When a Coincidence Is Suspicious: The Role of Mental Simulation

Dale T. Miller
Princeton University

William Turnbull and Cathy McFarland
Simon Fraser University
Burnaby, British Columbia, Canada

Five studies examined Kahneman and Miller's (1986) hypothesis that events become more "normal" and generate weaker reactions the more strongly they evoke representations of similar events. In each study, Ss were presented with 1 of 2 versions of a scenario that described the occurrence of an improbable event. The scenarios equated the a priori probability of the target event, but manipulated the ease of mentally simulating the event by varying the absolute number of similar events in the population. Depending on the study, Ss were asked to indicate whether they thought the event was due to chance as opposed to (a) an illegitimate action on the part of the benefited protagonist, or (b) the intentional or unintentional misrepresentation of the probability of the event. As predicted, the fewer ways the events could have occurred by chance, the less inclined Ss were to assume that the low-probability event occurred by chance. The implications of these findings for impression-management dynamics and stereotype revision are discussed.

Reactions to events depend not only on the events themselves, but on what the events bring to mind (Kahneman & Miller, 1986; Kahneman & Tversky, 1982b). The thoughts, images, and scenarios evoked by an event constitute what Kahneman and Miller (1986) have termed the event's *norm.* Events that elicit representations that are primarily similar to them are said to be *normal,* whereas events that elicit representations that are primarily dissimilar to them are said to be *abnormal.* The norms evoked by events include expectancies about events, with the consequence that events that disconfirm expectancies are judged to be less normal than events that confirm expectancies. Norms, however, are not only based on retrieved representations of what was expected to be, but also on constructed representations of what might have been. Kahneman and Miller used the term *postcomputed* to describe post hoc counterfactual representations that are generated by the events themselves, and distinguished this class of representations from precomputed representations such as expectancies.

Evidence for the role of postcomputed representations in the evaluation of events comes from a series of studies that investigated the link between counterfactual thinking and affective reactions. Consider the following vignette taken from Kahneman and Miller (1986, p. 145):

> Mr. Jones almost never takes hitch-hikers in his car. Yesterday he gave a man a ride and was robbed. Mr. Smith frequently takes hitch-hikers in his car. Yesterday he gave a man a ride and was robbed. Who do you expect to experience greater regret over the episode?

The vast majority of subjects who read this vignette predicted that Mr. Jones would experience more regret. Why? Mr. Jones's fate is the same as Mr. Smith's, and his actions would seem neither more foolish nor more immoral. According to Kahneman and Miller, people expect Mr. Jones to experience more regret because they assume that it is easier for him to imagine refraining from the action that led to the misfortune.

Regret is not the only emotion influenced by the postcomputed availability of counterfactual alternatives. Miller and McFarland (1986) showed that the sympathy observers felt for victims of misfortune was influenced by the ease with which they could mentally simulate or imagine more positive outcomes for the victims. An unforeseeable negative fate preceded by an unusual act generated more sympathy in observers than an unforeseeable negative fate preceded by a habitual act. Observers apparently feel sympathy for victims whose fates need not have happened as well as for those whose fates ought not to have happened (Miller, Turnbull, & McFarland, in press).

The routineness of the causal chain that produced the event is not the only factor that affects the availability of alternative

This research was supported by Social Sciences and Humanities Research Council Grants awarded to Dale T. Miller and to Cathy McFarland and William Turnbull.

This article profited from the helpful comments of Gary Wells and two anonymous reviewers and valuable discussions with Geoff Fong, Keith Holyoak, Daniel Kahneman, and Ziva Kunda.

Correspondence concerning this article should be addressed to Dale T. Miller, Department of Psychology, Green Hall, Princeton University, Princeton, New Jersey 08544-1010.

Journal of Personality and Social Psychology, 1989, Vol. 57, No. 4, 581-589

constructions. Kahneman and Tversky (1982a) proposed that it also is easier to imagine oneself abstaining from actions that one has carried out than it is to imagine carrying out actions that were not in fact performed. The corollary of this hypothesis—that consequences of actions will evoke stronger emotional responses than consequences of failures to act—was tested with the following vignette (Kahneman & Tversky, 1982a, p. 171):

> Mr. Paul owns shares in company A. During the past year he considered switching to stock in company B, but he decided against it. He now finds out that he would have been better off by $1,200 if he had switched to the stock of company B. Mr. George owned shares in company B. During the past year he switched to stock in company A. He now finds that he would have been better off by $1,200 if he had kept his stock in company B. Who feels greater regret?

The overwhelming majority of subjects predicted that Mr. George would experience more regret over his action than would Mr. Paul over his equally costly inaction. The finding that acts of commission produce greater regret than acts of omission was replicated by Landman (1988) and is in accord with formulations that distinguish omission from commission in attributions of causality and responsibility (Hart & Honoré, 1959; Heider, 1958).

The foregoing research supports norm theory by demonstrating that an event's normality, along with the affective reaction it generates, depends on the strength with which it evokes alternative events. In the present research, we sought to demonstrate that an event's normality also depends on the strength with which it evokes similar events. Whereas previous investigations of mental simulation have tested the hypothesis that events become less normal the easier it is to undo them mentally, the present research tested the hypothesis that events become more normal the easier it is to mentally replicate them. As an illustration of our approach, consider the following thought experiment.

Imagine that you have a young child who loves chocolate chip cookies. Imagine further that you buy your cookies in packages that include oatmeal as well as chocolate chip cookies. Your child's practice is to go to the cookie jar and select the chocolate chip cookies, leaving the oatmeal ones to go stale. One day you think of a strategy to cope with this problem. You tell your child to close his eyes before he reaches into the jar, taking whichever cookie he grabs. He agrees to this and heads to the kitchen and the cookie jar. Shortly, he comes back, exclaiming that he did what you said and just happened to get a chocolate chip cookie.

How would you react to the child's announcement? Would you accept his claim that, by coincidence, he had selected a coveted chocolate chip cookie or would you be suspicious that he had peeked? What factors might affect your judgment in this regard? Information you had about the child's prior behavior would be one obvious candidate for consideration. A prior history of honesty would leave you less suspicious than a prior history of dishonesty. Knowledge you had concerning the strength of the child's preference for chocolate chip over oatmeal cookies would probably also be relevant. Suspicion would be expected to increase as the child's assumed motivation to cheat increased. In addition to your knowledge of the child, your level of suspicion might well be influenced by your knowledge of the contents of the cookie jar. If 50% of the cookies were chocolate chip, you almost certainly would not be as suspicious as you would be if only 5% were. Indeed, an event could not even appropriately be termed a coincidence, let alone be expected to generate suspicion, if it were not at least moderately improbable.

The relative frequency of the two types of cookie may not be the only feature of the jar's contents that would affect your suspicion. Contrast the levels of suspicion you might experience in two slightly different versions of the scenario. In Version A, there is 1 chocolate chip cookie and 19 oatmeal cookies in the jar; in version B, there are 10 chocolate chip cookies and 190 oatmeal cookies in the jar. The relative frequency of the two types of cookie is the same in the two versions of the scenario (1 to 19), but we suspect that if you are like us (and as is seen shortly, like our subjects) you would be more suspicious when the child had claimed to draw the only remaining chocolate chip cookie (Version A) than merely 1 of the 10 remaining chocolate chip cookies (Version B). Assuming for a moment that you share our intuitions, why might your reactions to the two scenarios be different? Basing feelings of suspicion on factors such as the actor's past behavior, his motivation, and the likelihood of the event occurring by chance seems fair and rational. But what is rational or fair about being influenced by the absolute number of desired cookies in the jar?

The differential reactions generated by the two scenarios may not be accounted for by naive theories of suspicion, but norm theory does offer a possible account. The key to this account is the distinction between precomputed judgments of probability and postcomputed judgments of normality. The event of selecting a chocolate chip cookie may have been equally probable in the two scenarios (5%), but it was not equally normal. When there are 10 chocolate chip cookies, and hence 10 similar ways for the event to occur, the purported selection of a chocolate chip cookie is more normal than when there is only 1 chocolate chip cookie, and hence only 1 route or means by which the event could occur. Thus, the selection of the chocolate chip cookie in the 10/190 case generates less suspicion than it does in the 1/19 case, because there are more different ways for the event to happen in the former case than in the latter.

In this article, we report a series of five studies that take the same general format. Subjects read one of two versions of a scenario in which a highly improbable event either occurs or is alleged to have occurred. The probability of the target event is kept constant across the two versions, but the normality of the event is varied by manipulating the size of the relevant population, as occurred in the cookie-jar example. Suspicion that the occurrence of the target event was not simply a coincidence was assessed in a number of ways. In Studies 1 and 2, subjects rated how suspicious they were that the protagonist had obtained the desired outcome by some procedure other than the prescribed random one. In Study 3, subjects rated how suspicious they were that the protagonist had lied about the distribution of possible outcomes in the population; and in Study 4, subjects supplied their own estimate of the frequency of the target event in the population. Last, in Study 5, subjects rated their confidence in the claims of "experts" who offered different accounts of the frequency of the target attribute in the population.

Study 1

The cookie-jar scenario required us to evaluate the plausibility of two hypotheses. The first of these hypotheses suggests that the draw reflected chance; the second, that the draw reflected an illegitimate act (peeking). We predicted that one factor that affects people's preference for one of these hypotheses over the other is the number of ways they can imagine the event happening by chance. Specifically, we predicted that the more ways the event could have happened by chance, the more likely it is that individuals will believe that the unexpected event did happen by chance. We tested this hypothesis in Study 1 by using the cookie-jar scenario described earlier.

Method

Subjects. A total of 45 undergraduates (both men and women) served as subjects. They participated as part of a class demonstration.

Procedure. Subjects were asked to complete a questionnaire describing one of two scenarios identical in all respects except for the information pertaining to the normality of the target event. The scenario follows, with the manipulated information in parentheses:

> Imagine that you have a young child who loves chocolate chip cookies. Imagine further that you buy your cookies in packages that include oatmeal as well as chocolate chip cookies. Your child's practice is to go to the cookie jar and select the chocolate chip cookies, leaving the oatmeal ones to go stale. One day you think of a strategy to cope with the situation. You tell your child to close his eyes before he reaches into the jar, taking whichever cookie he grabs. He agrees to this and heads to the kitchen and the cookie jar. The jar contains 1(10) chocolate chip cookie(s) and 19(190) oatmeal cookies. Shortly, he comes back, exclaiming that he did just what you said and he selected a chocolate chip cookie.

After subjects had read the scenario, they were asked to indicate on a 9-point scale how suspicious they would be that the child peeked into the jar before selecting the cookie (1 = *Not at all suspicious*, 9 = *highly suspicious*).

Results

The probability of the child's randomly selecting a chocolate chip cookie in the two scenarios was identical (1/20 vs. 10/200). Consequently, subjects' suspicion about the child's honesty, if based solely on their estimates of the probability that he could have drawn a chocolate chip cookie by chance, should not have varied across the two conditions. Their level of suspicion should have varied across the two conditions, however, if it was also influenced by the normality of a target outcome (i.e., the selection of a chocolate chip cookie). Specifically, subjects should have expressed greater suspicion when there was only 1 way the outcome could have occurred than when there were 10 ways it could have occurred. The results confirmed this hypothesis. Subjects expressed greater suspicion about the child's honesty when there was only 1 chocolate chip cookie in the jar (M = 6.91) than when there were 10 (M = 5.13), $t(43) = 2.30$, $p < .05$.

Although the results are consistent with our analysis, two alternative explanations exist. First, the condition differences in level of suspicion may reflect the impact that the manipulation had on subjects' perceptions of the likelihood that the child would attempt to cheat (peek) in this situation. Specifically, subjects may have believed that a child is more likely to peek when there are only a few cookies in the jar than when there are many. The second explanation derives from research demonstrating that subjective probability estimates are sometimes influenced by the absolute number of instances in addition to the proportion of instances (Estes, 1976; Hamilton & Gifford, 1976; Silka & Albright, 1983). It is possible, therefore, that the present results may reflect the impact that the manipulation had on subjects' estimates of the subjective likelihood that the child would select the desired cookie. Thus, even if subjects could easily calculate that the objective probability of getting a chocolate chip cookie was 5% in both conditions, their subjective estimates may have been higher (and their resultant suspicion lower) when the absolute frequency of the desired cookie was high (10) than when it was low (1).

To test the first explanation, we presented two additional groups of subjects (20 in each group) with descriptions of the scenarios that left the outcome unstated. The scenarios were introduced with the following statement: "People often form expectancies about what they think is likely to happen in various situations. We are interested in the expectancies people form based on minimal information. Please read the following story and try to predict what will happen." After reading the scenario, subjects were asked to estimate the likelihood (on a 9-point scale ranging from *not at all likely* to *extremely likely*) that the child would attempt to peek. There were no significant differences in the likelihood estimates of the two groups (small group, $M = 6.05$, $SD = 2.56$; large group, $M = 6.05$, $SD = 2.70$; $t < 1$).

We used an additional two groups (20 subjects per group) to test the second alternative explanation. Each of these two groups was also presented with one of the two versions of the scenario (again omitting the alleged outcome); however, their task was simply to estimate the likelihood that the child would select a chocolate chip cookie, with no reference being made to the prospect of the child's peeking. Subjects answered this question on a 11-point scale that ranged from 0% to 100%.[1] There were no significant differences in the likelihood estimates of the two groups (small group, $M = 21.25\%$, $SD = 31.33$; large group, $M = 19.00\%$, $SD = 25.26$; $t < 1$). Note that the mean likelihood ratings differ markedly from those based on the objective probability (5%) of obtaining a chocolate chip cookie. Presumably this discrepancy reflects the impact that the nonnumerical aspects of the scenarios had on subjects' judgments of likelihood.

[1] It could be argued that the second of the control questions encompasses the first. That is, subjects' suspicion concerning the likelihood that the child would peek should be reflected in their responses to the question that asks them more generally about the likelihood of the desired outcome occurring (whether through peeking or chance). We included the former, more specific, question because we wanted to be sure we had included a direct and specific test of what we thought was the most plausible alternative account for the findings.

Discussion

The results provide evidence that two events that are judged to be equally likely can generate different degrees of suspicion. The shadow of reasonable doubt accorded to the target of suspicion extends as the normality of the event increases. The more easily subjects can imagine or mentally simulate something happening by one route (chance), the less suspicious they are that it occurred by some other route (peeking). Subjects' perceptions of the child's honesty were very possibly also affected by the probability that they assigned to the event. If 50% of the cookies were chocolate chip, then they surely would have expressed less suspicion. Moreover, considering probability in this manner would both be psychologically sensible and comport with legal philosophy. Exhortations to juries not to convict someone unless they are sure "beyond a reasonable doubt" invites them to consider the likelihood that a set of occurrences could have come about without the accused being guilty of the offense with which he or she is charged. However, neither normative nor legal dictates would seem to justify considering the number of different ways the event could occur when judging the probability that the target violated the prohibition against peeking.

Study 2

The results of Study 1 indicated that the fewer ways a low probability event could have occurred by chance, the less inclined subjects were to assume that the event occurred by chance. The task of judging the potential coincidental nature of outcomes is, of course, not restricted to contexts involving random draws. We are rarely certain that an outcome reflects only one particular cause, and thus continually have to decide whether an outcome that is consistent with one particular hypothesis constitutes evidence of the validity of that hypothesis or simply represents a coincidence (chance). Researchers are able to utilize statistics to help them make these decisions, but the layperson must rely on his or her own intuitions. Study 2 examines how the normality of an outcome affects subjects' judgments in one such situation.

Method

Subjects. Subjects were 40 undergraduates (both men and women) who were approached in a university cafeteria.

Procedure. Subjects were asked to complete a questionnaire describing one of two scenarios identical in all respects except for the normality of the target event. This scenario follows, with the manipulated information in parentheses:

> John S. is a supervisor in a local manufacturing firm. John is responsible for promoting the employees in his department. In the past he has been accused of being against equal rights and opportunities for women. There is (are) 1(10) male(s) and 9(90) females in his department who are potential candidates for promotion. John decides to give these employees a written examination to help with his decision. John grades these exams himself, and reports that the highest mark was obtained by a man, whom he promotes.

After subjects had read the scenario, they were asked to indicate on a 9-point scale how suspicious they would be that John's grading of the written examination was unfair (biased) (1 = *not at all suspicious,* 9 = *highly suspicious*).

Results and Discussion

We hypothesized that subjects would be more suspicious about the fairness of the examination grading when the man who allegedly received the highest score came from a department with only 1 man than when he came from a department in which there were 10 men. The results confirmed this hypothesis. Subjects expressed greater suspicion that there was a bias against women when there was only 1 man who could have received the highest score ($M = 6.10$) than when there were 10 men who could have received the highest score ($M = 4.80$), $t(38) = 2.11, p < .05$.

As with Study 1's results, alternative explanations can be proposed for these findings. First, the tendency to confuse frequency and proportion remains a possible alternative explanation here as well. In addition, our manipulation may have tapped naive social theories that predisposed subjects to experience differential suspicion in the two conditions. For example, subjects may believe that supervisors in small companies are more likely than those in large companies to be unfair and to act on personal bias. A second possibility is that subjects may believe that a supervisor would know a male candidate better when he was the only man than when he was 1 of 10 men, and thus would be better able to bias the test in his favor. To assess these possibilities, we presented a new sample of 80 subjects with either the small or large version of the scenario (omitting the outcome). Paralleling Study 1's procedure, these subjects were divided into four groups of 20 and were asked to indicate either the likelihood that the supervisor would lie about which candidate received the highest grade on the exam or simply the likelihood that a man would achieve the highest score on the exam. There were no condition differences for ratings of the likelihood that the supervisor would lie (small group, $M = 4.00$, $SD = 2.20$; large group, $M = 4.40$, $SD = 1.88$; $t < 1$) or for ratings of the likelihood that a man would achieve the highest exam score (small group, $M = 43.00\%$, $SD = 29.04$; large group, $M = 47.00\%$, $SD = 33.57$; $t < 1$).

Study 3

Studies 1 and 2 manipulated the number of legitimate ways a low-probability event could have occurred in an attempt to affect subjects' suspicion about the fairness of the process that generated the outcome. The subjects' task was to determine whether an event, which was both improbable and consistent with the protagonist's wishes, constituted a fortuitous coincidence or reflected a nonrandom, and hence illegitimate, procedure. By specifying the exact nature of the relevant population, these studies precluded subjects from being suspicious about another feature of the scenario: the alleged frequency (proportion) of the target attribute in the population. For example, subjects in Study 1 were not asked whether the purported random selection of the chocolate chip cookie affected their suspicion that there was more than the alleged number of chocolate chip cookies in the jar. Similarly, subjects in Study 2 were not asked

how the fact that a man had achieved the highest score on the exam affected their suspicion that there were more than the alleged number of male employees in the department. Had the scenarios not stipulated the frequency of the event in the distribution with such certainty, it is conceivable that the differential normality of the event would have resulted in differential suspicion about the event's frequency.

Study 3 explored the possibility that the normality of an improbable event affects people's inclination to question the alleged a priori probability of the event. Specifically, we hypothesized that the fewer similar instances there are purported to be in the population from which the improbable event springs, the more inclined people will be to believe that the event is, in fact, more probable than they had been led to believe.

Method

Subjects. Subjects were 30 sunbathers (both men and women) who were approached on a beach.

Procedure. Subjects were asked to complete a questionnaire describing one of two versions of a scenario identical in all respects except for the normality of the target event. The scenario follows, with the manipulated information in parentheses:

> Imagine that you rented a car from a company called "Rent a Clunker." You went to this company because its rates were much cheaper than other car rental companies but you were worried about the reliability of the cars. The manager assured you that his cars were very reliable, claiming that only 2(20) of his 20(200) cars had ever had problems. You choose a car and leave for a long trip. Within an hour the car breaks down.

After subjects read the scenario, they were asked to indicate on a 9-point scale how suspicious they would be that the manager had lied to them about the performance record of his cars (1 = *not at all suspicious,* 9 = *highly suspicious*).

Results and Discussion

We predicted that subjects would express more suspicion about the manager's integrity when he alleged there were only 2 unreliable cars in his fleet than when he alleged there were 20. Guiding this prediction was the assumption that encountering 1 of 2 possible unreliable cars would seem a less normal event than encountering 1 of 20 possible unreliable cars, even though the probability of encountering an unreliable car in the two cases was the same (10%). The results supported the hypothesis. Subjects expressed more suspicion in the small-fleet condition ($M = 7.60$) than in the large-fleet condition ($M = 5.87$), $t(28) = 2.14$, $p < .05$.

A new sample of 80 subjects was used to help rule out the possibility that the observed differences in suspiciousness reflected differential expectations that were engendered either by the numbers themselves or by the relation between the numbers and the specific content of the scenario (i.e., car rental companies). We presented 40 subjects (20 in each condition) with either the small or large population-size version of the scenario (omitting the outcome) and asked them to estimate the likelihood that the manager of the company had lied about the reliability of his cars. The remaining 40 subjects (20 in each condition) were presented with the two versions of the scenario and simply asked to estimate the likelihood that someone would get an unreliable car if they were to rent one from this company. There were no group differences in either the responses to the first question (small group, $M = 5.95$, $SD = 2.46$; large group, $M = 5.50$, $SD = 2.33$, $t < 1$), or in the responses to the second question (small group, $M = 16.50\%$, $SD = 13.87$; large group, $M = 17.00\%$, $SD = 16.57$; $t < 1$).

Study 4

In Study 3, subjects were asked how suspicious they would be that a person had intentionally underestimated the probability of a negative event, given the occurrence of this ostensibly improbable event. Study 4 provided a modified replication and extension of Study 3 by asking subjects to estimate the actual frequency of the target event. We hypothesized that subjects would estimate the actual incidence of a target characteristic in a population to be higher the less normal they found the occurrence of that characteristic to be.

Method

Subjects. Subjects were 30 sunbathers (both men and women) who were approached on a beach.

Procedure. Subjects were asked to complete a questionnaire describing one of two versions of a scenario identical in all respects except for the normality of the target event. The scenario follows, with the manipulated information in parentheses:

> Imagine that you have gone to visit a town on vacation. A friend has told you that 50% of cab drivers in this town are reckless. You call the Royal Taxi Company and are told by the owner that there are 10(1,000) taxis in his firm. Further, he assures you that most of his taxi drivers are good drivers, although he admits that 2 (200) of them are occasionally reckless.
>
> A couple of days later you hail a Royal taxi cab. During the ride you are almost hit by another car when your driver goes through a red light. To your dismay you have happened upon a reckless cab driver.

After subjects had read the scenario, they were asked to indicate on a 5-point scale what they thought the actual percentage of reckless drivers in this firm was (10%, 20%, 30%, 40%, or 50%).

Results and Discussion

We hypothesized that subjects would be less likely to think it was a coincidence that the first taxi driver they encountered was reckless when there allegedly were only 2 reckless drivers in the firm than when there allegedly were 200. This hypothesis was confirmed. Subjects estimated that the actual percentage of reckless drivers was higher when there were supposedly only 2 reckless drivers for them to encounter ($M = 45.33\%$) than when there were supposedly 200 reckless drivers for them to encounter ($M = 34.66\%$), $t(28) = 2.95$, $p < .01$.

Following the procedure of the previous studies, four additional groups of subjects were used. We presented 40 additional subjects (20 in each condition) with either the small- or large-population-size version of the scenario (omitting the outcome) and asked them to estimate the likelihood that the owner of the company lied about the number of reckless drivers in his firm. An additional 40 subjects (20 in each condition) were presented

with one of the two versions of the scenario (omitting the outcome) and were asked to indicate the likelihood that someone would encounter a reckless driver if they were to take a cab from that company. There were no group differences in either the responses to the first question (small group, $M = 5.45$, $SD = 2.11$; large group, $M = 5.20$, $SD = 2.42$; $t < 1$) or the responses to the second question (small group, $M = 22.0\%$, $SD = 13.22$; large group, $M = 27.00\%$, $SD = 18.09$; $t < 1$).

Study 5

Study 4 provided subjects with both a low- and a high-incidence hypothesis and asked them, on the basis of a single observation, to adjudicate between them. The low-incidence hypothesis was proposed by a target who had a vested interest in convincing people that the incidence was low, thereby raising the possibility that he willfully misrepresented the population incidence of the event (reckless drivers). Frequently, however, people receive conflicting hypotheses in circumstances in which the question of vested interest on the part of the informants does not arise. According to our analysis, even here the normality of the event should affect people's willingness to accept the possibility that an improbable event is a coincidence. To illustrate this hypothesis, consider the following thought experiment. Imagine that you visit a city in which there are members of an ethnic group with which you are not familiar. Although you have no personal experience with members of this ethnic group, assume that you have been given conflicting accounts of their social manner. One of these accounts suggests that most members of this group are extremely rude. The other account suggests that only a small minority (no more than 5%, say) of the group members are extremely rude and that the vast majority are very courteous. Now imagine that you have your first encounter with a member of this group and discover that this individual is extremely rude. This experience obviously would be consistent with the negative account that you had heard and inconsistent with the positive account. Nevertheless, it remains possible that the positive account was the correct one and that the first member of the group you encountered was atypical. Before you reflect on how prepared you would be to entertain the atypical-instance hypothesis, consider an additional piece of information. Imagine that the size of the ethnic group in the town was either small (fewer than 20 members) or very large (more than 50,000 members). How do you think your knowledge of the absolute number of rude group members would affect your preference for the high-incidence versus the low-incidence hypothesis on the basis of an encounter with 1 rude member? Specifically, how willing would you be to dismiss your experience as unrepresentative and to continue to entertain the hypothesis that the vast majority of the group members are courteous? The hypothesis guiding Study 5 suggests that your confidence in the accuracy of the high-incidence (negative) account would be higher when the group was small than when it was large. The larger the group, the more rude individuals there are for you to meet—even if they are a minority—and hence, the more normal it would be for you to encounter one.

Method

Subjects. Subjects were 36 undergraduates (both men and women) who participated as part of a class demonstration.

Procedure. Subjects were asked to complete a questionnaire describing one of two versions of a scenario identical in all respects except for the normality of the target event. The scenario follows, with the manipulated information in parentheses:

> Imagine that you are an anthropologist who decides to study a tribe living in a remote area of South America. Two other anthropologists, who have investigated the tribe previously, provide you with differing views of the friendliness of the tribe members. One investigator says that at least 20(200) of the 40(400) members are hostile and unfriendly, whereas the other investigator claims that only 2(20) of the 40(400) tribe members are hostile and unfriendly. You visit the tribe and the first tribe member you encounter is hostile and unfriendly.

After subjects had read the scenario, they were asked to answer the following question on a 9-point scale: "On the basis of your first encounter with a member of the tribe, how confident would you be that the investigator who had claimed that at least 20(200) of the tribe members were hostile and unfriendly had provided the more accurate of the estimates?" (1 = *Not at all confident*, 9 = *Extremely confident*).

Results and Discussion

We predicted that subjects would be more inclined to endorse a high-incidence over a low-incidence hypothesis the less normal the target event appeared in light of the latter hypothesis. This prediction was confirmed. Subjects were more confident that a hypothesis specifying a high population incidence of the observed characteristic (hostility) was correct when the low-incidence hypothesis specified that there were only 2 members of the population who had that characteristic ($M = 6.90$) than when it specified that there were 20 members who had that characteristic ($M = 5.24$), $t(34) = 2.35$, $p < .05$.

General Discussion

The thesis of this research is that events are evaluated in relation to the representations that they bring to mind, as well as in relation to the expectations that precede them. Reactions to events can be strong or weak depending on their relation to the alternatives that the occurrence of the event recruits from memory or generates in imagination. The more strongly events evoke similar instances, the more normal the events will seem, and the weaker the reactions that they evoke will be.

In the present five studies, we manipulated the strength with which an event evoked similar instances by varying the absolute (but not the relative) number of similar instances in the sample space. We assumed that the more similar instances there allegedly were in the sample, the more normal the occurrence of the event would be. More specifically, each study presented subjects with an improbable event and asked them, either explicitly or implicitly, to indicate how suspicious they were that the outcome represented something other than a coincidence. Common to all of these studies was the prediction that subjects' willingness to accept an event as a coincidence would be affected by the absolute number of ways in which the event could have occurred.

In Studies 1 and 2, we presented subjects with scenarios in which a target produced an improbable outcome that coincided with his or her wishes. Subjects' task was to decide whether this occurrence represented a coincidence or some nonrandom, and hence illegitimate, factor. Although the a priori probability of the event was identical in the two versions of the scenario, the more ways the outcome could have occurred by chance the less suspicious subjects were of the means by which the outcome came about.

In Study 3, subjects were presented with a scenario in which an event occurred that they had been led to believe was highly unlikely. Subjects were asked to decide whether the occurrence represented a coincidence or the fact that the commonness of the event in the distribution had been misrepresented to them by an informant. The event was equally unexpected in the two versions of the scenario, but differed in its normality. As predicted, the less normal the occurrence of the event was, the more suspicion subjects expressed about the informant's honesty.

In Studies 4 and 5, we presented subjects with scenarios in which an event occurred that one source had suggested was more probable than had another source. The subjects' task was to judge the accuracy of the different estimates they had received in light of the event's occurrence. The a priori probability values that the high- and low-incidence hypotheses assigned to the target event differed from one another but did not differ from their counterparts across the two versions of the scenario. Nevertheless, the fewer ways the scenario implied there were for the event in question to have occurred, the more subjects were inclined to accept the high estimate.

Alternative Interpretations

In each of the five studies presented in this article, we manipulated the absolute number of ways in which an event could occur and thereby presumably the ease of imagining the event's occurring. The variable of number seemed to us to be a relatively straightforward means of manipulating ease of imagining. However, by relying on number for our experimental manipulation we left our results vulnerable to alternative accounts that must be taken seriously. One of these alternatives concerns the psychophysics of numbers. Research indicates that people judge the distance between smaller numbers (e.g., 1 vs. 2) to be greater than the distance between larger numbers (e.g., 101 vs. 102, Banks, Fujii, Kayra-Stuart, 1976; Holyoak, 1978). On the basis of this finding, it could be argued that proportions derived from smaller populations (e.g., 1/20) may be perceived as being smaller than equivalent proportions derived from larger populations (e.g., 10/200). If this were true, subjects might understandably be more suspicious of a low-probability event stemming from a small population than of one stemming from a large population. The former event would, in effect, seem less probable. Although possible, this account does not seem plausible to us. In our research, subjects are not required to evaluate equivalent distances at lower versus higher values on the numerical scale. Furthermore, even if they did evaluate distances to assess their degree of suspicion (for example, by assessing the distance between the numerator and the denominator in the proportion), it is not clear that this would lead them to evaluate a proportion derived from a small population as being smaller than an equivalent proportion derived from a larger one. Indeed, it seems likely that the distance between 1 and 20 (the Study 1 proportion) is psychologically smaller, not larger, than that between 10 and 200 (after all, there are 19 units separating the first pair of numbers, but 190 units separating the second pair of numbers), in which case the latter proportion might be perceived to be smaller, not larger. Finally, the data from the control groups argue against the possibility that subjects perceived proportions based on smaller populations to be smaller than those derived from larger ones. If psychophysical effects were operative, we should have obtained condition differences on the measure that assessed subjects' predictions concerning the likelihood that the low-probability event would occur. No such effects were obtained.

An additional alternative account of our findings is derived from research demonstrating that people sometimes confuse frequency and proportion when estimating probability (Estes, 1976; Hamilton & Gifford, 1976; Silka & Albright, 1983). Specifically, it could be argued that when the absolute number of rare events was high, subjects may have evaluated the probability (at least the subjective probability) of obtaining a rare outcome to be greater than they did when the absolute number was low. As we have shown, however, the data from our control groups in Studies 1–4 do not support this account. Subjects' ratings of the likelihood that the rare event would occur did not differ across the experimental conditions in these studies.

One final account linked to number concerns the possibility that subjects may overextend the sample-size heuristic (Nisbett, Krantz, Jepson, & Kunda, 1983) to cases such as the present scenarios. Thus, even though our scenarios focused on populations rather than on samples, subjects may have ignored this fact and felt more confident in probabilistic hypotheses based on large numbers than in those based on small numbers. We know of no evidence to indicate that people actually do confuse samples and populations, but we cannot rule this possibility out. However, even if subjects did make this error, it is not clear that it constitutes a parsimonious account for our findings. Greater confidence in the accuracy of estimates pertaining to large as compared with small groups may plausibly explain the differences found on the measures used in Studies 4 and 5, but not those differences found on the measures used in Studies 1–3. For example, even if subjects in Study 3 thought that the judgments of the owners of a small and large car rental company differed in statistical reliability because of their inappropriate application of the sample-size heuristic, this would not seem to justify them thinking that one owner was more likely than the other to have *lied* to them.

Implications

The present results extend the empirical base of norm theory (Kahneman & Miller, 1986) in a number of respects. First, the focus of this research has been the ease of mentally replicating the target event rather than the ease of mentally undoing the target event. Previous investigations of mental simulation have manipulated factors designed to affect the ease with which an

alternative to a target event could be simulated; the present studies manipulated a factor that was designed to affect the ease with which the target event itself could be simulated. The easier it was for subjects to do this, the less strongly they reacted to the event.

Second, the events that these studies contrasted have not only equal probabilities, but precisely calculable probabilities. Previous research on norm theory (e.g., Kahneman & Miller, 1986; Kahneman & Tversky, 1982b) has contrasted events that were assumed not to vary in probability, but for which no precise probability estimates were calculable (e.g., people who died in a plane crash after having either changed or not changed flights at the last minute). Third, the present findings extend previous findings by revealing conditions under which postcomputational thought leads people to be unfair and irrational. Suspicion that a person had engaged in an illegitimate practice, or had misrepresented his or her experience, was based on the number of similar ways the event could have occurred, not solely on the probability of the event. In other words, subjects appeared to rely on the ease with which they could imagine or mentally simulate an event happening fairly in deciding the likelihood that it did happen fairly. Admittedly, there may be no normative standard for determining the appropriate level of suspicion that the selection of a chocolate chip cookie selection should generate (Study 1), but it does seem unfair to base suspicion on the absolute number of cookies in the jar. In fact, in a sample of more than 100 subjects to whom we presented the cookie-jar scenario, less than 5% thought it was "fair" to consider the absolute number of cookies in the jar when deciding whether or not the child had peeked.

In Studies 4 and 5, subjects were guilty of irrationality as well as unfairness. In these studies, subjects' willingness to accept hypotheses pertaining to the commonness of a particular event depended on the alleged absolute frequency of that event in the population. From the perspective of normative standards for judgments under uncertainty, subjects were no more justified in rejecting the low-incidence hypothesis in favor of the high-incidence hypothesis when the event was one of few possible events than when it was one of many.

The range of dependent measures used constitutes the fourth means by which the present results broaden the empirical base of norm theory. Previous research has linked normality to causal judgments (Wells & Gavanski, 1988; Wells, Taylor, & Turtle, 1987), perceptions of control (Turnbull, 1981), and luck (Johnson, 1986), as well as to counterfactual emotions such as regret (Kahneman & Tversky, 1982a) and sympathy (Miller & McFarland, 1986). The present research links normality to suspicion and belief revision.

Perhaps the most obvious social psychological implication of the present findings is the proposition that people will have more difficulty convincing others that an event that confirms a suspicion is just a coincidence in some circumstances than in others, even when the probability of the events is the same. This fact could influence, among other things, the way in which the person associated with the coincidence responds to the audience. One might, for example, expect especially defensive posturing from the child who draws the last chocolate chip cookie or the employer who finds that the employee who achieved the highest grade from him was the only male employee. Similarly, the manager of the car rental company might be especially vehement in protesting his honesty when a customer returns one of only two allegedly unreliable cars.

Our findings also have implications for in-group–out-group relations, particularly as they relate to group size. From Studies 4 and 5, it appears that an unexpected experience with a member of a small out-group will more likely lead people to revise their beliefs than will an equally unexpected experience with a member of a large out-group. People's willingness to change their beliefs in the face of evidence depends not only on how probable that evidence appears to be in light of their belief, but also on how normal it appears to be. Subjects in Study 4, for example, found it hard to believe that they "just coincidentally" encountered one of two reckless drivers in the taxi company, and thus assumed that there must be more than two. The more numerous the members of an out-group, the easier it is to believe that the "rotten apple" one has encountered is not representative of the rest of the barrel.

Future Directions

We have argued that people judge the plausibility of an improbable event's being a coincidence by the ease with which they can imagine or mentally simulate the occurrence of that event. Although the data are consistent with this argument, we provided no direct evidence from subjects to indicate that they actually engaged in this process. We did not seek this evidence for two reasons: (a) We doubted that subjects would have access to the process that mediated their judgments (Nisbett & Wilson, 1977; Wilson, 1985), and (b) even if they did have such access, we doubted that they would feel comfortable acknowledging it because we had evidence that they considered it inappropriate to be influenced in this manner. Still, future research should attempt to secure more direct, or at least more convergent, evidence for the process we have postulated. In doing so, it would be desirable to vary the ease of simulating the low-probability outcome by some means other than the number of ways in which it could have occurred. Varying the number of critical steps or causal joints in the scenario would be one such alternative. For instance, it should be more difficult to simulate the occurrence of a complex scenario than a simple one, irrespective of the a priori probability of the scenario.

References

Banks, W. P., Fujii, M., & Kayra-Stuart, F. (1976). Semantic congruity effects in comparative judgments of magnitudes of digits. *Journal of Experimental Psychology: Human Perception and Performance, 2,* 435–447.

Estes, W. K. (1976). The cognitive side of probability learning. *Psychological Review, 83,* 37–64.

Hamilton, D. L., & Gifford, R. K. (1976). Illusory correlation in interpersonal perceptions: A cognitive basis of stereotypic judgments. *Journal of Experimental Social Psychology, 12,* 392–407.

Hart, H. L. A., & Honoré, A. M. (1959). *Causation in the law.* London: Oxford University Press.

Heider, F. (1958). *The psychology of interpersonal relations.* New York: Wiley.

Holyoak, K. J. (1978). Comparative judgments with numerical reference points, *Cognitive Psychology, 10,* 203–243.

Johnson, J. T. (1986). The knowledge of what might have been: Affective and attributional consequences of near outcomes. *Personality and Social Psychology Bulletin, 12,* 51–62.

Kahneman, D., & Miller, D. T. (1986). Norm theory: Comparing reality to its alternatives. *Psychological Review, 93,* 136–153.

Kahneman, D., & Tversky, A. (1982a). The psychology of preferences. *Scientific American, 246,* 160–173.

Kahneman, D., & Tversky, A. (1982b). The simulation heuristic. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 201–208). New York: Cambridge University Press.

Landman, J. (1988). Regret and elation following action and inaction. *Personality and Social Psychology Bulletin, 13,* 524–536.

Miller, D. T., & McFarland, C. (1986). Counterfactual thinking and victim compensation: A test of norm theory. *Personality and Social Psychology Bulletin, 12,* 513–519.

Miller, D. T., Turnbull, W., & McFarland, C. (in press). Counterfactual thinking and social perception: Thinking about what might have been. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 23). Orlando, FL: Academic Press.

Nisbett, R. E., Krantz, D. H., Jepson, C., & Kunda, Z. (1983). The use of statistical heuristics in everyday inductive reasoning. *Psychological Review, 90,* 339–363.

Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review, 84,* 231–259.

Silka, L., & Albright, L. (1983). Intuitive judgments of rate change: The case of teenage pregnancies. *Basic and Applied Social Psychology, 4,* 337–352.

Turnbull, W. (1981). Naive conceptions of free will and the deterministic paradox. *Canadian Journal of Behavioural Science, 13,* 1–13.

Wells, G. L., & Gavanski, I. (1988). Mental simulation of causality. *Journal of Personality and Social Psychology, 56,* 161–169.

Wells, G. L., Taylor, B. R., & Turtle, J. W. (1987). The undoing of scenarios. *Journal of Personality and Social Psychology, 53,* 421–430.

Wilson, T. D. (1985). Strangers to ourselves: The origins and accuracy of beliefs about one's own mental states. In J. H. Harvey & G. Weary (Eds.), *Attribution: Basic issues and applications* (pp. 9–36) New York: Academic Press.

Received April 27, 1988
Revision received April 14, 1989
Accepted April 20, 1989 ■

Beutler, Levin, Tesser, and Miller Appointed New Editors, 1991–1996

The Publications and Communications Board of the American Psychological Association announces the appointments of Larry E. Beutler, University of Arizona; Joel R. Levin, University of Wisconsin; Abraham Tesser, University of Georgia; and Norman Miller, University of Southern California, as editors of the *Journal of Consulting and Clinical Psychology,* the *Journal of Educational Psychology,* the Attitudes and Social Cognition section and the Interpersonal Relations and Group Processes section of the *Journal of Personality and Social Psychology,* respectively. As of January 1, 1990, manuscripts should be directed as follows:

- For *Consulting and Clinical* send manuscripts to Larry E. Beutler, *Journal of Consulting and Clinical Psychology,* Department of Psychiatry, University of Arizona, College of Medicine, Tucson, Arizona 85724.
- For *Educational* send manuscripts to Joel R. Levin, Department of Educational Psychology, University of Wisconsin, 1025 West Johnson Street, Madison, Wisconsin 53706.
- For *JPSP: Attitudes* send manuscripts to Abraham Tesser, Institute for Behavioral Research, University of Georgia, 548 Boyd Graduate Studies, D. W. Brooks Drive, Athens, Georgia 30602.
- For *JPSP: Interpersonal* send manuscripts to Norman Miller, Department of Psychology, Seeley G. Mudd Building, University of Southern California, University Park, Los Angeles, California 90089.

Manuscript submission patterns make the precise date of completion of 1990 volumes uncertain. Current editors will receive and consider manuscripts until December 1989. Should any 1990 volume be completed before that date, manuscripts will be redirected to the newly appointed editor-elect for consideration in the 1991 volume.