

Studies of inference from lack of knowledge

DEDRE GENTNER and ALLAN COLLINS

Bolt Beranek and Newman Inc., Cambridge, Massachusetts 02138

Two experiments were performed to investigate the conditions affecting lack-of-knowledge inferences. The lack-of-knowledge inference is a metainference, an inference based on knowledge about one's own knowledge, in which the absence of information concerning a possible assertion is taken as evidence that the assertion is false. In Experiment 1, it was shown that lack of knowledge about an assertion decreases subjects' ratings of the likelihood that the assertion is true. The more important the assertion and the more expert the person who lacks knowledge, the more certain is the lack-of-knowledge inference, as measured by a decrease in the rated likelihood of the assertion. In Experiment 2, a hypothesis about the nature of self-judgments of expertise was tested. This intrinsic observation hypothesis states that expertise relevant to an assertion is determined by the number of similar assertions from the same domain that the person can retrieve. The hypothesis was ruled out by the results of Experiment 2. Other possibilities for the manner in which self-judgments of this kind are made are discussed.

People often must reason from incomplete knowledge (Carbonell & Collins, 1973; Collins, 1978; Collins, Warnock, Aiello, & Miller, 1975). They frequently have to make decisions without certain knowledge of all the relevant facts. One of the ways people might manage in spite of this is by taking advantage of metaknowledge about the mind: by using their patterns of knowledge as sources of information. Collins et al. (1975) have proposed as one example of this kind of reasoning the lack-of-knowledge inference. In this inference, a person who is trying to verify an assertion reasons that he or she would certainly remember the fact if it were true. Therefore, failure to find in memory any information one way or the other allows the conclusion that the assertion is false. In this way, the very incompleteness of a given data base could provide useful information to the reasoner.

The evidence concerning people's knowledge about the structure of their minds is mixed. Although people seem to be able to estimate fairly well what kind of efforts are required to commit something to memory (Brown, 1975; Flavell, Friedrichs, & Hoyt, 1970; Hagen, 1972), they seem to be systematically overconfident about the accuracy of their general knowledge (Fischhoff, Slovic, & Lichtenstein, 1977). Further, Nisbett and Wilson's (1977) examples of subjects' failures to introspectively report all or part of a given experience again cast doubt on the notion that people are able to derive useful information about their own mental structures. In

this paper, we investigate the lack-of-knowledge inference as a possible instance in which metaknowledge is used in reasoning.

A typical lack-of-knowledge inference is shown in the following protocol. We asked a subject if he had ever shaken hands with Richard Nixon. He was quite certain that he had not. When we asked why, he responded, "How do I know? [Laughter] It's not something that one would forget. I don't think I've ever seen him live, in person. I'm sure I haven't." When pressed as to why he was so sure, he said, "Oh, you want me to get that specific, huh? I don't think I've ever seen—I saw—I don't know—This isn't really an answer to the question.—All I can say is by elimination I know I have seen some presidents in the flesh." (He went on to describe seeing John Kennedy and Lyndon Johnson.)

From this and other protocols, two major factors emerge as affecting the certainty of the lack-of-knowledge inference. The first is the salience or importance of the issue. When asked about an assertion within some domain, people will take their lack of knowledge about that assertion as counterevidence if the fact is important enough that it should be memorable. For example, a person who cannot remember meeting Richard Nixon is fairly sure that he has not, but the fact that he cannot recall meeting an undersecretary in the diplomatic corps might not provide much evidence as to whether the meeting had actually taken place.

Similar patterns have been found in memory experiments. Brown, Lewis, and Monk (1977) found that, in recognition tests for lists of words, subjects rejected distractors judged to be memorable more confidently than distractors judged to be nonmemorable. They apparently reasoned that they would have recognized a memorable item had it appeared. Brewer and Treyns (1981) found similar results in a naturalistic setting. Subjects were asked to recall and/or recognize objects from a room and, subsequently, to rate the objects for

This research was supported by the Advanced Research Projects Agency of the Department of Defense under Contract MDA 903-77-C-0025. We would like to thank Lance Rips for his very helpful critique and Brenda J. Starr for her important contributions at every stage of this study. The views and conclusions contained in this document are those of the authors and should not be interpreted as necessarily representing the official policies, either expressed or implied, of the Advanced Research Projects Agency or the U.S. government.

saliency and for probability of being found in that context. Subjects were better able to recall and recognize high-salient objects than low-salient objects. Further, among objects not present, high-salient objects were most confidently rejected in the recognition test. Subjects apparently assumed in recognition that they had good positive memory for the high-salience objects that had been presented. This impression was fairly accurate, as evidenced by the correlation of both recall and recognition with saliency.

The second factor affecting certainty in the lack-of-knowledge inference is how much the person knows about the general topic. In the example of meeting public officials, a person who is an expert in foreign affairs would expect to remember meeting not only Richard Nixon but even an unimportant diplomat. In that case failure to remember meeting an undersecretary in the diplomatic corps might indeed allow the conclusion that the meeting had not taken place. Thus both the saliency of the topic and the expertise of the judge can contribute to subjective negative likelihood.

Normally, a person has several lines of incomplete knowledge that can be brought to bear on a given issue. In addition to any negative likelihood stemming from a lack-of-knowledge inference, the person may have counterinformation that suggests that the assertion is true. For example, if our subject had lived in San Clemente during the time that Nixon lived there, the probability would be greater that he might have met Nixon. Given such a source of positive likelihood, would the negative inference still have force? Our protocols suggest that both inferences would be made: the negative inference from the lack of memory for the event and the positive inference from the knowledge of the opportunity for the event. When a person's memory contains both positive and negative evidence with respect to a given assertion, these counterfactors must somehow be weighed in the decision process.

The experiments reported here are concerned both with the factors affecting the strength of a lack-of-knowledge inference and with the way in which negative evidence is integrated with positive evidence about an assertion. Although our chief interest is in people's judgments of their own knowledge of the world, the difficulty of manipulating this knowledge led us to adopt an indirect approach. We instead gave subjects scenarios that we hoped would lead them to inspect the knowledge attributed to the protagonists of the episodes. The basic paradigm was to lay out a scenario for subjects and then ask them to rate the likelihood of the protagonist's having had certain experiences in that context. Half the time they simply rated the likelihood of the experience; the other half of the time, they were told that the person could not remember having had the experience. To the extent that the lack-of-knowledge inference operates in this vicarious introspection, subjects' ratings should be depressed by the information that the person has no memory of the experience. The basic

data, then, were the differences in likelihood between the simple ratings and the ratings in the lack-of-knowledge condition.

The experiences themselves were varied to be either of high saliency or of low saliency. Here, we assumed that, as in the Brewer and Treyens (1981) and Brown et al. (1977) studies, our subjects would feel that the more salient or memorable a given fact or event, the less likely it should be that someone could have stored it at some time and then have forgotten it. Therefore, we expected high-salience items to be most affected by the lack-of-knowledge information. That is, we expected that the difference in likelihood ratings between the simple scenarios and the lack-of-knowledge scenarios would be greater for high-salience items than for low-salience items.

The other factor varied in the lack-of-knowledge condition was whether the protagonist was portrayed as an expert or as a novice in the general topic area. Here, we predicted that the fact that an expert cannot remember some possible fact would count as strong evidence against the truth of that fact. On the other hand, that a novice cannot recall some possible fact would count very little about the truth or falsity of that fact. Therefore, our prediction was that lack-of-knowledge information would depress likelihood ratings more when the protagonist was portrayed as an expert than when the protagonist was portrayed as a novice.

EXPERIMENT 1

Method

Subjects read a series of short scenarios. For each scenario, the subjects rated the likelihood of the event or state of affairs described. The likelihood scale ranged from -5 (absolutely certain no) to +5 (absolutely certain yes).

Subjects. The subjects were 14 undergraduates from Lesley College, Cambridge, Massachusetts, who had not served in other inference experiments. They were paid for their participation.

Materials. Ten basic scenarios, each with six versions, were prepared. The 10 scenarios are summarized in Table 1.¹ Three independent variables entered into the construction of the versions. The first of these, called judgment type, refers to whether the scenario was a simple likelihood judgment based on a positive induction (Rips, 1975; Collins, Note 1) or was a judgment of likelihood given lack of knowledge on someone's part. Scenario 1 exemplifies the simple likelihood judgment and Scenario 2 shows the lack-of-knowledge judgment: (1) "If someone has seen an oriole and a starling, what is the likelihood that the person has seen a robin at some time?" (2) "Your niece is very fond of animals and knows a lot about birds. She only remembers seeing an oriole and a starling in real life. She does not remember ever seeing a robin. What is the likelihood that she has seen a robin at some time and just forgets seeing it?"

The two other factors were saliency and expertise. Saliency refers to the importance for memorability of the target item. For example, in the bird scenario, the target item was varied between robin (a highly salient bird) and wren (a less salient bird). Within the lack-of-knowledge condition, the factor of expertise refers to whether the protagonist of the scenario is described as an expert in the subject matter or as a novice. In Scenario 2, above, the niece was described as an expert. The novice version of the scenario is: (3) "Your five-year-old niece knows very little about

Table 1
Summaries of Scenarios Used in Experiment 1

Topic	Items Experienced	Target	
		High Salience	Low Salience
zoo	weasel zebra	elephant	lynx
farm chores	collect eggs feed pigs	milk cows	brush horses
burglary	camera transistor radio	TV	Timex watch
birds	oriole starling	robin	wren
minerals	copper zinc	silver	titanium
American offices	President Head CIA	Vice President	Governor—Maryland
advertising firm	new lawyer sell more stock	contract	stamp budget
clothing items	scarf socks	sweater	bathrobe
newspaper	first moon landing J. Edgar Hoover's death	Nixon in China	Swine flu epidemic
Cinderella	stepsisters prince's messenger	fairy godmother	coachman

birds. She only remembers seeing an oriole and a starling in real life. She does not remember ever seeing a robin. What is the likelihood that she has seen a robin at some time and just forgets seeing it?"

Design. The design was a 2 by 2 within-subjects factorial design of judgment type by salience. One level of the judgment-type factor, the lack-of-knowledge level, was further subdivided into two levels of expertise.

The central prediction was that there would be a main effect of judgment type: ratings of simple likelihood should exceed ratings of likelihood given lack of knowledge. The prediction for salience was that increasing the salience of the target would depress likelihood ratings in the lack-of-knowledge condition, but not in the simple likelihood condition. We expected subjects to think that the niece could have seen a wren and forgotten about it, but not a robin. Thus we expected an interaction between judgment type and salience. The expertise prediction was that if people believe that an expert is less likely than a novice to have forgotten something in the area of his or her expertise, then the negative effect of lack of knowledge should be greater if the protagonist is an expert. Thus, we predicted a main effect of expertise within a lack-of-knowledge judgment.

Procedure. Each subject rated four versions of each scenario: a high-salience and a low-salience version of the scenario in both the simple likelihood condition and the lack-of-knowledge condition. For each subject, half of the lack-of-knowledge versions involved novice protagonists and half involved expert protagonists. However, on any one scenario, a subject saw only one level of expertise, so that it was never necessary for the subject to revise the character of the protagonist. For example, in the bird scenario, a subject might rate high-salience and low-salience targets in the simple likelihood version and in the novice lack-of-knowledge version. In another scenario, the subject would rate the simple likelihood version and the expert lack-of-knowledge version. Subjects were divided into two groups in order to counterbalance the stories that occurred in the expert and the novice versions. In addition to the 40 experimental versions of scenarios (4 versions \times 10 scenarios), subjects rated 10 filler scenarios of varied construction. These were included to make it difficult for subjects to discover the patterns in the stimuli. These 50 items were presented in random order, except that the first scenario was always a filler item and no two ver-

sions of the same scenario were presented sequentially. The same random order was used for all subjects.

Results

The results, shown in Figure 1, agree quite well with the predictions of the theory. Subjects considered it far less likely that a fact was true when they were given lack-of-knowledge information than when they rated simple likelihood. The mean rating of simple likelihood was 1.26, and the mean rating given lack-of-knowledge information was $-.78$. This shift from positive to negative likelihood is evidence of our subjects' belief that lack of knowledge constitutes strong evidence against a fact.

This effect, as predicted, was stronger for expert characters than for novice characters. Within the lack-

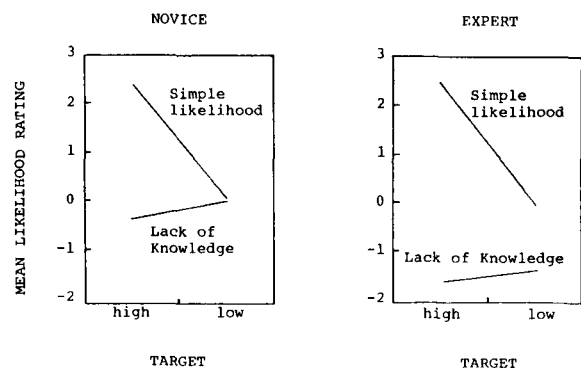


Figure 1. Results of Experiment 1: Mean ratings of the likelihood of target assertions of high or low importance, both for the simple likelihood condition and for novice or expert scenarios in the lack-of-knowledge condition.

of-knowledge condition, the mean likelihood rating for a fact, given that a novice could not recall the fact, was about neutral ($-.10$); the rating when an expert could not recall the fact was strongly negative (-1.46). Thus an expert's failure to recall a fact is stronger evidence against the fact than a novice's failure to recall it.

Finally, there was the predicted interaction between salience and judgment type. In the simple likelihood condition, high-salience facts were considered more likely to have been experienced than low-salience facts. The mean ratings were 2.39 for high-salience targets and .12 for low-salience targets. This comparison reversed in the lack-of-knowledge likelihood ratings. The mean ratings given lack of knowledge were $-.97$ for high-salience targets and $-.59$ for low-salience targets. Thus our subjects considered lack of knowledge to be very strong evidence against a proposed fact, particularly against an important and memorable fact. Within the lack-of-knowledge condition, expertise did not interact with salience. The overall ratings of both high-salience and low-salience facts are much lower when an expert experiences lack of knowledge than when a novice experiences it, but the difference between high- and low-salience targets is about the same for experts and for novices.

Analyses: Significance tests of variance. The design was somewhat complex, owing to the embedding of expertise within the lack-of-knowledge condition. Therefore, several analyses of variance were performed. To test the significance of judgment type and of salience, the novice and expert versions of the lack-of-knowledge scenarios were separately compared with the simple likelihood versions. In both cases, the analysis was a within-subjects design of judgment type (simple likelihood vs. lack-of-knowledge likelihood) crossed with salience (high or low). As predicted, judgment type was significant as a main effect in both the novice and expert analyses [$F(1,13) = 20.85$, $p < .001$, and $F(1,13) = 29.23$, $p < .001$, respectively].

The second key prediction, the interaction of judgment type and salience, was also significant in both analyses [$F(1,13) = 15.05$, $p < .005$, and $F(1,13) = 46.67$, $p < .001$, respectively]. The factor of salience was significant as a main effect in the expert analysis [$F(1,13) = 8.51$, $p < .05$] but not quite significant in the novice analysis [$F(1,13) = 4.49$, $p < .10$]. There were no other significant effects.

Item analyses were also performed, although they are questionable for this study, since the stimuli were not randomly selected but were, rather, carefully constructed. The results of the item analyses were quite consistent with those of the subject analyses. Judgment type was significant in all four analyses, as was the interaction of Salience by Judgment Type; the factor of salience itself reached significance only when expert protagonists were used in the lack-of-knowledge versions.

A second kind of analysis of variance was used to test the significance of the salience and expertise factors

within the lack-of-knowledge condition. This design was also a 2 by 2 within-subjects design. The effects of expertise were significant [$F(1,13) = 5.64$, $p < .05$], but not the effects of salience or the interaction between expertise and salience. The analysis over items yielded the same pattern of significance.

Discussion

The predictions of the lack-of-knowledge hypothesis were borne out quite well in Experiment 1. The subjects considered the fact that a person could not recall an event or fact to be evidence that it was not true, particularly if the person was portrayed as an expert in the topic matter. Moreover, the more salient the fact, the more strongly the subjects believed that the protagonist's not knowing it constituted negative evidence. They appeared to have a metarule about the human mind to the effect that people can retrieve important things if those things are true. This effect is stronger for experts than for novices regardless of the salience of the fact. The more someone knows about a topic area, the stronger can be the assumption that if a given fact were true, the person would know it.

In Experiment 1, we manipulated how much the protagonist knew about the area by describing the person as an expert or novice. However, differences in expertise can apply within individuals as well as between individuals. What is the nature of our intuitions about how much we know in various areas? How does someone decide to take his own or someone else's lack of knowledge in one particular domain as evidence against a fact but in another domain lack of knowledge as simply ignorance? In other words, how do we judge expertise in the course of an inference?

In some cases, we probably begin with a notion of how much we know in an area. We may think of ourselves as expert car mechanics or as film buffs and therefore be confident of our judgments in these areas. But what about the areas in which we have no explicit preconceived rating of our knowledge? One possibility is that our judgment of expertise is made in a manner intrinsic to the inference process. We may observe the intermediate products of our inference processes and from these observations derive a judgment of how expert we are in the domain. For example, when searching for information about an assertion in a given domain, we might notice whether we (or whoever is seeking the information) are able to find a great deal of information in that domain or only a small amount. Then, if we do not find any direct information about the particular assertion in question, we judge this lack of knowledge as more or less telling, depending on the amount of other information found in the domain. The more related information we have found, the more strongly we assume that we would also have found the assertion itself if it were true; therefore, the more strongly our lack of knowledge supports the negative conclusion about the assertion.

Experiment 2 was undertaken in order to investigate

the notion of intrinsic observation, as well as to broaden somewhat the other findings of Experiment 1. (This experiment actually preceded Experiment 1 chronologically but logically follows it.)

EXPERIMENT 2

Method

The basic methodology was similar to that of Experiment 1; subjects read scenarios and rated the likelihood of various facts or events. The major difference between the two studies was that the factor of set size was introduced in Experiment 2 instead of expertise, as in Experiment 1. The factor of set size refers to the number of items (e.g., number of birds) comparable to the target item that the protagonist has experienced (in the simple likelihood condition) or can remember (in the lack-of-knowledge condition). The factor of set size is different in structure from the comparable factor of expertise in Experiment 1; expertise applied only within the lack-of-knowledge condition, whereas set size in Experiment 2 applies in the simple likelihood condition as well. Another difference in methodology from Experiment 1 is that the basic task was preceded by an importance rating task and a simple likelihood rating task.

Subjects. The subjects were 14 undergraduates from Lesley College, Cambridge, Massachusetts, who had not served in other inference experiments. They were paid for their participation.

Materials. There were five scenarios, each with eight versions, which represented all possible combinations of three factors. These were judgment type (simple likelihood or likelihood given lack of knowledge), salience (high or low), and set size (two items or five items). The factors of judgment type and salience were as in Experiment 1. The factor of set size refers to the number of other items of the same kind that could be recalled or had been experienced by the protagonist of the scenario. For example, in the zoo scenario, the small set size included weasel and zebra, as in Experiment 1, and the large set size included weasel, zebra, gnu, seal, and lion. Table 2 shows two sample scenarios.

All eight versions of a given scenario were presented sequentially, instead of being randomly mixed as in Experiment 1. The order of versions was the same for all scenarios; Table 3 shows the order of presentation of scenario versions. Subjects rated a total of 48 versions of scenarios (8 versions each of the five experimental scenarios plus a sixth scenario that had to be omitted due to an error in stimulus construction). A summary of all the scenarios used is given in Table 4.

Procedure. Subjects performed the tasks in the order (1) importance ratings of all the items, (2) likelihood ratings of all items, in the absence of set-size information, and (3) the

Table 2
Sample Scenarios Used in Experiment 2

Condition 1 : 2 Likelihood High	
Walter has been watching a TV series on important American political officials. He has seen a television series about the President and the Head of the CIA. What is the likelihood that he has seen a television series about the Vice President?	
Condition 8 : 5 Lack of Knowledge Low	
Given that he [Walter] only remembers seeing a television series on the President, Head of the CIA, the Chief Justice, the Secretary of Defense, and the Representative from Alaska, and does not remember seeing a series on the Governor of Maryland, what is the likelihood that he has seen a series on the Governor of Maryland and has just forgotten seeing it?	

Table 3
Order of Presentation of Scenario Versions in Experiment 2

Order of Scenarios	Type of Judgment	Set Size	Salience of Target
1	Likelihood	2	High
2	Likelihood	2	Low
3	Lack of Knowledge	2	High
4	Lack of Knowledge	2	Low
5	Likelihood	5	High
6	Likelihood	5	Low
7	Lack of Knowledge	5	High
8	Lack of Knowledge	5	Low

basic task: ratings of simple likelihood and likelihood given lack-of-knowledge for the high- and low-salience targets and different set sizes.

Pretask ratings of importance and likelihood. Our predictions depended on the assumptions that the high-salience target was more important or memorable within the category than the low-salience target and that, on the average, the items in the small and large memory sets were intermediate in salience, since they served as background information. To verify this, the same subjects were asked to rate the importance of each of these seven items: the high and low target and the five members of the large set (which included the two items of the small set). For each scenario, additional items were also rated, to bring the total to 11-15 items/scenario. The ratings were on a scale of 1 (of almost no importance) to 5 (extremely important). The importance ratings were set in the general context of the scenario, but without the protagonist. For example, in the scenario concerning a television series on American politics, the question was put: "How important is each of the following offices in the overall American political system?"

The next task was for subjects to give a priori likelihood ratings of each event, using scenarios similar to those used in the experimental task, but without mentioning the protagonist's state of knowledge, for example: "Walter is watching a television series about the most important American public officials. How likely is it that each of the following offices will be included?"

After these tasks, subjects proceeded to the basic scenarios task, as described above.

Results and Discussion

Ratings of importance and likelihood. The subjects' importance ratings confirmed that the experimental stimuli corresponded to the desired salience order. The mean importance rating for the high-salience targets was 4.4, and for the low-salience targets, 3.4. The standard deviation was 1.1. The mean importance ratings for both the small and large sets were 3.8; these were intermediate between those of the high and low targets, as desired. The mean a priori likelihood rating (with a scale of -5 to +5) for the high-salience target was 4.5, and for the low-salience target, .2, with a standard deviation of 2.9. The mean rating for the small set was 2.2, and for the large set, 2.9. The a priori likelihood results are discussed further, below.

Major predictions. The major results are shown in Figure 2. These results are in agreement with two major findings of Experiment 1. First, the likelihood of an inference, given lack-of-knowledge information, is lower than the likelihood of a simple positive induction. The mean values are 1.73 for simple likelihood and .76 for

Table 4
Summary of the Scenarios Used in Experiment 2

Set		Target	
Small	Large	High	Low
		Burglary	
camera transistor radio	camera transistor radio stereo one hundred dollars electric toothbrush	TV	Timex watch
		Political Officials	
President Head of CIA	President Head of CIA Chief Justice Secretary of Defense Representative from Alaska	Vice President	Governor of Maryland
		Discussion Topic for Advertising Agency	
choice of lawyer sell more stock	choice of lawyer sell more stock enlarge catalog stock market sinking how to attract more clients	new contract with IBM	annual budget for stamps
		Minerals	
gold diamonds	gold diamonds uranium copper zinc	oil	titanium
		Zoo	
weasel zebra	weasel zebra gnu seal lion	elephant	lynx

likelihood given lack of knowledge. Second, the drop in likelihood, given lack-of-knowledge information, is more profound for high-salience targets than for low-salience targets. The mean values for high-salience and low-salience targets in the simple likelihood condition are 2.83 and .64, respectively; in the lack-of-knowledge condition, the mean values are 1.30 and .23. Thus, for high-salience targets, lack-of-knowledge information

depresses the simple likelihood by 1.53 points; for low-salience targets, the drop is .41 points. A three-factor within-subjects analysis of variance confirmed the significance of the main effects of judgment type [$F(1,13) = 26.54, p < .001$] and salience [$F(1,13) = 19.53, p < .001$] and of the interaction between judgment type and salience [$F(1,13) = 14.00, p < .005$]. Thus, the results of Experiment 2 replicate the findings of Experiment 1 with respect to judgment type and salience.

The other significant effects are the main effect of set size and the interactions between set size and judgment type and between set size and salience. These patterns are discussed in some detail below. Briefly, the set-size effect appears to be due to higher overall likelihood ratings for the large set size than for the small set size. The interaction of set size and judgment type appears to be due to a greater drop in likelihood from the simple likelihood ratings to the lack-of-knowledge ratings in the small set size than in the large set size.

Finally, the interaction between set size and salience appears to be due to a greater difference between low- and high-salience targets for the small set size than for the large set size. Focusing on the simple likelihood curve, notice that its slope is steeper for the small-set

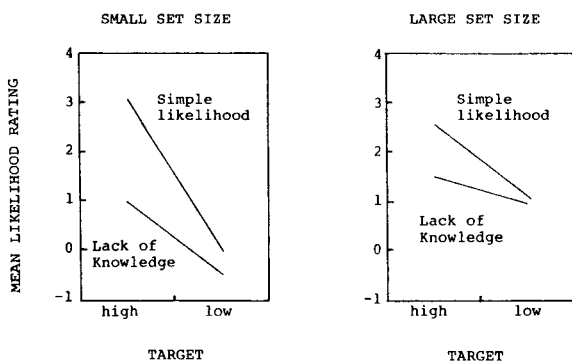


Figure 2. Results of Experiment 2: Mean ratings of the likelihood of target assertions of high or low importance for both the simple likelihood condition and the lack-of-knowledge condition, in the context of either a small set or a large set of similar assertions.

than for the large-set condition. There is an interesting connection here with the results of the a priori likelihood pretask. The slope of the simple likelihood curve is even steeper than that for the two-item set size: The mean likelihood for high-saliency items is 4.2, and for low-saliency items, .09. Since the simple likelihood task was, in effect, a zero set-size condition, it appears that there is a steady progression in steepness of curve from zero set size, the steepest curve, through set sizes of two and five. The difference between high-saliency and low-saliency targets is greater the smaller the set size. It appears that the greater the number of items in the set, the less effect the particular qualities of the target item have.

The factor of set size was introduced in Experiment 2 to test whether a person judges expertise in a domain by intrinsic observation during the search process. The results are discrepant from the predictions in two ways. First, they are contrary to the predictions of the intrinsic observation hypothesis. According to this hypothesis, when considering a memory search in a given domain, people use the number of similar items recalled from the domain as a measure of expertise in the area. The more things recalled, the higher the expertise judgment, and therefore, the more likely people are to attribute a failure to recall an item to the nonexistence of the item, rather than to faulty knowledge. If subjects had applied this reasoning in our study, then the lack-of-knowledge inference should have been stronger in the five-item case than in the two-item case. This prediction was not confirmed.

As Figure 2 shows, the negative effect of lack-of-knowledge information as against simple likelihood information was less for the large set than for the small set. This is indicated by the fact that the interaction between set size and judgment type was significant in the direction opposite to that predicted [$F(1,13) = 8.76$, $p < .05$]. It is further confirmed by the significant main effect of set size [$F(1,13) = 6.87$, $p < .05$], which was due to the fact that the mean likelihood was higher for large set size for both the lack-of-knowledge condition and the simple likelihood condition.

A second discrepancy arises if we consider this experiment as a replication of Experiment 1. For both of the set sizes, the slope of the lack-of-knowledge curve appears to be different from the slope of the corresponding curve in Experiment 1. In Experiment 1, the slope of the lack-of-knowledge curve was opposite to that of the simple likelihood curve: High-saliency targets, given lack of knowledge, were considered less likely than low-saliency targets, as reflected in a strong interaction between judgment type and saliency. Although this interaction was also significant in Experiment 2 [$F(1,13) = 28.95$, $p < .001$], the effect was not as large. The lack-of-knowledge curves in Experiment 2 sloped in the same direction as the simple likelihood curve: The effect of lack of knowledge was not great enough to reverse the relative likelihood of high- and low-saliency targets.

Thus it appears that (1) the difference between the two set-size conditions was opposite to that predicted and (2) the lack-of-knowledge effect in Experiment 2 was not as strong as that in Experiment 1. These discrepancies seem to disconfirm the intrinsic observation hypothesis. The number of items recalled from a domain does not seem to bear a strong relation to the subjective expertise of the recaller. However, we should consider other possible explanations for the differences between the expertise manipulations in the two experiments.

First, in making predictions concerning the set-size manipulation, it must be borne in mind that this manipulation affects not only the strength of the lack-of-knowledge inference, but also the strength of the positive induction. Research on concept formation supports the intuition that the likelihood of an event is judged greater if many similar events are known to have occurred (Hunt, Marin, & Stone, 1966). Thus, in our experiment, if subjects hear that five items of a certain kind were experienced, their estimate of the simple likelihood of another similar item should be greater than if they hear that only two items were experienced. This means that the strengths of the positive induction and the negative lack-of-knowledge inference vary in the same manner: Both should increase in strength as set size increases from two to five.

Thus the first discrepancy, that the likelihood drop given lack of knowledge is less for the large set than for the small set, might be partly accounted for by assuming that the positive induction increases more than the negative inference as set size increases from two to five. However, this cannot be the whole explanation. Examination of the data revealed that the increase in likelihood owing to the induction from five rather than two items (in the simple likelihood condition) was not sufficient to account for the increase in likelihood in the lack-of-knowledge condition. The average simple likelihood across high and low targets was 1.66 for the small set and 1.81 for the large set (an increase of .15). Lack-of-knowledge likelihood is .34 for the small set and 1.18 for the large set (an increase of .84). The inductive increase of .15 in simple likelihood from small set to large set is too small to account for the increase of .84 in the lack-of-knowledge case. It appears we cannot assume that the lack-of-knowledge inference is merely swamped by the positive induction in the five-item case. It is actually weaker for five items than for two. Why might this be?

One possible contribution to the first discrepancy between the set-size manipulation and the expertise manipulation (i.e., the weaker effect of lack-of-knowledge for large than for small set) is that the subjects may have applied their psychological intuitions about memory. In typical free recall situations, a person is less likely to recall another item within a category after five items than after two items (Anderson, 1976; Moeser, 1979; Postman, 1961; Smith, Adams, & Schorr, 1978). If subjects applied this reasoning, it would weaken the lack-of-knowledge inference in the five-item set, since

memory interference in the five-item case would provide an alternate explanation for the protagonist's failure to remember the target event. This explanation requires us to posit fairly sophisticated processing among our subjects. Further, it does not account for the second discrepancy, the slope difference: that the likelihood, given lack of knowledge, is higher for high-salience than for low-salience targets in Experiment 2.

Perhaps, instead, subjects went in the other direction and simplified the complex scenarios by ignoring the metajudgment aspects. For example, subjects were told that the protagonist could remember seeing a set of five birds, but not the target bird, and were asked to judge how likely it was that the protagonist had in fact seen the target and forgotten it. They might instead simply judge the likelihood that someone could have seen the five and not the target. However plausible this sort of simplification process might seem, it does not fit the data. Any such strategy would predict lower ratings for high-salience targets than for low-salience targets, since it is more unlikely that someone might never have seen a robin, for example, than a wren. Thus, simplification of the problem statement does not seem to account for the two discrepancies.

One possibility that can account for both discrepancies is that the set-size manipulation in Experiment 2 was weak as an expertise manipulation, because it did not lead subjects to think about the state of knowledge of the protagonist. In Experiment 1, each scenario explicitly referred to the state of knowledge of the protagonist (i.e., as a novice or expert). These descriptions were not used in Experiment 2. Instead, it was assumed that subjects would infer the state of knowledge from the number of items recalled. This could account for the second discrepancy, relating to the difference in the slope for the lack-of-knowledge condition between the two experiments. To the extent the subjects failed to focus on the protagonist's state of knowledge, the lack-of-knowledge inference would be weakened. This would result in the slope of the lack-of-knowledge curve in Experiment 2 being more like the slope for the simple likelihood judgment, which reflects a pure positive induction.

It also can account for the first discrepancy, relating to the decrease in the effect of the lack-of-knowledge manipulation for large set size (contrary to the predictions of the intrinsic observation hypothesis). The supposition here is that the larger number of items caused subjects to focus more on the size of the stimulus set, and hence on the positive induction, in the large set-size condition than in the small set-size condition. Thus, the positive induction may have been relatively stronger with the large set size, and the lack-of-knowledge inference may have been relatively stronger with the small set size. Therefore, given a large set size, the curve for the lack-of-knowledge condition would move toward the curve for the simple likelihood condition. This has two consequences with respect to the first

discrepancy: (1) The overall mean for the lack-of-knowledge condition would be greater for large set size than for small set size, and (2) the slope of the curve would be less for large set size than for small set size, since, as noted above, the increase in set size decreases the slope for the simple likelihood judgment, which is our measure of a pure positive induction. In summary, the argument is that the first discrepancy arises because the increase in set size acted to make the lack-of-knowledge condition more like the simple likelihood condition. This was because it focused subjects' attention still more strongly on the inductive inference, as opposed to the state of knowledge of the protagonist.

Thus, the lack-of-knowledge inference, though reliably present, appears to have been generally weaker, relative to the positive induction, in Experiment 2 than in Experiment 1. A possible area for further research is investigating the conditions that affect the relative strengths of the individual inferences that are combined in a person's overall likelihood judgment. This suggests an inquiry into variables that seem relevant to stimulus structure—and thus affect the positive induction—vs. those that seem relevant to subjective expertise—and thus affect the negative lack-of-knowledge inference. One stimulus relationship that might affect induction is the similarity between the target and the set. Kahneman and Tversky (1973) have shown that subjective judgments of the likelihood of an event may depend on the degree of similarity between the target event and the events known to have occurred, rather than on the mere number of events of that category known to have occurred. It seems likely that the strength of the positive induction inference could be more effectively manipulated by the degree of similarity between the target item and the other items discovered in memory than by the mere number of such items (Collins, Note 1).

A study by Rips (1975) suggests another variable that may affect the strength of the positive induction inference. He showed that induction is sensitive to category structure. Subjects in his experiments were more likely to infer that a property held true for a target item if they heard that it held true for a typical member of the same category than if they heard it held true for an atypical member.

The discussion of some of the possible factors relevant to stimulus structure makes it clear that the global variable of salience needs to be reconsidered. Salience was described earlier as "importance" or "memorability," and it was quantitatively estimated by subjects' importance ratings. But this is clearly an oversimplification. Among the factors that could enter into salience judgments are frequency, typicality, availability or memorability, similarity to the target item, and indeed any number of other conceptual relationships between the target item and the explicit or implicit context set (Kahneman & Tversky, 1972). It would be useful to obtain direct and indirect judgments of more of these interrelated factors in further research.

CONCLUSIONS

The two experiments described here give support to the notion of a lack-of-knowledge inference that takes place when people believe that a person would know if an item were present in memory. In this inference, the failure to find anything in memory about an assertion leads to the conclusion that the assertion is false. Our results indicate that the negative conclusion is stronger the more important the item and the greater the expertise of the person. This lack-of-knowledge judgment is not an all-or-none phenomenon; as we saw in the studies here, the negative implications of the lack-of-knowledge inference must be combined with the positive implications of various other reasoning processes, such as induction.

The lack-of-knowledge inference is related to the availability notion studied by Tversky and Kahneman (1973). They found that highly available events, that is, events that could easily be brought to mind (such as the existence of a word beginning with the letter k), were judged as more probable or more frequent than were events that came to mind less easily (such as a word ending with the letter k), even though the latter are more frequent. An interesting facet of both the availability notion and the lack-of-knowledge notion is that people assume that negative information is less likely to be stored than is positive information (Collins et al., 1975). Thus, if nothing has been stored as to whether an assertion is true, the reasoner decides that the assertion is false and has therefore not been stored as either false or true.

One of the most intriguing aspects of the lack-of-knowledge inference is the kind of self-observation strategies it requires. To make such an inference, people must evaluate the pattern of information in their own or another person's data base. The results of Experiment 1 showed that the factor of expertise in Experiment 1 is an example of this kind of evaluation. Our subjects clearly felt that if the protagonist was an expert, then any lack of knowledge in his or her domain of expertise should count heavily against the truth of the assertion, more so than for a novice. This has intuitive appeal and seems a sensible inference rule. However, it leaves open the question of what people do when they lack explicit information about the expertise of the protagonist. An important example of this situation is the introspective case: How does a person who experiences a lack of knowledge make a judgment of his or her own expertise in the relevant domain in order to know how to interpret that lack of knowledge?

In Experiment 2, we tested a simple hypothesis concerning the on-line evidence people might bring to bear on the question of their own or someone else's expertise. The intrinsic observation hypothesis states that, when trying to find an item in memory, people judge their expertise in the area by the number of other items they can remember. This hypothesis predicts that

subjects will judge an item as less likely if the protagonist who cannot recall the item is able to recall a great many other items. The experiments reported here provided no support for this hypothesis. Two things should be noted, however. First, our method required subjects to vicariously imagine the protagonist's state of knowledge. A more direct methodology, such as the Brewer and Treyns (1981) paradigm discussed above, might yield different results. Second, the version of the intrinsic observation hypothesis tested here is only one particularly simple model of how intrinsic judgments of expertise might be made. It is possible that the issue is more subtle than can be accounted for by any of the factors so far mentioned (i.e., number of comparable items recalled, similarity of other items recalled to the target, or relative typicality, availability, or memorability). Judgments of expertise may be based on structural considerations, such as the richness of the relational network in the domain or the amount of redundancy in the relevant knowledge.

The results of these studies indicate that metajudgments based on the importance of the material and the subjective expertise of the thinker affect the way in which inferences are made and combined in reasoning. We hope that further research will lead to greater understanding of the nature of self-knowledge and its role in inference.

REFERENCE NOTE

1. Collins, A. *Human plausible reasoning* (Final Rep.). Department of Defense, Defense Advanced Research Projects Agency, February 1978.

REFERENCES

- ANDERSON, J. R. *Language, memory and thought*. Hillsdale, N.J.: Erlbaum, 1976.
- BREWER, W. F., & TREYENS, J. C. The role of schemata in memory for places. *Cognitive Psychology*, 1981, 13, 207-230.
- BROWN, A. L. The development of memory: Knowing, knowing about knowing and knowing how to know. In H. W. Reese (Ed.), *Advances in child development and behavior* (Vol. 10). New York: Academic Press, 1975.
- BROWN, J., LEWIS, V. J., & MONK, A. F. Memorability, word frequency and negative recognition. *Quarterly Journal of Experimental Psychology*, 1977, 29, 461-473.
- CARBONELL, J. R., & COLLINS, A. Natural semantics in artificial intelligence. *Proceedings of the Third International Joint Conference on Artificial Intelligence*, 1973, 344-351.
- COLLINS, A. Fragments of a theory of human plausible reasoning. In D. Waltz (Ed.), *Theoretical issues in natural language processing, II*. Urbana: University of Illinois, 1978.
- COLLINS, A., WARNOCK, E. H., AIELLO, N., & MILLER, M. L. Reasoning from incomplete knowledge. In D. Bobrow & A. Collins (Eds.), *Representation & understanding: Studies in cognitive science*. New York: Academic Press, 1975.
- FISCHOFF, B., SLOVIC, P., & LICHTENSTEIN, S. Knowing with certainty: The appropriateness of extreme confidence. *Journal of Experimental Psychology: Human Perception and Performance*, 1977, 3, 552-564.
- FLAVELL, J. H., FRIEDRICH, A. G., & HOYT, J. D. Developmental changes in memorization processes. *Cognitive Psychology*, 1970, 1, 324-340.
- HAGEN, J. W. Strategies for remembering. In S. Farnham

- Diggory (Ed.), *Information processing in children*. New York: Academic Press, 1972.
- HUNT, E. B., MARIN, J., & STONE, P. *Experiments in induction*. New York: Academic Press, 1966.
- KAHNEMAN, D., & TVERSKY, A. Subjective probability: A judgment of representativeness. *Cognitive Psychology*, 1972, 3, 430-454.
- KAHNEMAN, D., & TVERSKY, A. On the psychology of prediction. *Psychological Review*, 1973, 80, 237-251.
- MOESER, S. D. The role of experimental design in investigations of the fan effect. *Journal of Experimental Psychology: Human Learning and Memory*, 1979, 5, 125-134.
- NISBETT, R. E., & WILSON, T. D. Telling more than we know: Verbal reports on mental processes. *Psychological Review*, 1977, 84, 231-259.
- POSTMAN, L. The present status of interference theory. In C. N. Cofer (Ed.), *Verbal learning and verbal behavior*. New York: McGraw Hill, 1961.
- RIPS, L. J. Inductive judgments about natural categories. *Journal of Verbal Learning and Verbal Behavior*, 1975, 14, 665-681.
- SMITH, E. E., ADAMS, N., & SCHORR, D. Fact retrieval and the paradox of interference. *Cognitive Psychology*, 1978, 10, 438-464.
- TVERSKY, A., & KAHNEMAN, D. Availability: A heuristic for judging frequency and probability. *Cognitive Psychology*, 1973, 5, 207-232.

NOTE

1. The scenarios are obtainable from the authors.

(Received for publication April 27, 1979;
revision accepted January 18, 1981.)