Bulletin of the Psychonomic Society
1978, Vol. 11 (3), 197-199

Perception as hypothesis testing

TIMOTHY A. SALTHOUSE
University of Missouri, Columbia, Missouri 65201

and

WARREN L. DANZIGER
Washington University, St. Louis, Missouri 63103

Three experiments were conducted to investigate the previously reported finding that the accuracy of perceptual recognition decreases as the amount of experience with degraded versions of a visual stimulus increases. Since this result is apparently predictable only from the view of perception as an active process of hypothesis generation, this negative-effect-of-prior-experience phenomenon has important implications for theoretical conceptualizations of perception. None of the current experiments yielded any evidence of less accurate perceptual identification with increased number of incomplete versions of the stimuli when accuracy was assessed with a cumulative measure of identification accuracy, but two of the experiments did provide such evidence when accuracy was assessed with a conditional measure of accuracy. Consideration of the complete pattern of results led to the conclusion that there is no real evidence that perception is impaired because of early experience with ambiguous versions of a subject.

An apparently well-documented but largely ignored phenomenon which is of great importance to the theoretical conceptualization of the perceptual process is that prior impoverished experience with a visual stimulus impairs the subsequent identification of that stimulus. Independent studies (i.e., Bruner & Potter, 1964; Potter, 1966; Wyatt & Campbell, 1951) have reported that "if a subject is initially exposed to a blurred image that he cannot recognize, subsequent recognition of the image in clearer form is substantially delayed" (Bruner & Potter, 1964. p. 424).

Although certainly not a detailed explanation, an interpretation of this phenomenon can be derived from the conceptualization of perception as an active process involving the generation, testing, and verification of hypotheses. Perhaps Gregory (1970, 1974) is currently the strongest proponent of this view, but the notion that perception is something more than the registering and cataloging of sensations and that the perceiver is not merely a passive recipient of environmental information is at least as old as experimental psychology (e.g., Boring, 1942), and has been reflected in many perceptual theories (e.g., Allport, 1955).

One assumption of this conceptualization is that perception involves testing hypotheses, and that once a hypothesis is established it may prevent or delay the acceptance of an alternative hypothesis. This interpretation thus attributes the negative effect of prior experience to the formation of incorrect hypotheses about the identity of the stimulus that, once established, become resistant to modification. Identification is therefore impaired as a result of prior experience because the subject is seeking to confirm or disconfirm an inappropriate hypothesis; without this prior experience, the subject is not encumbered by an inappropriate hypothesis and he need only evaluate the currently available information without any predispositions.

Such an interpretation of this phenomenon is far from satisfactory since it provides no explanation of how hypotheses are generated in the first place, nor any description of the mechanism of hypothesis generation and verification. Nonetheless, it appears that only an interpretation of this type can account, even if only superficially, for the negative-effect-of-prior-experience phenomenon. The existence of this phenomenon, therefore, can be taken as evidence in support of the conceptualization of perception as a process of hypothesis testing.

Before fully accepting the phenomenon and its important theoretical implications, it is desirable to obtain assurance that the result is reliable and not an artifact of some procedural or methodological peculiarity. For this purpose, the current experiments were designed and conducted.

In all three of the earlier studies on this phenomenon, the stimulus materials were photographs or photographic slides of complex scenes (e.g., "an aerial view of a street light on the campus at night"). The procedure consisted of initially presenting the photograph or slide in a very blurred version and then gradually reducing the blur to a point at which the subject was asked to identify the

This research was supported by a grant to the senior author from the Research Council of the Graduate School, University of Missouri, Columbia, and by NIA Grant AG-00008, administered by J. Botwinick, Washington University, St. Louis, Missouri 63103.

stimulus. The present experiments utilized simple line drawings of common objects and animals as the stimulus materials, and manipulated the impoverishment of the stimuli by spatial masking of different percentages of the stimulus. If the negative-effect-of-prior-experience phenomenon is to be confirmed, one would expect the accuracy of identifying these incomplete stimuli to be inversely related to the amount of prior experience (i.e., the number of previous incomplete versions) with the stimuli.

METHOD

Subjects

There were 36, 24, and 36 college-student participants in Experiments 1, 2, and 3, respectively.

Stimuli

Forty line drawings of common objects and animals similar to those found in children's picture-word flash cards were used as stimuli. For example, the first five stimuli were a key, a pipe, a saw, a bell, and a sailboat. All drawings occupied an area of approximately 4 x 6 in., centered in a sheet of 8½ x 11 in. white paper in a loose-leaf notebook. The impoverished versions of the drawings were produced by overlaying a mask consisting of horizontal and vertical strips of paper on top of the figure and then photocopying the figure through the mask. Various combinations of horizontal and vertical strip widths from .125 to 1 in. were used to create different degrees of impoverishment. The percentage of the total paper area that was not masked by the strips ranged from 4.4% to 25.0%, with intermediate values of 6.3%, 8.8%, 10.7%, 12.5%, 15.6%, 17.8%, and 20.9%. All masks were independently constructed and applied; hence, although each allowed increasing amounts of the total area of the figure to be transmitted, the particular stimulus regions transmitted were not necessarily the same as the percentage area increased.

The incomplete versions were always presented in consecutively increasing order, with all of the versions of one figure presented before any of the versions of another figure were presented.

Procedure

Experiment 1. Subjects were presented with one, four, or seven incomplete versions of the stimulus figures and were asked to identify the figure after viewing the last incomplete version. The incomplete versions presented were Version 7 (17.8% completeness) for the subjects with only one incomplete version, Versions 4 (10.7% completeness) through 7 for the subjects with four incomplete versions, and Versions 1 (4.4% completeness) through 7 for the subjects with seven incomplete versions. Thus, all subjects were required to identify the stimulus figure after viewing the same incomplete version (i.e., Version 7), but they differed in the number of previous incomplete versions of the figure they had seen.

Experiment 2. Subjects were required to write an identification response to every incomplete version in order to monitor conditional probabilities of correct identification. Incomplete Versions 8 (20.9% completeness) and 9 (25.0% completeness) and the complete version of the figure were presented after Version 7 to provide further assessments of conditional identification accuracy.

Experiment 3. Subjects were provided with a verbal response alternative for each figure and instructed to indicate whether that response was correct (YES) or incorrect (NO) on each incomplete version. The incorrect alternatives were selected from the most frequent wrong responses to the figures in Experiment 2. Each subject received one-half of the figures with the correct response alternative and one-half with the incorrect alternative. The assignment of correct alternative to stimulus figure was balanced across subjects.

The task in all experiments was self-paced and the subjects were allowed unlimited time to view the drawings and write their responses. A response was required to each stimulus figure even if it was only a guess.

RESULTS

The data from each experiment were initially analyzed by determining the percentage of figures correctly identified on Version 7 (17.8% completeness) for each subject and submitting these percentages to an analysis of variance. The analyses were not significant for the data of Experiment 1 [$F(2,33) < 1.0$] or Experiment 2 [$F(2,46) = 1.83$, $p > .15$], but the analysis in Experiment 3 was significant [$F(2,70) = 4.27$, $p < .05$]. The mean percentages are displayed in the top three rows of Table 1; it can be seen that the differences are in the direction of the lowest amount of experience producing the lowest level of accuracy, the opposite of the result reported by earlier investigators.

A second analysis of the data from Experiments 2 and 3 involved comparisons of conditional percent correct values. These values were computed for each subject at each incomplete version number by dividing the number of stimulus figures identified on that version by the number of figures not previously identified in earlier versions.

The conditional percent correct data from Versions 7 through 9 were analyzed in an analysis of variance, with the three incomplete versions and the three levels of experience as factors in each experiment. The level-of-experience factor [Experiment 2, $F(2,184) = 15.27$, $p < .0001$; Experiment 3, $F(2,280) = 6.10$, $p < .005$], the incomplete version factor [Experiment 2, $F(2,184) = 11.40$, $p < .0001$; Experiment 3, $F(2,280) = 4.93$, $p < .01$], and the interaction of the two factors [Experiment 2, $F(4,184) = 5.43$, $p < .0005$; Experiment 3, $F(4,280) = 7.52$, $p < .0001$] were all statistically significant. Separate analyses on the data from each incomplete version revealed that only on Version 7 was the level-of-experience factor statistically significant [Experiment 2, $F(2,46) = 41.30$, $p < .0001$; Experiment 3, $F(2,70) = 29.55$, $p < .0001$]. The mean conditional percent correct values for Version 7 are displayed in the bottom two rows of Table 1.

Table 1
Cumulative and Conditional Identification Accuracy on Version 7

	Experiment	Amount of Prior Experience		
		High	Moderate	Low
Cumulative Percent Correct	1	68.3	67.1	64.0
	2	64.7	66.0	54.2
	3	82.0	83.1	75.6
Conditional Percent Correct	2	11.4	26.8	54.2
	3	29.3	32.3	75.4

DISCUSSION

None of the three experiments reported in this project provided any indication that the cumulative accuracy of identifying a degraded visual stimulus was inversely related to the amount of experience with more severely degraded versions of the stimulus. However, Experiments 2 and 3 indicated that subjects with no prior experience with incomplete versions of the stimuli were more accurate with the measure of conditional percent correct than subjects with prior experience. Since the advantage of no prior experience holds only for the first exposure of an incomplete version of the stimulus, and is not apparent in the cumulative percent correct measure of accuracy, a simple interpretation of this finding seems possible: The subjects with prior experience with the stimuli identified some of the stimuli on incomplete versions prior to the critical version, so the conditional percent correct measure for those subjects includes a higher proportion of difficult-to-identify stimuli than that for the subjects for whom the critical version was the first exposure to the stimulus. The important point is that the total number of figures identified on or before the critical version is the same for all subjects regardless of the amount of prior experience with the stimuli.

The experiments therefore lead to one or more of the following conclusions: (1) that the negative-effect-of-prior-experience phenomenon is unreliable and cannot be replicated; (2) that the phenomenon is specific to a particular type of stimuli and does not have general applicability; (3) that the phenomenon is demonstrable only with the conditional measure of accuracy which leads to artifactual differences when there are no differences in the cumulative measure of accuracy. Regardless of the specific conclusion, however, the phenomenon that greater amounts of experience with an ambiguous stimulus impair subsequent identification of the stimulus can no longer be interpreted as supporting the active hypothesis-testing view of perception. Of course, this does not mean that active theories of perception are not viable, but only that this particular phenomenon cannot be considered as evidence for that position.

REFERENCES

ALLPORT, F. H. *Theories of perception and the concept of structure.* New York: Wiley, 1955.

BORING, E. G. *Sensation and perception in the history of experimental psychology.* New York: Appleton-Century-Crofts, 1942.

BRUNER, J. S., & POTTER, M. C. Interference in visual recognition. *Science*, 1964, **144**, 424-425.

GREGORY, R. L. *The intelligent eye.* New York: McGraw-Hill, 1970.

GREGORY, R. L. Choosing a paradigm for perception. In E. C. Carterette & M. P. Friedman (Eds.), *Handbook of perception: Historical and philosophical roots of perception* (Vol. 1). New York: Academic Press, 1974.

POTTER, M. C. On perceptual recognition. In J. S. Bruner, R. R. Olver, & P. M. Greenfield (Eds.), *Studies in cognitive growth.* New York: Wiley, 1966.

WYATT, D. F., & CAMPBELL, D. T. On the liability of stereotype or hypothesis. *Journal of Abnormal and Social Psychology*, 1951, **46**, 496-500.

(Received for publication December 27, 1977.)