

On the Psychology of Experimental Surprises

Paul Slovic and Baruch Fischhoff
Decision Research, A Branch of Perceptronics
Eugene, Oregon

Studies of the psychology of hindsight have shown that reporting the outcome of a historical event increases the perceived likelihood of that outcome. Three experiments show that similar hindsight effects occur when people evaluate the predictability of scientific results—they tend to believe they “knew all along” what the experiments would find. However, the hindsight effect was reduced by forcing people to consider how the research could otherwise have turned out. Implications for the evaluation of scientific research by lay observers are discussed.

In an era of reduced public support and increased public scrutiny of scientific research, psychologists are under increasing pressure to make certain that their work is viewed as both important and informative. The lay test of importance is typically some aspect of personal or social relevance. The test for informativeness is some variant of the questions “Did I learn anything new from this research?” or “Did the results surprise me?”

Relevance judgments are, of course, a highly individual matter. Assuming that a project has been properly explained, scientists have little basis professionally for reproving an individual who believes that the problem was not worth the resources invested in it. However, when an observer claims that the results of a study were highly predictable—and thus the study need not have been conducted—there may be grounds for contention. Recent results by Fischhoff

(1975a, 1975b) and Fischhoff and Beyth (1975) have shown that (a) reporting the outcome of a historical event increases the perceived likelihood of that outcome and (b) people underestimate the effect of outcome knowledge on their perceptions. As a result, people believe that they would have seen in foresight the relative inevitability of the reported outcome, which in fact was only apparent in hindsight. Thus, they exaggerate the predictability of reported outcomes.

It seems plausible that similar effects might occur when viewing the results of scientific research. Once we hear an experiment's findings, we may tend to feel as though we “knew all along” that it would turn out that way. If this happens and people systematically exaggerate the predictability of the findings, that bias could reasonably lead people to be unduly critical in their evaluation of such “uninformative” research.

The experiments reported below examine the existence and workings of hindsight bias in lay assessments of scientific research. Subjects in Experiments 1 and 2 read descriptions of a number of studies from different scientific disciplines, each of which had two possible outcomes. *Foresight* subjects were told that a single trial was about to be tested in each study. For both possible outcomes, they were asked to indicate the probability that that outcome would be obtained on a specified number of additional

Support for this research was provided by the Advanced Research Projects Agency of the Department of Defense and was monitored by the Office of Naval Research under Contract N00014-76-C-0074 under subcontract to Oregon Research Institute from Decisions and Designs, Inc.

We would like to thank Barbara Combs, Lewis Goldberg, Sarah Lichtenstein, Michael Posner, and several reviewers for comments on earlier drafts.

Requests for reprints should be sent to Baruch Fischhoff or Paul Slovic, Decision Research, A Branch of Perceptronics, 1201 Oak Street, Eugene, Oregon 97401.

replications *if* it were obtained on the first trial. *Hindsight* subjects were told that the first trial had already been conducted and had produced one of the possible outcomes. They were asked how likely it was that this outcome would be obtained on the same specified number of additional replications.

Thus, both groups were asked to assess the probability of a number of future replications, conditional on the outcome obtained on a first trial. Formally, these conditional probabilities should be the same for subjects in both groups. We hypothesized, however, that hindsight subjects, told the outcome obtained on the first trial, would exaggerate its inevitability and thus the probability that it would be replicated on future trials. Foresight subjects, we believed, would be less sanguine about the prospects of successful replications. One reason for such an effect is that foresight is a perspective conducive to seeing how a study could go either way; whereas in hindsight, we may be so intent on explaining the reported result that we can no longer see how the study could, in past or future, turn out otherwise.

These experiments differ from previous hindsight work in one important respect. In order to demonstrate inconsistency (bias) between hindsight and foresight judgments, Fischhoff (1975b) forced hindsight subjects to ask themselves "What would I have thought had I not been told how things turned out?" A variety of recent studies (reviewed in Fischhoff, 1977) have shown that hypothetical judgments of this sort are very difficult. With the present design, it is the foresight subjects who are required to perform hypothetical judgments, asking themselves "What will I think if things turn out a particular way?" Thus, the present experiments will show whether hindsight effects are obtained with this rather different design.

Experiment 1

Method

Design. Subjects received brief descriptions of studies drawn from biology, psychology, and meteorology, which they were told either would soon be conducted (foresight) or had recently been

conducted (hindsight). For each study, foresight subjects were told that two outcomes were possible with the first trial; hindsight subjects were told that one of those two possible outcomes had been obtained. Foresight subjects were asked to (a) assign a probability to each of the possible first-trial outcomes, (b) explain why each outcome might occur, and (c) estimate the probability that each of the two possible outcomes would be replicated in all, some, or none of a fixed number of replications if it were obtained on the initial trial. Hindsight subjects were asked to (a) explain why the reported outcome had occurred and (b) estimate the probability that it would be obtained in all, some, or none of the replications. The dependent variable for all groups was the conditional probability of replicating the outcome of the initial trial.

Stimuli. The four studies, along with the possible outcomes considered and the number of replications, were as follows:¹

1. Virgin rat. Several researchers intend to perform the following experiment: They will inject blood from a mother rat into a virgin rat immediately after the mother rat has given birth. After the injection, the virgin rat will be placed in a cage with the newly born baby rats, after removal of the actual mother. The possible outcomes were (a) the virgin rat exhibited maternal behavior or (b) the virgin rat failed to exhibit maternal behavior. Subjects estimated the probability of the initial result being replicated with all, some, or none of 10 additional virgin rats.
2. Hurricane seeding. A team of government meteorologists recently seeded a tropical storm, which had reached hurricane status, with large quantities of silver-iodide crystals (the same type of crystals that are used to seed clouds in attempts to produce rain). The possible outcomes were (a) the hurricane increased in intensity or (b) the hurricane decreased in intensity. Subjects estimated the probability of the initial result being replicated in all, some, or none of six additional hurricanes.
3. Gosling imprinting. A goose egg was placed in a soundproof, heated box from time of laying to time of cracking. Approximately 2 days before it cracked, the experimenter began intermittently to play sounds of ducks quacking into the box. On the day after birth, the gosling was placed on a smooth floor equidistant from a duck and a goose, each of which was in a wire cage.

¹ For stylistic purposes, the tenses of the verbs used in these descriptions varied between experiments and between hindsight and foresight versions of the same experiment. Fischhoff (1976) has found that the tense used in describing events has no effect on their perceived likelihood.

The gosling was observed for 2 minutes. The possible outcomes were (a) the gosling approached the caged duck or (b) the gosling approached the caged goose. Subjects estimated the probability of the initial result being replicated with all, some, or none of 10 additional goslings.

4. The Y test. In the pretest of an experiment that she intends to run in the future, an experimenter placed a 4-year-old child in front of an easel with a large Y on it, with a dot in the lower left-hand third of the letter. The child was then taken around to the back of easel where he saw another Y. He was asked to draw a dot in the "same position" on that Y as the one he had just seen. The possible outcomes were (a) the child placed a dot in Area A (the lower left-hand third) or (b) the child placed a dot in Area B (the upper third). Subjects estimated the probability that the initial result would be replicated with one additional child. The lower right-hand was labeled Area C.

The hurricane seeding study was loosely based on Howard, Matheson, and North (1972); the imprinting study on Grier, Counter, and Shearer (1967); and the Y test on Smothergill, Hughes, Timmons, and Hutko (1975). The virgin rat study was invented.

The virgin rat study was presented to one set of foresight and hindsight groups. The other three studies were presented together to a second set of foresight and hindsight groups. These hindsight subjects received either Outcome (A) of each of Studies 2, 3, and 4 or Outcome (B) for each.

Instructions. All subjects received the same general instructions: "The following questionnaire concerns your scientific intuitions. We would like to ask you a number of questions about possible results of several experiments in different areas that have recently been conducted or will be in the near future. We thank you for your cooperation." Each study appeared on a separate page, with the description at the top. Questions were presented in the following format (using the virgin rat example):

Foresight

1a. What is the probability that the virgin rat will exhibit maternal behavior? _____ Why do you think that this might happen?

1b. What is the probability that the virgin rat will not exhibit maternal behavior? _____ Why do you think that this might happen?

2. If the virgin rat does exhibit maternal behavior, what is the probability that in a replication of this experiment with 10 additional virgin female rats,

a. all will exhibit maternal behavior? _____

b. some will exhibit maternal behavior? _____

c. none will exhibit maternal behavior? _____

(Note: These three probabilities should total 100%.)

3. Question 3 was identical to Question 2, except that it began "If the virgin rat does not exhibit maternal behavior. . . ."

Hindsight (after being told either that the initial virgin rat exhibited maternal behavior or that it failed to exhibit maternal behavior)

1. Why do you think that this happened?

2. What is the probability that in a replication of this experiment with 10 additional virgin female rats

a. all will exhibit maternal behavior? _____

b. some will exhibit maternal behavior? _____

c. none will exhibit maternal behavior? _____

(Note: These three probabilities should total 100%.)

Subjects. All 184 subjects were paid volunteers who responded to an ad in the University of Oregon student newspaper. The present task was the first of several performed during a 2-hour session. Group size varied from 24 to 37 subjects.

Results

The second and fourth columns of Table 1 present the mean probability of replication assigned by the foresight and hindsight groups in Experiment 1. The italicized rows of Table 1 present the mean judged probability of the initial outcome being obtained on all subsequent replications. In six of eight cases (two from each of four studies), this probability was significantly larger for hindsight than foresight subjects. Thus, subjects who were told that a study had "worked" once in the past found its working consistently in the future more likely than those asked "If it works once, how likely is it to work again consistently?" For the three studies with multiple replications (virgin rat, hurricane seeding, and gosling imprinting), the mean probability of an initial outcome always being replicated was .38 for the foresight group and .55 for the hindsight group; the mean probability of its never being replicated was .19 and .10, respectively.

Discussion

Why should these formally equivalent conditional probabilities be judged differently by hindsight and foresight subjects?

Table 1
Mean Probabilities for Experiments 1 and 2

Initial result and kind of replication	Foresight		Hindsight	
	Two alternatives (Experiment 1) ^b	One alternative (Experiment 2) ^c	One alternative (Experiment 1) ^d	Two alternatives (Experiment 2) ^e
Virgin rat study				
Shows maternal behavior				
a. <i>All show maternal behavior</i> ^a	.30*	.27	.44	.47*
b. Some show maternal behavior	.42	.50	.49	.40
c. None show maternal behavior	.29***	.23	.07	.13***
Fails to show maternal behavior				
a. All show maternal behavior	.10	.09	.09	.11
b. Some show maternal behavior	.33	.32	.24	.31
c. <i>None show maternal behavior</i>	.57	.58	.67	.58
Hurricane seeding study				
Intensity increases				
a. <i>All increase</i>	.29***	.31	.56***	.35
b. Some increase	.49	.52	.37*	.51
c. None increase	.22***	.17	.08**	.14
Intensity decreases				
a. <i>All weaken</i>	.34*	.32	.47	.50**
b. Some weaken	.47	.49	.40	.38
c. None weaken	.19	.19	.13	.12
Gosling imprinting study				
Approaches goose				
a. <i>All approach goose</i>	.40***	.33	.73*	.59*
b. Some approach goose	.41**	.51	.22	.34
c. None approach goose	.19**	.15	.05	.07**
Approaches duck				
a. <i>All approach duck</i>	.40	.40	.42	.33
b. Some approach duck	.46	.44	.43	.55
c. None approach duck	.16	.16	.15	.12
Y-test study				
Places dot in Area A				
a. <i>Places in Area A</i>	.29***	.30	.58***	.38*
b. Places in Area B	.13	.12	.12	.18
c. Places in Area C	.57***	.58	.30**	.44**
Places dot in Area B				
a. Places in Area A	.26	.29	.24*	.29
b. <i>Places in Area B</i>	.16***	.25*	.36	.31***
c. Places in Area C	.58***	.46*	.39	.39***

Note. Sample size varies from 24 to 41 subjects.

^a Italicized initial result indicates outcome reported to have happened (hindsight) or considered as happening (foresight).

^b Significance tests in this column refer to differences from one-alternative hindsight.

^c Significance tests in this column refer to differences from two-alternative foresight.

^d Significance tests in this column refer to differences from two-alternative hindsight.

* $p < .05$.

** $p < .01$.

*** $p < .001$.

Two possibilities occur to us. One, suggested in the introduction, is that hindsight subjects unduly concentrate their attention on the reported outcome, thereby failing to see how the initial trial could have gone the other way. A second possibility is that the conditional judgments that the foresight subjects make ("If it were to work once, what is . . . ?") are quite difficult and confusing. Thus, when they attempt to consider the possible occurrence of two different out-

comes, they may be unable to devote to either the attention given by hindsight subjects to their one alternative. As a result, foresight subjects may be unable to assess properly the impact that the result on the first trial should have on their perceptions. In summary, the *availability-of-reasons* explanation attributes the discrepancy to hindsight subjects' failure to consider the feasibility of alternative outcomes. The *conditionality* explanation attributes the effect

to the inability of foresight subjects to consider multiple contingencies. Both may be true.

Experiment 2 tests these hypotheses by replicating Experiment 1 with the following differences: (a) Foresight subjects were required to consider the probability of replicating only one of the possible outcomes; (b) hindsight subjects were required to explain not only why the reported outcome happened, but also "Had the experiment worked out the other way, how would you explain it?" These *one-alternative* foresight subjects should be able to devote the same undivided attention to their one possible outcome that the one-alternative hindsight subjects in Experiment 1 could devote to their one reported outcome. If the conditionality hypothesis is correct, they should respond more like one-alternative hindsight subjects than like the two-alternative foresight subjects in Experiment 1. According to the availability-of-reasons hypothesis, *two-alternative* hindsight subjects forced to consider why the unreported outcome might have occurred should respond like foresight subjects.

Experiment 2

Method

Experiment 2 was identical to Experiment 1 except for two changes. The first was that foresight subjects estimated the probability of replicating only one of the two possible outcomes for each study. They were asked, in effect, either "If the study works, how likely is that result to be replicated?" or "If the study doesn't work, how likely is that result to be replicated?" The (two-alternative) foresight group in Experiment 1 answered both these questions. Second, after the two-alternative hindsight group of Experiment 2 was asked "Why did the study work out this way?" they were also asked "Had the study worked out the other way, how would you explain it?" Like the one-alternative hindsight subjects of Experiment 1, they estimated the probability of replication only for the reported outcome. There were 151 subjects, recruited in the same manner as Experiment 1, who participated in Experiment 2. All subjects considered either Outcome A or Outcome B for each of the four studies.

Results

Columns 3 and 5 of Table 1 present the mean probabilities from Experiment 2.

Comparing columns 2 and 3, we see that the responses of one- and two-alternative foresight subjects were generally indistinguishable. Reducing the number of alternatives considered did not systematically increase the perceived probability of replicating the initial outcome. In 18 of 24 cases, the one-alternative foresight mean was closer to the two-alternative foresight mean than to the one-alternative hindsight mean. Thus, there is no evidence that attentional problems were responsible for the hindsight-foresight discrepancy.

The second manipulation, forcing two-alternative hindsight subjects to consider how the first trial of each study could have turned out otherwise, produced a marked difference. Comparing column 5 with columns 2 or 3 of Table 1, there was still a substantial hindsight effect for five of the eight outcomes considered (see Footnote c to Table 1). The size of the effect, however, was reduced. For four of the eight outcomes (see Footnote d to Table 1), the mean probability of consistently replicating the reported outcome was significantly lower for two-alternative than for one-alternative hindsight subjects. In general, the means of the two-alternative hindsight subjects lie between those for the one-alternative hindsight subjects and both foresight groups. These results strongly support the availability-of-reasons hypothesis.

Further evidence of the effect of reason availability on probability judgments was sought by looking at hindsight subjects' confidence in replication as a function of their ability to supply reasons for the different outcomes. Subjects who only provided a reason for the *reported* outcome assigned a mean probability of replication of .56; those who provided reasons for both alternatives assigned a mean probability of only .40; whereas those who only provided a reason for the *unreported* outcome found replication even less likely (mean probability = .24).

Experiment 3

One possible fault with Experiments 1 and 2 is that the descriptions of the various

studies were only one paragraph in length. Although such brevity is typical of the media reports of scientific research from which the public presumably receives most of its information about science, one could have more confidence in the present hindsight effect if it were obtained from people who had considered the research in question more thoroughly. Experiment 3 provided an opportunity for such consideration by presenting subjects with simulated (although necessarily incomplete) manuscripts in a journal review format.

Method

Design. Subjects were asked to read and evaluate scientific manuscripts in a manner similar to that of professional reviewers. Hindsight reviewers received manuscripts with introduction, method, and results sections. For foresight subjects, the results section was missing. Each manuscript was composed so that there were two possible outcomes for the study in question. There were two separate hindsight groups, each receiving one of the possible outcomes presented as if it had actually happened.

Subjects were asked to evaluate the manuscripts on seven 7-point scales, two of which were designed to be sensitive to hindsight-foresight differences. One was *surprisingness of results*, in which hindsight subjects assessed the surprisingness of the reported outcome, and foresight subjects assessed how surprising each of the two possible outcomes would seem were they obtained. The second sensitive question was *stability of results*, in which hindsight subjects assessed the likelihood that the reported results would be obtained in an exact replication of the same study, and foresight subjects answered the same question for each of the possible results. The remaining five scales were used as fillers and as tests for other possible changes between hindsight and foresight. They referred to clarity of the introduction, clarity of the research question, clarity of the method, adequacy of the method to test the research questions, and personal interest in the study.

Stimuli. Three experiments from diverse areas of psychology were used.² One called "Scientific Ambiguity and Attitudinal Conflict," described an unpublished study that we had recently completed. In that study, subjects first indicated their position on several environmental issues, including nuclear power; later they were asked to guess whether an ambiguous statement about nuclear power was offered by an opponent or proponent of nuclear power. The introduction advanced the hypothesis that people would interpret ambiguous statements as supporting their own positions. The two pos-

sible outcomes were confirmation and disconfirmation of the hypothesis.

The second and third studies were elaborated versions of the gosling imprinting and Y-test studies used in Experiments 1 and 2. No hypothesis was advanced for either of these studies. The two outcomes used for the imprinting experiment were (a) follow the duck or (b) follow the goose. The two outcomes for the Y test were (a) place the dot in Area A (the lower left-hand corner of the letter Y) or (b) place the dot in Area C (the lower right-hand corner).

The studies were chosen because they were unfamiliar to our subjects, yet comprehensible without prior knowledge of the area. The studies were written to show that there were two possible outcomes, each of which could conceivably be obtained.

Procedure. Subjects were told about the review process for scientific manuscripts and then were asked to perform a task similar to that of actual reviewers. They read the three studies in the order given above, evaluating each before going on to the next. Foresight subjects considered both possible outcomes for each study; hindsight subjects considered either the first outcome or the second outcome for each.

Subjects. There were 128 paid subjects who participated, having responded to an advertisement in the University of Oregon student newspaper. They were assigned to the foresight group or one of the two hindsight groups according to their preference for experimental date and hour.

Results

If these reviewers are susceptible to a hindsight bias, then hindsight subjects should find the reported results less surprising and more likely to be replicated than the foresight subjects anticipated they would appear. Table 2 presents the relevant group means for the two outcomes used for each of the three studies. In five of the six cases, hindsight subjects found the reported outcome less surprising than did foresight subjects; in three of these cases, this difference was statistically significant. In five of six cases (not the same five), hindsight subjects found the reported outcome more replicable; again three differences were statistically significant. There were no systematic differences on the five filler questions.

² Copies of these descriptions and the accompanying questionnaires are available upon request.

Table 2
Mean Ratings for Experiment 3

Outcome	Surprisingness of results (1 = very surprising)		Stability of results (1 = very likely to replicate)	
	Foresight	Hindsight	Foresight	Hindsight
Ambiguity experiment				
Confirm hypothesis	5.57	5.42	3.09	2.76
Disconfirm hypothesis	3.02	4.67***	4.41	2.91***
Gosling imprinting experiment				
Follow goose	4.05	4.28	2.77	2.81
Follow duck	3.66	4.62*	3.11	2.93
Y-test experiment				
Place dot in Area A	5.34	6.04*	2.75	1.73**
Place dot in Area C	3.23	3.69	3.52	2.80*

- * $p < .05$ (one-tailed).
 ** $p < .01$ (one-tailed).
 *** $p < .001$ (one-tailed).

General Discussion

Reported outcomes seem less surprising in hindsight than in foresight. Summarizing the results of the present experiments and previous work, this is apparently true whether the outcomes are historical developments or the results of scientific experiments, whether hindsight subjects perform hypothetical or straightforward judgments, whether foresight subjects engage in cognitively complex or cognitively simple tasks, and whether the antecedent events are presented in greater or lesser detail.

Although we have termed these results "hindsight effects," one might argue that they reflect foresight subjects' inability to see how things will look in the future as much as hindsight subjects' inability to see how things looked in the past. These two interpretations roughly parallel the conditionality and availability-of-reasons hypotheses offered following Experiment 1. The fact that simplifying the foresight task in Experiment 2 failed to influence judgments—whereas helping hindsight subjects recruit reasons for the unreported outcome did reduce confidence in replication—suggests that hindsight is the problematic perspective. These results are not, however, conclusive.

It is possible that conditional tasks, however structured, cause difficulties.

Before a firmer conclusion may be drawn about the conditionality hypothesis, further research is needed on how people make such hypothetical judgments about the way that possible futures will appear should they come about (the sort of task posed to our foresight subjects). Aside from their theoretical interest, such judgments are an important component of contingency planning (as in "What will it be like?" "How shall I respond?" "What should I do to be prepared when I take my first job; when my first child arrives home; when Portugal withdraws from NATO?"). If such hypothetical judgments are not performed well, then contingency plans based upon them may appear grossly inappropriate when foreseen contingencies do arise but do not appear "in the flesh" the way they were supposed to look. Brown (1975), in one of the only studies addressed to this question, suggests that this is often the case even with formal contingency planning.

Although these results have shown the robustness of hindsight bias—particularly the fact that it can be demonstrated without forcing subjects to make hypothetical judg-

ments—the results have not completely clarified its operation. Two basic types of explanation were advanced by Fischhoff (1975b). One type attributes the effect to the way people approach hindsight tasks. The other attributes it to the reorganization of one's perceptions that follows being told what has happened, which makes it impossible to retrieve information available in foresight however one approaches a problem.

In Fischhoff (1977), subjects performing a task prone to hindsightlike bias were directly told about the bias and its extent and encouraged to avoid it. That manipulation failed, and its failure was taken as support for the idea that being told what has happened irrevocably alters one's cognitive representation of the event in question. The partial success of the debiasing manipulation in the present Experiment 2 suggests that there are ways to help people retrieve some of the perspective available in foresight. Whether there are even more effective manipulations or whether some of that foresight perspective is forever lost is a topic for future research.

Implications

What can we, as scientists, do if we feel that the public judges the informativeness of our work unfairly? One response is to stress in our writing and speaking the unpredictability of our results. A classic example of this approach is Paul Lazarsfeld's (1949) article "The American Soldier—An Expository Review," in which he first presented the reader with several obvious findings from the study and then revealed that the opposite results were, in fact, true. Thus Lazarsfeld forced his readers to consider the plausibility of alternative outcomes in a man-

ner considerably more extreme than that adopted in Experiment 2.

Writing to highlight unpredictability might not only be useful to the image of our profession, but it could also be therapeutic to our readers. People who exaggerate the predictability of past events in general and research results in particular overstate the extent of their own knowledge. Showing them when some surprise is in order should help them appraise their need for learning in general and scientific research in particular.

References

- Brown, R. V. *Modeling subsequent acts for decision analysis*. McLean, Va.: Decisions and Designs, 1975.
- Fischhoff, B. Hindsight: Thinking backward? *Psychology Today*, 1975, 3, 4. (a)
- Fischhoff, B. Hindsight \neq foresight: The effect of outcome knowledge on judgment under uncertainty. *Journal of Experimental Psychology: Human Perception and Performance*, 1975, 1, 288-299. (b)
- Fischhoff, B. Temporal setting and judgment under uncertainty. *Organizational Behavior and Human Performance*, 1976, 15, 180-194.
- Fischhoff, B. Perceived informativeness of facts. *Journal of Experimental Psychology: Human Perception and Performance*, 1977, 3, 349-358.
- Fischhoff, B., & Beyth, R. "I knew it would happen"—remembered probabilities of once-future things. *Organizational Behavior and Human Performance*, 1975, 13, 1-16.
- Grier, J. B., Counter, S. A., & Shearer, W. M. Pre-natal auditory imprinting in chickens. *Science*, 1967, 155, 1692-1693.
- Howard, R. A., Matheson, J. E., & North, D. W. The decision to seed hurricanes. *Science*, 1972, 176, 1191-1202.
- Lazarsfeld, P. The American soldier—An expository review. *Public Opinion Quarterly*, 1949, 13, 377-404.
- Smothergill, D. W., Hughes, F. P., Timmons, S. A., & Hutko, P. Spatial visualizing in children. *Developmental Psychology*, 1975, 11, 4-13.

Received November 12, 1976 ■