE, M. *A cognitive theory of learning*. Hillsdale, N.J.: Erlbaum, 1975.
- MANDLER, G. Consciousness: Respectable, useful, and probably necessary. In R. Solso (Ed.), *Information processing and cognition: The Loyola symposium*. Hillsdale, N.J.: Erlbaum, 1975.
- NISBETT, R. E., & WILSON, T. D. Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, 1977, **84**, 231-259.
- POSNER, M. I., & SNYDER, C. R. Attention and cognitive control. In R. L. Solso (Ed.), *Information processing and cognition: The Loyola symposium*. Hillsdale, N.J.: Erlbaum, 1975.
- REBER, A. S., KASSIN, S. M., LEWIS, S., & CANTOR, G. On the relationship between implicit and explicit modes in the learning of a complex rule structure. *Journal of Experimental Psychology: Human Learning and Memory*, 1980, **6**, 492-502.
- SHEVRIN, H., & DICKMAN, S. The psychological unconscious: A necessary assumption for all psychological theory? *American Psychologist*, 1980, **35**, 421-434.
- SHIFFRIN, R. M., & SCHNEIDER, W. Controlled and automatic human information processing: II. Perceptual learning, automatic attending, and a general theory. *Psychological Review*, 1977, **84**, 127-190.
- WILSON, T. D., HULL, J. G., & JOHNSON, J. Awareness and self-perception: Verbal reports on internal states. *Journal of Personality and Social Psychology*, 1981, **43**, 53-71.

NOTES

1. Concept usage, such as defining a concept, evaluating the typicality of instances, estimating feature frequencies, or classifying stimuli, may require conscious attention even though concept learning does not.

2. Conceptual memory was defined in terms of the difference in mean recognition confidence rating given to both old and new concept exemplars presented at test vs. the mean confidence rating given to new nonexemplars. In all conditions, subjects gave roughly equivalent confidence ratings to the old and new exemplars, indicating a lack of specific memory of old items.

3. These means are averages from three experiments reported by Bourne et al. (Note 2). Kellogg et al. (1978) reported comparable data for the 50% hypothesis-selection condition. Hypothesis recognition performance was statistically equivalent across pre- and postsolution trials. The means presented here are collapsed across the pre- and postsolution data.

(Received for publication May 12, 1981;
revision accepted November 23, 1981.)