

The "Same Person" Heuristic: An Attributional Procedure Based on an Assumption About Person Similarity

Charles G. Lord and Daniel T. Gilbert
Princeton University

A personal anecdote, attribution theory, and evidence from personality and social psychology suggested that attributors would be more convinced by within-persons evidence than by statistically equivalent between-persons evidence. Experiment 1 supported this prediction and suggested that attributors may rely on a cognitive heuristic—be more convinced of a stimulus' effectiveness when a change in stimuli changes the behavior of the same rather than different persons. It also suggested that this problem-solving shortcut may be based on the assumption that the same person is more likely to behave similarly from one situation to the next than are different persons, that is, that persons are more similar to themselves behaviorally than they are to each other. In Experiments 2 and 3, between-persons evidence was more convincing when it concerned similar persons, and within-persons evidence was more convincing when it concerned consistent persons, even when the persons were similar or consistent on dimensions seemingly irrelevant to the evidence in question. These results suggest that the greater convincingness of within-persons relative to between-persons evidence may rely on a problem-solving shortcut that is based on an assumption about person similarity.

A graduate student and a professor were discussing evidence that California sunshine puts people in a good mood. "When I moved across the country," said the graduate student, "I noticed that the people who waited on me in stores and restaurants in California were more pleasant than those on the East Coast." "I have a better story than that," countered the professor. "When the moving men picked up our furniture on the East Coast they were thoroughly obnoxious. We drove across the country to California and arrived to find the moving men sitting on the front lawn in the sunshine, eating fried chicken. They greeted us warmly and were quite genial and accommodating in delivering our furniture. These were the *same men*, mind you." The graduate student and the professor agreed that the latter evidence was the more compelling, and they speculated that this was because within-persons evidence usually generates greater confidence than between-persons evidence.

This hypothesis, that within-persons evidence is more convincing than between-persons evidence, may be derived from traditional models of causal attribution. Kelley (1967), for example, noted that the average person determines the cause of an event by reference to three dimensions: distinctiveness, consensus, and consistency. Suppose one knows that Fred feels happy when he is basking in the California sunshine. Does one attribute his good mood to something about Fred or to the effect of California sunshine? One is more likely to credit the California sunshine if Fred is not happy under other climatic conditions (distinctiveness), if other persons also feel happy in the California sunshine (consensus), and if Fred always feels happy in the California sunshine (consistency).

Table 1 presents the patterns of information necessary for drawing conclusions about distinctiveness, consensus, and consistency and does so in terms of whether an event involves the same or different persons, behaviors, stimuli, and modalities. As can be seen in Table 1, the professor's within-persons evidence was identical to high distinctiveness information in that the same persons were said to have responded differently to different stimuli, a pattern of information used frequently in ex-

We thank Daryl Bem, Ned Jones, Lee Ross, and two anonymous reviewers for helpful comments and suggestions on earlier versions of the manuscript.

Requests for reprints should be sent to Charles Lord, Department of Psychology, Princeton University, Princeton, New Jersey 08544.

perimentation as proof of a stimulus' or manipulation's effectiveness. Notice, however, that this information by itself conveys neither high nor low consensus or consistency in that it matches none of those four patterns. The graduate student's between-persons evidence, in which different persons were said to have responded differently to different stimuli, also represents a pattern of information used frequently in experimentation but by itself conveys neither high nor low distinctiveness, consensus, or consistency and thus should be less convincing than within-persons evidence as proof of a stimulus' or manipulation's effectiveness (cf. McArthur, 1972).

The attributor who lacks information about one or more of Kelley's (1967) dimensions is apt to rely on expectations, schemata (Kelley, 1971), and heuristics. A heuristic is a problem-solving procedure or shortcut that is based on an unjustified or unjustifiable assumption and thus may lead either to the goal or down a blind alley. We contend that within-persons evidence generates greater intuitive confidence than between-persons evidence because attributors rely on a heuristic assumption about person similarity. The labels "same" and "different" in Table 1 may be considered as the two endpoints of a "similar-dissimilar" continuum. Thus "same persons respond differently to different stimuli" (within-persons evidence) may be more convincing as proof of one stimulus' greater effectiveness than "different persons respond differently to different stimuli" (between-persons evidence) because the attributor expects similar persons to behave

similarly and dissimilar persons to behave dissimilarly. That is, the attributor relies on a problem-solving procedure or guideline—is more convinced of a stimulus' effectiveness when a change in stimuli changes the behavior of the same rather than different persons—that is based on an assumption that the same person is more likely to behave similarly from one situation to the next than are different persons.

This assumption of greater intrapersonal than interpersonal similarity is so basic that it may affect causal attributions without the attributor being aware of its influence or questioning its validity. It is an assumption that seems to require no justification, yet it is an assumption for which empirical justification is somewhat lacking. Mischel (1968), in reviewing the accumulated evidence on this issue, concluded that the assumption of cross-situational similarity within the behaviors of an individual has received less than overwhelming empirical support. In fact, attributors may mistakenly infer cross-situational similarity from cross-temporal similarity (Mischel & Peake, 1982), a problem-solving shortcut analogous to inferring an event's probability from the ease with which it comes to mind, as in the "availability heuristic" (Tversky & Kahneman, 1973). Although others have argued that cross-situational similarity, when examined idiographically rather than nomothetically, is more a fact than an assumption (Bem, 1983; Bem & Allen, 1974; Lord, 1982), both sides agree that attributors expect more cross-situational similarity within

Table 1
Patterns of Information in Distinctiveness, Consensus, Consistency, Within-Persons Evidence, and Between-Persons Evidence

Information	Person	Behavior or response	Entity or stimulus	Time or modality
Distinctiveness				
High	Same	Different	Different	Same
Low	Same	Same	Different	Same
Consensus				
High	Different	Same	Same	Same
Low	Different	Different	Same	Same
Consistency				
High	Same	Same	Same	Different
Low	Same	Different	Same	Different
Within-persons evidence	Same	Different	Different	Same
Between-persons evidence	Different	Different	Different	Same

the behaviors of the same person than objectively exists. In addition, many studies have shown that attributors underestimate the degree to which different persons behave similarly in response to compelling social stimuli (Asch, 1951; Bierbrauer, 1973; Milgram, 1963; Nisbett, Borgida, Crandall, & Reed, 1982). Inflated expectations for intrapersonal similarity may combine with deflated expectations for interpersonal similarity to influence causal attributions.

Experiment 1 was designed to test whether these speculations address a real phenomenon, that is, whether within-persons evidence does seem more convincing to the attributor than between-persons evidence, even when the two kinds of evidence are equated for statistical significance.

Experiment 1

Method

Overview. Sixteen Princeton University undergraduate students read summaries that displayed the data from two fictitious studies. One study employed a within-persons format; the other employed a between-persons format with twice as many subjects. In terms of raw scores, the between-persons study produced larger differences. In terms of statistical significance, the two studies produced identical confidence levels. The students were asked how much confidence the researcher was entitled to hold in each result, which study they personally found more convincing, and how likely it was that each result would affect their own behavior. They were also asked to estimate how much each result depended on the effect of the experimental manipulation, ability, luck, and effort.

Materials. Between-persons and within-persons versions were prepared for each of two fictitious studies, one comparing the effects of two different drugs on cutaneous sensitivity and the other comparing the effects of two kinds of background music on keypunching accuracy. For example, the between-persons version of the drug study stated that

researchers who are interested in helping blind persons read Braille faster have studied two drugs that increase skin sensitivity. ANCYCLOMYCRIN and PHYLOPOROZINE each increase skin sensitivity without any known side effects, and both can be given in controlled doses that wear off completely within an hour. However, the manufacturers of these drugs both claim that their respective products are superior in producing skin sensitivity. A researcher at a well known school of medicine recently taught 8 sighted persons to read Braille, then randomly assigned 4 of them to receive ANCYCLOMYCRIN and the other 4 to receive PHYLOPOROZINE. After the drugs had taken effect, each participant was presented with a Braille character for 2000 milliseconds. The duration of these "touches" was precisely controlled by a machine called a cutachistoscope, which can present a character to a

person's fingertips for an exact, specified length of time. The researcher measured the number of Braille characters that volunteers correctly identified out of 175. The results were as follows:

Subject	Sex	Age	# Characters Correct With
			Ancyclomycin
James W.	M	28	104
Roger J.	M	34	104
Doris G.	F	30	64
Cathy F.	F	32	128
			Phyloporozine
Scott H.	M	26	128
Charles D.	M	36	160
Joanne C.	F	32	120
Elizabeth R.	F	30	152

The within-persons version was identical except that the researcher was said to have taught only four sighted persons to read braille. All four took both drugs 3 hours apart, with two persons randomly selected to take Ancyclomycin first and the other two selected to take Phyloporozine first. Thus, the data table listed only four names (those in the top four rows above), each with two scores. Cathy F's Phyloporozine score was changed to 144 (a smaller raw score difference) in the within-persons version so that a *t* test in either version would yield the same probability level of $p = .05$.

The between-persons version of the music study explained that although background music was known to decrease keypunching errors, an important question was whether the "step-wave pattern of the flute" or the "sine-wave pattern of the violin" was more helpful. Therefore, researchers had piped flute music into one keypunching room and violin music into a second, with the following error rates per thousand characters:

Subject	Sex	Age	# Errors With
			Violin Music
Rodney S.	M	26	32
Dale L.	M	32	40
Sandy P.	F	28	30
Rita M.	F	30	38
			Flute Music
Thomas N.	M	24	26
Jeff K.	M	34	26
Wendy D.	F	30	16
Nancy R.	F	28	32

The within-persons version was identical except that the researchers piped flute and violin music into the same room at two different times and measured error rates for only four workers. The data were identical except that Rita M.'s error rate with flute music was changed to 36. Because these data are linear transforms of those for the drug study, the probability levels were identical for both studies in both versions.

Procedure. The students were told that "psychologists are often asked to evaluate research in a variety of disciplines for review committees whose members may have little training in behavioral analysis. Therefore, it is important to us to know just how convincing the research

Table 2

Comparisons of Within-Persons and Between-Persons Evidence: Experiment 1

Question	Within-persons evidence	Between-persons evidence	<i>p</i> <
How confident should the researcher be? (in percentage)	61.9	40.6	.01
Which study was more convincing? (proportion choosing each) ^a	81.2	6.2	.01
How likely would you be to choose the "winning" drug/music for yourself? (10 = extremely likely)	8.63	7.26	.05
How much was James W.'s (Rodney S.'s) behavior due to (10 = very much)			
Ability	6.88	7.50	.06
Luck	3.75	4.31	<i>ns</i>
Effort	6.00	6.44	<i>ns</i>
The drug/music	6.44	5.25	.06

^a Two of 16 students marked the midpoint (i.e., preferred neither study).

we analyze is to the intelligent non-psychologist." All students then received booklets that described both a drug study and a music study. One was the within-persons version and the other the between-persons version. For half of the students the first study was within persons and the second study was between persons. For the other half, this order was reversed. The drug study was always presented first. Which drug (music) proved more effective was also counterbalanced.

Dependent measures. After reading both studies, the students answered the following questions: First, they indicated what percentage of confidence the researcher(s) should place in each study's results. Next, they indicated which study was more convincing on an 11-point scale labeled "Drug Study More Convincing" on one end, "Studies Equally Convincing" in the middle, and "Music Study More Convincing" on the other end. Finally, they were asked the extent to which they would base their own behavior on each study's evidence. They were asked how likely they would be to use Phylporozine (or Ancylo-mycin) if they were to become blind and how likely they would be to choose to listen to flute (or violin) music while working if they were to become keypunch operators. On a second page, the students were asked for each study to attribute a specific subject's performance to "luck or chance," "his or her effort," "his or her ability," and "the effect of the drug (music)," each on a separate 11-point scale ranging from 0 ("little or not at all") to 10 ("very much or completely"). After completing all dependent measures, the students were paid, debriefed, and asked not to discuss the experiment with other students.

Results and Discussion

The primary hypothesis was that within-persons evidence is more convincing than statistically equivalent between-persons evidence. This hypothesis was tested by answers to three questions. First, would our student subjects believe that investigators who produce within-persons experimental evidence are entitled to hold greater confidence in their results? Sec-

ond, when asked directly to compare the convincingness of two studies that produced statistically equivalent results, would the students find the within-persons study more convincing? Third, would the students be more likely to base their own behaviors on within-persons evidence?

Confidence. As shown in Table 2, students who read both studies felt that researchers who produced within-persons evidence were entitled to place more confidence in their results (61.9%) than researchers who produced between-persons evidence (40.6%). This main effect of type of evidence was significant according to a 2 (type of evidence: within, between) \times 2 (pairing: drug within and music between, drug between and music within) \times 2 (more effective drug: Ancylo-mycin, Phylporozine) \times 2 (more effective music: flute, violin) analysis of variance (ANOVA), $F(1, 8) = 12.84, p < .01$, and was independent of study content.¹

Convincingness. Ratings on the bipolar convincingness scale were converted to a single score that expressed the degree of preference for the within-persons evidence. Thus, convincingness scores could range from -5 (between-persons study much more convincing)

¹ The only other significant effect was an uninterpretable Type of Evidence \times Pairing \times More Effective Music interaction, $F(1, 8) = 5.38, p < .05$. The difference between types of evidence was most pronounced (67% within persons, 17% between persons) when flute music produced fewer errors in a between-persons design.

to +5 (within-persons study much more convincing). These scores were significantly slanted in favor of the within-persons evidence, $t(15) = 4.62, p < .001$. Thirteen of 16 students marked the convincingness scale on the within-persons side, 2 chose the middle, and only 1 found the between-persons evidence more convincing ($\chi^2 = 13.98, p < .001$).

Behavior. Perhaps the most important measure was the question about the students' own behaviors. It was quite possible that as an abstract intellectual exercise, the students would find the within-persons evidence more convincing, but would the evidence prompt them to choose different behavioral alternatives? A Type of Evidence \times Pairing \times More Effective Drug \times More Effective Music ANOVA of the behavioral question revealed only a significant main effect of type of evidence, $F(1, 8) = 8.07, p < .05$. As shown in Table 2, the students were more willing to translate their convictions into behavioral intentions when the convictions were based on within-persons rather than between-persons evidence.

Causal attributions. A similar ANOVA was performed on each of the four attributional scales: effect of the drug (or music), luck, ability, and effort. Table 2 shows that the experimental manipulation, whether drug or music, was believed to have influenced the within-persons outcome marginally more than the between-persons outcome, $F(1, 8) = 5.23, p < .06$. In addition, all three other explanations of the results—luck, effort, and ability—were thought to have been more influential in producing the between-persons than the within-persons evidence, although this effect reached marginal significance only for ability, $F(1, 8) = 5.00, p < .06$. A significant positive correlation was found between the strength of a student's preference for the within-persons study and the extent to which that student believed the experimental manipulation to have been more responsible for the within-persons than for the between-persons result, $r(15) = .50, p < .05$. In other words, those students who found the within-persons evidence more convincing also attributed the within-persons result more to the drug (or music).

Summary. Why did our subjects find the within-persons evidence more convincing than the between-persons evidence? By standard statistical criteria no such difference should

have been found, for we had equated the statistical confidence levels of the two studies. In addition, we emphasized random assignment in both studies because it has been shown that attributors take random assignment into account in making causal attributions (Wells & Harvey, 1977). Moreover, the greater mean differences between the raw scores, and the greater number of subjects in the between-persons study should have produced an appearance of stronger, more compelling evidence than that found in the within-persons study.²

In the spirit of Tversky and Kahneman (1974), we should like to suggest that our subjects relied on an effort-reducing heuristic—were more convinced of a stimulus' effectiveness when a change in stimuli changed the behavior of the same rather than different persons—and that this heuristic is based on the assumption that the same person is more likely to behave similarly from one situation to the next than are different persons. If our reasoning is correct, that is, if the greater convincingness of within-persons evidence rests on an assumption about person similarity, then manipulations of perceived person similarity should be effective in increasing or decreasing the convincingness of either form of evidence. More specifically, between-persons evidence should be more convincing when it involves persons who are more rather than less similar to one another, and because consistency is a form of similarity to oneself, within-persons evidence should be more convincing when it involves consistent rather than inconsistent persons. Experiments 2 and 3 tested these hypotheses.

Experiment 2

This experiment was designed to determine whether between-persons evidence is more convincing when it involves similar rather than dissimilar persons.

² An alternative explanation is that subjects may have viewed the between-persons data as two overlapping distributions (in the drug example, Cathy F's score is higher than Joanne C's) and the within-persons data as one distribution of uniformly positive change scores. Interestingly, the classification of the within-persons data as "change scores" may itself reflect the operation of the "same person" heuristic.

Method

Overview. Twenty-four Princeton University undergraduate students read summaries that displayed the data from two fictitious between-persons "pretests." Although the pretests had yielded identical results, one had used as subjects persons who appeared similar on an irrelevant dimension, whereas the other had used as subjects persons who appeared dissimilar on a different, also irrelevant, dimension. The students were asked to make situational and dispositional attributions for the two outcomes, to determine which pretest entitled the researcher to hold greater confidence in the outcome, and to estimate the extent to which their own behavior would be affected by knowledge of each outcome.

Materials. Two pretest descriptions were prepared. Each described a pretest in which—to determine whether large mood shifts were a side effect of the drug Trianil—two persons were told that they were taking a placebo. One was actually given Trianil and the other the placebo. Both persons recorded their moods each hour for 24 hours. The person who had ingested Trianil experienced "many large mood shifts" and the person who had ingested the placebo experienced "a few small mood shifts" over the 24-hour period. Also, in each pretest the two persons were said to have been selected at random from a list of individuals whose names were available because they happened to have participated in an unrelated prior experiment. Two such "prior experiment" lists were used. One was from an experiment on "how accurately subjects could identify true north from a strange location," and the other was from an experiment on "how much subjects admired their parents in important matters."³ The pretest persons from the "true north" experiment were said to have erred by an average of 78° and 72°, respectively. The pretest persons from the "parental admiration" experiment were said to have admiration scores of 88 and 80, respectively. In the similar-persons study the other scores on the prior-experiment list ranged from 9° to 174° or 2 to 196 admiration points; in the dissimilar-persons study the other scores on the prior-experiment list ranged from 72° to 79° or 80 to 89 admiration points. We chose this particular way of operationalizing perceived similarity to avoid confounds that would inevitably have occurred had the two pretest subjects differed across the similar-persons and dissimilar-persons versions of each pretest (e.g., two men versus one man and one woman). We solved this problem by giving the "subjects" the same names and scores in the two versions and changing only the scores of the other persons who happened to have participated with them in the prior experiment.

Procedure. The students were told that we often conduct "pretest samples" before launching extensive formal research programs and that we wanted their opinions about two such two-person samples to "help us decide whether it is useful or not to finance larger studies." All students received booklets that described two pretests that, as described above, differed only in the "unrelated prior-experiment list" from which the two pretest persons had been randomly selected. For half of the students the similar persons were chosen from the true-north list and the dissimilar persons were chosen from the parental-admiration list; for the other half, these pairings were reversed. For each group, half of the subjects learned that the high scorer (78° or 88 admiration points) had received the drug, and

the other half learned that the low scorer (72° or 80 admiration points) had received the drug. The overall design crossed two pairings (north-similar/admiration-dissimilar, north-dissimilar/admiration-similar) with two levels of drug taker (high scorer, low scorer) and two levels of person similarity (similar persons, dissimilar persons), with repeated measures on the third factor.

Dependent measures. After reading both pretests, in each of which the person who took Trianil had larger mood shifts, our student subjects were asked—both on 9-point scales and as percentages constrained to add to 100%—how much they attributed each pretest's results to individual differences and how much they attributed each pretest's results to a side effect of the drug. They were also asked to indicate on a 9-point scale how likely it was that "If my own physician prescribed Trianil for some ailment, I would speak up and request a different drug." The students were then asked to compare the two pretests directly on how confident the researcher should be that a side effect of the drug Trianil had caused the mood swings. This rating was on a 9-point bipolar scale, with one of the two pretests at each end. Finally, as a manipulation check, the students were asked, again on 9-point scales, to rate the "overall general similarity" of the two persons involved in each pretest. Following this, the experimenter explained that we had invented the data and the drug's name, and he conducted a thorough debriefing.

Results and Discussion

We were interested in whether confidence in between-persons evidence varies as a function of the perceived similarity of the persons involved. Thus, we attempted to present persons who would be perceived as either similar or dissimilar to one another. Analysis of the manipulation-check measure indicated that we were successful, $F(1, 20) = 9.55, p < .01$. In fact, only 3 of 24 subjects perceived the dissimilar persons as more similar. As will be discussed below, these three subjects may be regarded as the "exceptions that prove the rule."

Attributions. As shown in Table 3, the similar-persons pretest prompted greater attribution to the drug (6.38) than did the dissimilar-persons pretest (5.75), producing the predicted person similarity main effect in a Pairing \times Drug Taker \times Person Similarity ANOVA, $F(1, 20) = 5.23, p < .05$. In this, as in the other analyses to be reported, pairing and drug taker showed neither significant main effects nor significant interactions with person similarity, an indication that perceived simi-

³ These dimensions were selected because 8 of 10 pretest subjects believed that they had no predictive validity for mood stability.

Table 3

Comparisons of Similar and Dissimilar Persons: Between-Persons Evidence in Experiment 2

Question	Similar persons	Dissimilar persons	<i>p</i> <
To what extent were the results due to Trianil? (9 = very much)	6.38	5.75	.05
To what extent were the results due to personal disposition? (9 = very much)	4.92	5.42	.10
What percentage of the results was due to Trianil?	56.4	51.3	.10
In which pretest's results should the researcher be more confident? (proportion choosing each) ^a	66.7	12.5	.01
How likely would you be to speak up to avoid taking Trianil if your physician prescribed it? (9 = very likely)	6.50	6.08	.01

^a Five of 24 students marked the midpoint (i.e., preferred neither pretest).

larity is sufficient in itself, even on irrelevant dimensions, to alter reactions to between-persons evidence. When attributions to individual differences were subjected to the same analysis, the effect of person similarity was marginal, $F(1, 20) = 3.05$, $p < .10$, but the means were in the predicted direction (similar persons, $M = 4.92$; dissimilar persons, $M = 5.42$). As will be recalled, however, three subjects perceived the "dissimilar" persons as more similar to one another than the "similar" persons. It could be argued that the hypothesis of a greater attribution to individual differences for dissimilar persons would be better tested by changing the experimenter's labels for the pretest persons to the students' own labels. The same analysis performed on these converted data yielded a significant effect of person similarity, $F(1, 20) = 5.61$, $p < .05$. In fact, although not employed elsewhere, a switch of these three subjects' ratings to accord with their own perceptions of person similarity would have increased, and not once decreased, the significance of the statistics reported. A Pairing \times Drug Taker \times Person Similarity ANOVA of the percentage of attribution ratings failed to achieve person-similarity significance, $F(1, 20) = 3.20$, $p < .10$, although the means were again ordered in the predicted direction. Large mood shifts were attributed to the drug at a rate of 56.4% (and to individual differences at 43.6%) based on pretest scores from similar persons but attributed to the drug at a rate of 51.3% (and to individual differences at 48.7%) based on pretest scores from dissimilar persons.

Confidence. When asked to compare the two pretests directly in terms of how confident the researcher should be, the students displayed

an overwhelming preference for the similar-persons evidence, $F(1, 20) = 11.50$, $p < .01$. In fact, only three students marked the scale on the side favoring the dissimilar-persons pretest, and these were the same three students who perceived the dissimilar persons as more similar to one another.

Behavior. Most importantly, the students expressed a greater willingness to translate their convictions into behavior when those convictions were based on between-persons evidence obtained from similar persons, $F(1, 20) = 4.90$, $p < .05$, by indicating that they would be more likely to speak up and request another drug should their own doctor prescribe Trianil.

Summary. The results of postexperimental interviews of Experiment 1 led us to believe that within-persons evidence is more convincing than between-persons evidence because of the same-person heuristic. Perceivers expect that the same person will behave more similarly from one situation to the next than will different persons and are therefore more impressed when a stimulus or manipulation changes the behavior of the same rather than different persons. We believed that the result depended on a basic and unquestioned assumption about person similarity rather than on complex inferential reasoning or a failure to understand random assignment. This led us to predict that any stimulus would appear to be a more plausible cause when it changed the behavior of persons who were perceived as similar to one another than when it changed the behavior of persons who were perceived as dissimilar to one another. Experiment 2 supported this hypothesis. Our subjects were more likely to make an attribution to Trianil,

felt that scientific researchers were entitled to hold greater confidence in their results, and rated their own behavior as more likely to be affected when between-persons evidence was obtained from persons who were perceived as similar rather than dissimilar to one another. An analogous prediction is that a stimulus will appear to be a more plausible cause when it changes the behavior of persons who are perceived as consistent than when it changes the behavior of persons who are perceived as inconsistent.

Experiment 3

Method

In an attempt to model as closely as possible the California sunshine anecdote that generated this line of research, we asked 20 undergraduate students to read "pretest examples" that concerned a new type of lighting. We told them that

psychologists have long observed that natural sunlight seems to put people in a good mood. Industrial manufacturers have attempted to create a lighting source that captures the "psychological feel" of natural sunlight, and claim that these bulbs do indeed induce a better overall mood than normal indoor lighting. (You may have seen these marketed as "sun bulbs" or Natursun bulbs.) Our pretest subject came into the laboratory for a one-hour session on two consecutive days, and was given a simple proofreading task to perform. The room contained a normal lightbulb on the first day, which was replaced with a "sun bulb" on the second day. After each session the subject completed several scales and questionnaires, including a mood index. (He did not know what the experiment was about, nor did he know that the lightbulbs had been switched.)

The normal-bulb rating was always "slightly bad mood" and the sun-bulb rating was always "extremely good mood." Analogous to Experiment 2, the fictitious subject in one pretest was made to appear consistent on one unrelated dimension, and the fictitious subject in the other pretest was made to appear inconsistent. For example, Dick H. was said to have been in an unrelated previous experiment in which he tried to detect the letter *k* in a rapid visual display. He was always said to have detected 164, 186, 174, 160, and 176 *ks* on 5 consecutive days of testing. These data were depicted both in a table and graphically on what looked like a handwritten notebook page torn from an experimenter's records. Dick H. was made to look consistent by showing on the same notebook page a second subject, John C., who had identified 202, 140, 217, 104, and 197 *ks* on the same 5 days. In the inconsistent version John C.'s scores were 173, 172, 171, 171, and 173. In both versions subjects were told that "John C. did not participate in the sun bulb experiment and his data can therefore be ignored." The second unrelated previous experiment in-

volved the number of nonsense syllables recalled over 10 trials per day for 5 days.⁴ The recall scores presented for the nonsense-syllable experiment were linear transforms of the scores for the visual-detection experiment.

All subjects read about a pretest involving a consistent subject and a pretest involving an inconsistent subject. Dependent measures were much the same as in Experiment 2, except that the word *consistent* was substituted for *similar* in the manipulation check, which was taken last, just before debriefing. The complete design involved two pairings (*ks* consistent, nonsense syllables inconsistent and vice versa), two orders (*ks* first, nonsense syllables first), and two levels of consistency (consistent, inconsistent), with repeated measures on the last factor.

Results and Discussion

We attempted to present persons who would be perceived as either consistent or inconsistent. Analysis of the manipulation-check measure indicated that we were successful, $F(1, 16) = 22.47, p < .001$. No effects of pairing or order appeared in any of the analyses to be reported.

Attributions and convincingness. As shown in Table 4, our student subjects attributed the pretest person's better mood more to the sun bulb when he had been consistent (relative to a second person whose data they were told to ignore) in a previous experiment on what we assumed was an unrelated topic, $F(1, 16) = 6.02, p < .05$. They also attributed the pretest person's better mood less to his personality when he had been consistent than when he had been inconsistent, $F(1, 16) = 5.61, p < .05$. The students also attributed a greater percentage of cause to the effect of the sun bulb when it improved the mood of a consistent person, $F(1, 16) = 8.93, p < .01$. In addition, a significant majority found the pretest involving the consistent person more convincing ($\chi^2 = 11.34, p < .01$).

Behavior. Finally, the students indicated that they would be more likely to install sun bulbs in their own rooms based on the pretest

⁴ We did not consider it necessary to pretest to determine whether students consider the number of *ks* detected or number of nonsense syllables recalled to be relevant to the effect of light bulbs on mood shifts. It could be argued, however, that subjects believed that consistent performance at any task implies consistent moods. Unfortunately, this potential confound exists in any method of operationalizing within-persons consistency.

Table 4
Comparisons of Consistent and Inconsistent Persons: Within-Persons Evidence in Experiment 3

Question	Consistent person	Inconsistent person	<i>p</i> <
To what extent was change in persons's mood due to sun bulb? (11 = very much)	7.40	6.40	.05
To what extent was change in person's mood due to disposition? (11 = very much)	6.05	7.40	.05
What percentage of change was due to sun bulb?	54.6	38.0	.01
In which pretest's results should the researcher be more confident? (proportion choosing each) ^a	70.0	5.0	.01
How likely would you be to install the sun bulb in your own room? (11 = very likely)	6.05	4.90	.05

^a Five of 20 students marked the midpoint (i.e., preferred neither pretest).

involving the consistent person, $F(1, 16) = 6.08$, $p < .05$.

Summary. As predicted, the students placed greater confidence in "scientific" conclusions generated by within-persons evidence when the persons involved were perceived as more consistent, or more similar, to themselves. This result was predicted by our reasoning that the same-person heuristic is an attributional problem-solving procedure that is based on a basic assumption about person similarity.

General Discussion

The results of Experiments 2 and 3 suggest that within-persons evidence may not always generate greater confidence than between-persons evidence. The important variable, it seems, is perceived person similarity. If enough time intervenes in the within-persons evidence so that the person is seen as very dissimilar to himself or herself, and if the persons involved in the between-persons evidence are extremely similar, then the normal pattern of confidence ratings may be reversed. Certainly, no one doubts that a difference between the behaviors of identical twins would be more convincing than a difference between the behaviors of a person at age 3 years and the same person at age 43 years. What we have labeled the same-person heuristic may thus be viewed as only one manifestation of a general assumption about person similarity. In most everyday contexts, however, we maintain that within-persons evidence is accorded greater

credence than it merits relative to between-persons evidence.⁵ In fact, although we have no data to support this claim, our debriefing sessions convinced us that far from engaging in complex statistical reasoning, our subjects paid little attention to the relative size of the numerical differences in each condition of Experiment 1 and instead relied on a cognitively efficient heuristic.

The potency of the same-person heuristic in everyday life is illustrated not only by advertising campaigns, which use predominantly within-persons evidence, but also by the news. Consider the well-publicized trial of an East Coast man accused of putting his wife in a coma with insulin injections. The defense and the prosecution each presented several witnesses who spoke for and against the defendant's character. The jury seemed little moved by this between-persons evidence. The testimony that may have decided the issue came in the form of within-persons evidence, when the defendant's former mistress testified that upon first hearing of the charges she was convinced of her lover's innocence. When asked for her present opinion, she replied "I'm not sure." One could argue that this testimony had simply added one point pro and one point

⁵ One might alternatively contend that between-persons evidence is associated with diminished confidence relative to within-persons evidence (cf. Edwards, 1968). Although we favor the increased confidence in the within-persons-evidence explanation, we can offer no proof that the "action" in Experiment 1 was not all on the between-persons side.

con to the long list of character opinions. The person she was at one time believed the man innocent; the person she was many months later had doubts. The jury did not see it that way. If one can believe news reports and analyses, the jury seemed more impressed by this one difference of opinion within what they regarded as the same person than by all the other differences of opinion.

The same-person heuristic may be involved in the fundamental attribution error (Ross, 1977), which consists of "the tendency to attribute behavior exclusively to the actor's dispositions and to ignore powerful situational determinants of the behavior" (Nisbett & Ross, 1980, p. 31). If we witness a student deliver a pro-Castro speech, we assume that the student has a pro-Castro attitude even when we know that he was told to do so (Jones & Harris, 1967). The usual explanations are perceptual, in terms of behavior "engulfing the field" (Heider, 1958), or cognitive, in terms of failure to adjust a first impression sufficiently even after learning that the behavior was constrained (Jones, 1979; Quattrone, 1982). We suggest that a question about an actor's attitude given his behavior in one situation may be a question about generalization of the behavior to other, in this case nonconstraining, situations. The attributor who relies on the same-person heuristic will conclude, with only a minimum of inferential reasoning, that the actor will behave in a similar manner in other situations. After all, he *is* the same person.

The same-person heuristic may also enter into the explanation of actor-observer differences (Jones & Nisbett, 1971). The phenomenon, which has been verified repeatedly, is that we attribute our own behaviors relatively more to the situation and less to personal causes than we do for the observed behavior of other persons. This discrepancy would make sense if we expect less consistency from ourselves than from others, and in fact we do (Nisbett, Caputo, Legant, & Maracek, 1973). Thus, we may recognize that we behave like somewhat different persons in different contexts but continue to believe on an intuitive level that other persons are always the same. This is not to say that we view ourselves as inconsistent or that we never apply the same-person heuristic to ourselves. Many experi-

ments guided by self-perception theory (Bem, 1972) have demonstrated that we readily draw confident conclusions about the cross-situational similarity of our own behaviors and attitudes. For a variety of reasons, however, we seem more prone to apply the heuristic to others. We may change, but others are expected to exhibit the consistency of the Rock of Gibraltar, and anything that can move the Rock must be a powerful force indeed.

In fact, the effects of within-persons evidence and between-persons evidence on causal attributions are roughly parallel to the effects of target-based and category-based expectancies (Jones & McGillis, 1976). The former are expectancies based on prior behavior of the same actor, whereas the latter are expectancies based on distinguishing characteristics of the actor's group. Also related are discussions of differences between temporal comparison, for example, running to beat your "personal best," and social comparison, for example, running to beat the field (Albert, 1977; Suls & Mullen, 1982). It is interesting in light of Experiments 2 and 3 that Goethals and Darley (1977) singled out perceived similarity as the most important determinant of the attributional impact of information relevant to social comparison.

When two behaviors are discrepant, the most likely causes are situational differences and personal dispositions. The more similar the two persons seem, the more confident we are that the difference between their behaviors was situationally caused. The reason is that our expectations and assumptions have been violated more severely. Suppose we learn that two discrepant behaviors have occurred: A person moved from the East Coast to California and a month later a person moved from California to the East Coast. Whether we make a dispositional or a situational attribution for the difference between these two behaviors is influenced by whether the same person or different persons were involved. If Fred moved to California and then back to the East Coast, a situational attribution, perhaps in terms of a lost job, seems necessary. If Fred moved to California and Ted to the East Coast, the discrepancy is either ignored or put down to different personalities. Notice, however, that these attributions—the former situational and the

latter dispositional—depend on assumptions about person similarity. When Fred is known to be consistent, or when Fred and Ted are known to be dissimilar, the attributions become even more situational and even more dispositional, respectively. When Fred is inconsistent, or when Fred and Ted are known to be similar, the tendency to make more situational attributions for the same person is weakened and may (conceivably) be reversed. This hypothesized reversal highlights our contention that in explaining why two behaviors differed, the question of whether the events involved the same or different persons may be important only because of a basic and usually unquestioned assumption about person similarity.

It appears that in addition to attributing human motives and dispositions to inanimate objects (Heider & Simmel, 1944), one may also attribute some of the characteristics of things to persons. One may expect others to remain steadfast and unswerving, varying little from day to day or situation to situation, and thus one may be overly impressed when some external agent alters the “invariant” course of their behavior and violates one’s expectations. At an intuitive level it indeed *feels* as though a situational explanation is required to explain a discrepancy between behaviors involving the same person, as though “sameness” were the rule and difference the deviation. Although this simple heuristic often serves one well, it is fascinating to speculate about why the human organism has adopted this rule of thumb rather than the equally plausible Heraclitean perspective that human beings are more like rivers than like rocks and that never do we meet the same person twice.

References

- Albert, S. Temporal comparison theory. *Psychological Review*, 1977, 84, 485–503.
- Asch, S. Effects of group pressure upon the modification and distortion of judgment. In H. Guetzkow (Ed.), *Groups, leadership, and men*. Pittsburgh, Pa.: Carnegie Press, 1951.
- Bem, D. J. Self-perception theory. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 6). New York: Academic Press, 1972.
- Bem, D. J. Toward a response style theory of persons in situations. In R. A. Dienstbier & M. M. Page (Eds.), *Nebraska Symposium on Motivation 1982: Personality—Current Theory and Research*. Lincoln: University of Nebraska Press, 1983.
- Bem, D. J., & Allen, A. On predicting some of the people some of the time: The search for cross-situational consistencies in behavior. *Psychological Review*, 1974, 81, 506–520.
- Bierbrauer, G. *Effects of set, perspective, and temporal factors in attribution*. Unpublished doctoral dissertation, Stanford University, 1973.
- Edwards, W. Conservatism in human information processing. In B. Kleinmuntz (Ed.), *Formal representation of human judgment*. New York: Wiley, 1968.
- Goethals, G. R., & Darley, J. M. Social comparison theory: An attributional approach. In J. M. Suls & R. L. Miller (Eds.), *Social comparison processes: Theoretical and empirical perspectives*. Washington, D.C.: Hemisphere, 1977.
- Heider, F. *The psychology of interpersonal relations*. New York: Wiley, 1958.
- Heider, F., & Simmel, M. An experimental study of apparent behavior. *American Journal of Psychology*, 1944, 57, 243–259.
- Jones, E. E. The rocky road from acts to dispositions. *American Psychologist*, 1979, 34, 107–117.
- Jones, E. E., & Harris, V. A. The attribution of attitudes. *Journal of Experimental Social Psychology*, 1967, 3, 1–24.
- Jones, E. E., & McGillis, D. Correspondent inferences and the attribution cube: A comparative reappraisal. In J. H. Harvey, W. J. Ickes, & R. F. Kidd (Eds.), *New directions in attribution research* (Vol. 1). Hillsdale, N.J.: Erlbaum, 1976.
- Jones, E. E., & Nisbett, R. E. *The actor and the observer: Divergent perceptions of the causes of behavior*. Morristown, N.J.: General Learning Press, 1971.
- Kelley, H. H. Attribution theory in social psychology. In D. Levine (Ed.), *Nebraska Symposium on Motivation* (Vol. 15). Lincoln: University of Nebraska Press, 1967.
- Kelley, H. H. *Causal schemata and the attribution process*. Morristown, N.J.: General Learning Press, 1971.
- Lord, C. G. Predicting behavioral consistency from an individual’s perception of situational similarities. *Journal of Personality and Social Psychology*, 1982, 42, 1076–1088.
- McArthur, L. A. The how and what of why: Some determinants and consequences of causal attribution. *Journal of Personality and Social Psychology*, 1972, 22, 171–193.
- Milgram, S. Behavioral study of obedience. *Journal of Abnormal and Social Psychology*, 1963, 67, 371–378.
- Mischel, W. *Personality and assessment*. New York: Wiley, 1968.
- Mischel, W., & Peake, P. K. Beyond déjà vu in the search for cross-situational consistency. *Psychological Review*, 1982, 89, 730–755.
- Nisbett, R. E., Borgida, E., Crandall, R., & Reed, H. Popular induction: Information is not necessarily informative. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases*. Cambridge University Press, 1982.
- Nisbett, R. E., Caputo, C., Legant, P., & Maracek, J. Behavior as seen by the actor and as seen by the observer.

- Journal of Personality and Social Psychology*, 1973, 27, 154-164.
- Nisbett, R., & Ross, L. *Human inference: Strategies and shortcomings of social judgment*. Englewood Cliffs, N.J.: Prentice-Hall, 1980.
- Quattrone, G. A. Overattribution and unit formation: When behavior engulfs the person. *Journal of Personality and Social Psychology*, 1982, 42, 593-607.
- Ross, L. The intuitive psychologist and his shortcomings: Distortions in the attribution process. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 10). New York: Academic Press, 1977.
- Suls, J., & Mullen, B. From the cradle to the grave: Comparison and self-evaluation across the life-span. In J. Suls (Ed.), *Psychological perspectives on the self* (Vol. 1). Hillsdale, N.J.: Erlbaum, 1982.
- Tversky, A., & Kahneman, D. Availability: A heuristic for judging frequency and probability. *Cognitive Psychology*, 1973, 5, 207-232.
- Tversky, A., & Kahneman, D. Judgment under uncertainty: Heuristics and biases. *Science*, 1974, 185, 1124-1131.
- Wells, G. L., & Harvey, J. H. Do people use consensus information in making causal attributions? *Journal of Personality and Social Psychology*, 1977, 35, 279-293.

Received September 7, 1982

Revision received January 14, 1983 ■

Manuscripts Accepted for Publication in the Section Attitudes and Social Cognition

- Causes and Effects of Causal Attribution. Reid Hastie (Department of Psychology, Northwestern University, Evanston, Illinois 60201).
- Analogical Reasoning in Social Judgment: The Importance of Causal Theories. Stephen Read (Department of Psychology, Northwestern University, Evanston, Illinois 60201).
- The Effects and Aftereffects of Stressor Expectations. Shirlynn Spacapan (Department of Psychology, Carnegie-Mellon University, Pittsburgh, Pennsylvania 15213) and Sheldon Cohen.
- Self-Presentation, Audience Transfer, Choice, and Cognitive Dissonance: Induced Compliance and the Two-Experiment Ploy. Roy F. Baumeister (Department of Psychology, Case Western Reserve University, Cleveland, Ohio 44106) and Dianne M. Tice.