Do People Use Consensus Information in Making Causal Attributions?

Gary L. Wells
Ohio State University

John H. Harvey
Vanderbilt University

Nisbett and Borgida's failure to obtain effects for consensus information on person versus situation attributions was questioned regarding their operationalization of consensus. Nisbett and Borgida's finding was replicated, but an operationalization of consensus designed to reflect Kelley's conceptual criterion of perceived covariation and to represent more levels of consensus resulted in significant consensus effects on person versus situation attributions. In Study 2, consensus information was found to affect specific traits judged to be relevant to the given behavior. In addition, the second study found significant effects for consensus information on the subjects' predictions of target persons' behavior. Included were discussions of (a) issues regarding operationalization of the consensus variable, (b) issues regarding operationalization of dependent variables in consensus research, and (c) the relationship between consensus research in the area of attribution and base-rate research in the area of the psychology of prediction (Kahneman & Tversky).

In his influential attribution statement, Kelley (1967) proposed that individuals explain their own and others' behavior in a manner similar to that of a scientist identifying cause and effect relations. In general, an effect is attributed to the particular causes with which it covaries. Kelley suggested that the individual makes use of three different kinds of information: consistency (Does the actor always respond in the same way to the stimulus?), distinctiveness (Does the actor respond in that way to all or only a few stimuli of that type?), and consensus (Do all or only a few people respond to the stimulus in the same way as does the actor?). Each of these three pieces of information can provide covariation information. For example, behavior (the effect) probably will be perceived
as covarying with the stimulus (a possible cause) as opposed to the person (another possible cause) when most other people behave in that manner toward the same stimulus (high consensus).

Recently, Nisbett and Borgida (1975) addressed the issue of whether consensus information has any effect on causal attribution, and they concluded that "there is considerable evidence suggesting that it does not" (p. 932). We will argue that Nisbett and Borgida's "no effect" conclusion regarding the consensus variable may be premature in view of the particular operations that have been employed to investigate consensus effects. The literature on which Nisbett and Borgida based their conclusion is discussed below. It is our contention that in most studies, no effect for consensus was found because of inadequate operationalization of either the dependent or the independent variable.

The Dependent Variable Problem in Consensus Research

Nisbett, Borgida, Crandall, and Reed (1976) manipulated consensus in three different studies in a way designed to make the subjects feel that their depressive mood was widely shared by people in that particular situation. For example, it was emphasized to new assistant professors that negative reactions in their first year on a faculty were the norm (high consensus), whereas a control group of assistant professors was given no such consensus information. Nisbett et al. reasoned that knowledge indicating that negative affect was widely shared would shift the causal locus away from internal causes to external, situational causes. This prediction was certainly theoretically relevant; however, Nisbett et al. chose to use measures such as mood change, satisfaction with teaching, number of papers written, and number of cigarettes smoked as the dependent variables. There were no effects for consensus information on any of these measures. Nisbett et al. did not specify how these dependent variables are theoretically related to attribution theory or how unambiguous predictions for the variables can be derived from attribution theory. Why should making a situational attribution improve mood state? New faculty members who learn that it is the situation that is causing their feelings of unhappiness will not necessarily show elation at such knowledge; after all, the cause of unhappiness still exists. There is therefore no reason to expect these new faculty members to reduce smoking, increase production, and the like. A more direct line of inquiry based on attribution theory would have been to ask about the subjects' causal attributions for the negative affect. Attribution theory would then predict that the subjects given high-consensus information would attribute their depression more to situational factors than would subjects having low-consensus information.

Another study cited by Nisbett and Borgida (1975) as evidence that consensus has no effect on attribution was conducted by Miller, Gillen, Schenker, and Radlove (1973). These investigators varied consensus information concerning the behavior of participants in the Milgram (1963) obedience study. All their subjects read the procedure section of Milgram's experiment. Some subjects also read the results, including the information that 65% of the participants had delivered the maximum possible shock. The subjects were then told of two persons who had delivered the maximum possible shock. Once again, however, the dependent variable did not directly measure person versus situation attributions. Instead, the subjects rated those two persons on a number of trait scales (e.g., aggressiveness, warmth, likability, dependency). Nisbett and Borgida (1975) have argued that the data in Miller et al.'s study represent a failure for the consensus prediction because consensus had an effect on only 1 of 11 trait ratings. But why these data should necessarily represent a failure of the consensus variable is unclear. Attribution theory predicts that subjects provided with low-consensus information will perceive the cause of behavior to be due to personal factors; however, the specification of which of the possible multitude of personal factors will be affected has not yet been well articulated in attribution theory. If the dependent variable in the study by Miller et al. had been a person versus situation attribution, the varia-
ble would have been more representative of a theoretical dependent variable directly related to attribution theory.

In conclusion, predictive clarity would be facilitated if some external, empirical criterion or criteria were developed to provide a basis for the substitution of dependent variables such as trait ratings and affective responses for the person–situation dependent measure in attribution research. Bem (1972) has suggested that the theoretical relationship between important independent variables in attribution theory and behavioral and physiological responses has not been well articulated by theorists. We are suggesting that Bem’s point may be extended to encompass a host of cognitive–affective responses as well.

The Independent Variable Problem in Consensus Research

Whereas the studies just cited can be questioned in terms of the dependent variables used to assess the effects of consensus, other studies require examination with regard to their operationalization of the independent variable of consensus. Nisbett et al. (1976) conducted a study in which the subjects drank a neutralizing solution between trials in which they taste-tested crackers. Subsequently, the subjects were told they had drunk more (or less) than other people in the experiment. Later, the subjects were asked to estimate how pleasant they found the taste of the solution, how thirsty they were, and how thirsty the crackers made them. Nisbett et al. found no significant effects for their manipulation. Nisbett and Borgida (1975) have cited this study as evidence that consensus has no effect on attributions. However, it is unclear how these operations constituted a manipulation of the degree of consensus because in both experimental conditions the subjects were always in a low-consensus situation. Although the manipulation in this study is certainly related to attribution theory, it does not provide a strong test of the consensus variable.

The three previously mentioned studies by Nisbett et al. (1976) on depressive mood states can also be questioned with regard to the manipulation of consensus. In particular, there is no evidence presented in these studies that the consensus information (i.e., that depression is the norm) was incongruous with the subjects’ a priori beliefs. It is conceivable that the subjects already felt that their depressive mood state was widely shared.

Nisbett and Borgida (1975) cited McArthur (1972) as an example of research in which consensus was not found to “have the anticipated effects on the perception of others” (p. 933). They argued that even though McArthur found that consensus significantly affected causal attributions, it accounted for only 3% of the variance, whereas consistency and distinctiveness accounted for 20% and 10%, respectively. However, there are many ways to operationalize a conceptual variable (Mills, 1969). Certain operations of the same conceptual independent variable may have more impact on a dependent variable than other operations have, and until a wider range of operationalizations is sampled, percentage of variance data should be viewed with caution in considering the relative impact of theoretical independent variables. In support of our argument in favor of a cautious approach, it should be noted that Ruble and Feldman (1976) recently reported evidence that with stimuli similar to McArthur’s, consensus effects can be quite robust when the consensus information is presented subsequent to consistency and distinctiveness information.

Finally, Nisbett and Borgida placed considerable confidence in their “no effect” conclusion on the basis of their own studies (Nisbett & Borgida, 1975). They gave subjects a description of two experiments: (a) the high-fear condition of a study by Nisbett and Schachter (1966) concerned with individuals’ tolerance of shock and (b) the emergency condition of a study by Darley and Latané (1968) concerned with the helping of a person in distress. After reading a description of each experiment, some subjects were provided with the results (consensus information). For example, half of the subjects were told that in the shock experiment 16 of 34 participants had shocked all the way to a jolt causing the entire arm to jerk; the other half of the
subjects were not given the results. Similarly in the case of the helping experiment, some subjects were provided with the results (e.g., 6 people had never helped a confederate who was having an obvious seizure), and others were not given the results. Subsequently, all subjects were told of one participant in each experiment who had acted in the most extreme manner. For the shock study, they were told that Bill J. had accepted shock all the way to the 37th step, causing his entire arm to jerk. For the helping study, the subjects were told that Greg R. had never helped the confederate in distress. They then were asked to explain the behavior of each participant in the two experiments by indicating to what extent the situation or the individual’s personality had been the cause of his behavior. The consensus information was effective in the sense that the subjects receiving that information indicated that more participants acted in the extreme manner (i.e., the way Bill J. and Greg R. acted) than did the subjects given no consensus information. However, the consensus information had no effect on causal attributions, and Nisbett and Borgida (1975) therefore concluded that “subjects’ judgments about the role of an individual's dispositional qualities in producing extreme behavior . . . are unaffected by knowledge that extreme behavior was modal in the situation” (p. 942).

Attribution theory, however, predicts an effect for consensus only because it is a variable that is informative regarding covariation of an internal or external factor with a particular effect. If a sample of people is perceived as somehow biased, the fact that the entire sample acts in a similar way may be seen as uninformative regarding the question of whether or not the behavior is covarying with the situation or the person. Since it was not indicated that Nisbett and Schachter (1966) and Darley and Latané (1968) had randomly selected their participants, the subjects in Nisbett and Borgida’s (1975) study were free to assume that the behavior was covarying with the subject-selection procedure, with subject self-selection, and/or with the situation.

Kelley (1967) defined consensus as the case in which “attributes of external origin are experienced the same by all observers” (p. 197). Did Nisbett and Borgida (1975) operationalize consensus in a manner consistent with this definition? In their discussion of how to determine whether the conceptual variable has been operationally realized, Aronson and Carlsmith (1968) suggested that the experimenter measure other dependent variables that are theoretically tied to the conceptual variable. In terms of this suggestion, Nisbett and Borgida’s consensus manipulation probably would have had little effect on subjects’ responses to an inquiry about how many people would shock (or help) at various levels if all people of similar sex and age were placed in that experimental situation. Again, when subjects are not aware of the random sampling procedure, they may believe that the results (consensus manipulation) are due to a systematically biased sample. This expectation should be particularly true for the kinds of behavior described in the Nisbett and Borgida study, since such behavior generally is perceived as atypical of a “normal” group of people. As teachers of social psychology can attest, students are usually amazed at how Milgram found so many obedient and aggressive people in his 1963 research. Only when they are explicitly informed of the fact that Milgram’s subjects generally were representative of the population do they begin to attribute the high administration of shock to the situation.

Study 1

The first study reported here was designed to replicate Nisbett and Borgida’s (1975) study using the description of Nisbett and Schachter’s (1966) procedure. In the present study, however, half of the subjects were informed of the random, representative sampling of participants. In Nisbett and Borgida’s study, the consensus information was that 16 persons of a possible 34 had acted in an extreme manner (taking shock all the way to the 37th level). In the present study, we added high consensus information, stating that 26 people had acted in that extreme manner, and low consensus information, stating that 1 person acted in that extreme manner. Nisbett and Borgida’s consensus-infor-
Consensus and Attributions

283

mation condition will hereafter be referred to as moderate consensus. The design in the present study was a 4 (High, Moderate, Low, and No Consensus) × 2 (Knowledge of Random Sampling, No Knowledge of Random Sampling) between-subjects factorial.

Specifically, we expected that without knowledge of random, representative sampling, consensus information would not significantly affect person–situation causal attributions (as in Nisbett & Borgida, 1975). However, the subjects who were informed of random, representative sampling were expected to attribute the target person's behavior increasingly to the situation (and decreasingly to the person) under increasing levels of consensus.

In line with Aronson and Carlsmith's (1968) suggestion of probing other measures as a way of determining whether the conceptual variable has been operationally realized, the subjects were asked what proportion of people would accept shock at various levels if all people of similar sex and age were placed in the experiment. We expected that when subjects were not told of the random, representative sampling procedure, their predictions would not be greatly affected by consensus information. On the other hand, we anticipated that when subjects were given random, representative sampling knowledge, they would predict behavior in line with their particular consensus manipulation. Finally, the subjects were asked to rate the target person on likability and on the six personality traits employed by Nisbett and Borgida (1975).

Method

Subjects

The subjects were male and female students from various sections of introductory psychology. Participation constituted partial fulfillment of a course requirement. The subjects participated in small groups set up in such a way that communication between subjects could not occur. There were 13 to 18 subjects in each condition.

Procedure

The procedure described by Nisbett and Borgida (1975) was followed as closely as possible. Upon arrival, the subjects were seated and the experimenter explained that although most psychology experiments involved people like themselves, relatively little was known about college students' perceptions of the experimental situation. It was further explained that in order to learn more about the perceptions of college subjects, an experiment would be described to them in detail and they would be asked for their own impressions.

Manipulation of sampling knowledge. For one half of the subjects, the cover page read, "The next two pages describe an experiment conducted at Long Island University. Read it carefully. You will be asked a few questions about the behavior of subjects in this experiment. The subjects were 34 18- to 22-year-old males." These subjects formed the no-knowledge-of-random-sampling group. For the other half of the subjects, the following information was added:

It is necessary in most psychological research to use random samples of the population in order to assure that the subjects are representative and typical of that population. The present study was

<table>
<thead>
<tr>
<th>Behavior category</th>
<th>High consensus</th>
<th>Moderate consensus*</th>
<th>Low consensus</th>
<th>No consensus-information*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Refused to participate</td>
<td>1</td>
<td>2</td>
<td>26</td>
<td>—</td>
</tr>
<tr>
<td>Stopped shock at tingling fingers</td>
<td>2</td>
<td>1</td>
<td>3</td>
<td>—</td>
</tr>
<tr>
<td>Stopped shock at jolt felt through whole hand</td>
<td>2</td>
<td>6</td>
<td>2</td>
<td>—</td>
</tr>
<tr>
<td>Stopped shock at jolt causing entire forearm to jerk</td>
<td>3</td>
<td>9</td>
<td>2</td>
<td>—</td>
</tr>
<tr>
<td>Went all the way causing entire arm to jerk</td>
<td>26</td>
<td>16</td>
<td>1</td>
<td>—</td>
</tr>
</tbody>
</table>

* These data are identical to those presented by Nisbett and Borgida (1975) in their consensus and no-consensus conditions.
interested in 18- to 22-year-old males, and in order to assure that the subjects were representative of 18- to 22-year-old males, the 1970 census was used. The names of all males in this age range from the Long Island area were placed in a central register. Thirty-four were randomly selected and represented a cross section (e.g., high, medium, low income; large, medium, and small families, etc.) of the population. Each male was contacted by phone and asked to participate, but none was told of the nature of the experiment. Only after arriving for the experiment were they told what they would be required to do.

Subjects given this information formed the knowledge-of-random-sampling group.

**Description of experiment.** Subjects read a description of the high-fear condition in Nisbett and Schachter (1966). Specifically, they were told that the experiment concerned "skin sensitivity," and the test of sensitivity was electric shock. The participants in the original experiment had been given a brief lecture on the importance of shock sensitivity research for the prevention of electrical accidents and for the control of shock dosage for patients undergoing electroshock therapy. They also had been told that the experimenters were interested in the skin sensitivity caused by a particular drug (which, it turned out, had no detectable effect). The shock had been described as extremely painful but causing no permanent damage. The description our subjects read indicated that the first two or three shocks had been too weak to be felt but that the shocks had then progressed from a tingle in the fingers to a jolt causing the entire arm to jerk; the participants had been told to report when they first felt the shock, when the shock first became painful, and when the shock was too painful to endure.

**Table 2**

<table>
<thead>
<tr>
<th>Consensus</th>
<th>No knowledge of random sampling</th>
<th>Knowledge of random sampling</th>
</tr>
</thead>
<tbody>
<tr>
<td>High consensus</td>
<td>26.1a 18</td>
<td>55.3b 15</td>
</tr>
<tr>
<td>Moderate consensus</td>
<td>23.0a 15</td>
<td>47.8b 15</td>
</tr>
<tr>
<td>Low consensus</td>
<td>5.4a 17</td>
<td>6.1a 13</td>
</tr>
<tr>
<td>No consensus-information</td>
<td>17.8a 17</td>
<td>10.9a 16</td>
</tr>
</tbody>
</table>

*Note.* Means not sharing a common subscript differ at the .05 level by a Newman-Keuls test. This test was performed by using the harmonic mean of the number of subjects per cell.

**Manipulation of consensus.** A fourth page was attached to the end of the description of the experiment giving the results that served as consensus information. Table 1 presents the consensus information given to the subjects according to condition. The no-consensus-information subjects received no results.

**Dependent variables.** The description of the experiment and the results were not available to the subjects while they answered the dependent variable questionnaire. All subjects received the same questionnaire. The first item involved a brief description of one of the participants in the experiment: Bill J. was a sophomore majoring in history who had gone all the way up to the 37th step in the series, although the last few shocks had clearly caused him considerable pain. The subjects were asked to rate, on a 7-point scale, the extent to which Bill's personality was responsible for his going all the way up in the series versus the extent to which the situation was responsible (this attributional question was identical to the one employed by Nisbett and Borgida, 1975). In addition, the subjects were asked to rate Bill's likability on a 7-point scale and to rate him on several personality traits (conforming, weak-willed, apathetic, cruel, fearful, timid). Finally, the subjects were provided with the five categories of behavior in Table 1 and asked what they thought the distribution of behavior would be (in percentages adding to 100%) if all 18- to 22-year-old males participated in this experiment.

**Results**

Table 2 presents data concerning the subjects' responses to the question "What do you think the distribution of behavior would be if all 18- to 22-year-old males participated in this experiment?" These data represent the subjects' mean predictions of the percentage of persons who would take shock all the way to the 37th step (i.e., perceived consensus). On this question, a 4 (Consensus Level) X 2 (Knowledge of Random Sampling) analysis of variance produced a significant interaction, \( F(3,118) = 4.57, p < .05 \). A Newman-Keuls analysis was performed, and the results for this analysis are presented in Table 2. Significant differences between the knowledge-of-random-sampling group and the no-knowledge-of-random-sampling group were found at the high- and moderate-consensus levels. Also, as expected, consensus information significantly affected the predictions of subjects in the knowledge-of-random-sampling conditions: Both high and moderate consensus information led to higher predictions than did either
low or no consensus information. In the no-knowledge-of-random-sampling conditions, consensus information did not significantly affect predictions; however, there was evidence that at the extremes (low vs. high consensus), which were not employed in Nisbett and Borgida's (1975) study, a consensus effect tended to occur (although nonsignificant by the Newman-Keuls analysis).

Mean responses to the 7-point-scale question concerning causal attribution are presented in Table 3. An analysis of variance resulted in a significant 4 (Consensus Level) × 2 (Knowledge of Random Sampling) interaction, \( F(3, 118) = 8.06, p < .01 \). The conceptual analysis outlined earlier calls for a specific pattern of interaction. That pattern corresponds to a 1 df contrast of the linear trend interaction in the 3 × 2 design that omits the no-consensus-information conditions. Because of unequal cell size and unequal spacing of the levels of consensus information (regarding the number of participants who had taken shock to the extreme), the linear trend coefficients for the 3 × 2 design were computed according to Gaito (1965). As predicted, this analysis yielded a significant linear interaction, \( F(1, 87) = 14.1, p < .01 \). The means and Newman-Keuls analysis are presented in Table 3.

Neither the 4 × 2 analysis nor the 3 × 2 linear interaction analysis resulted in significant effects on any of the trait ratings or on the likability measure.

**Discussion**

The results of the first study provide generally strong support for the predictions of a consensus effect when subjects were given information about the random sampling procedure associated with the shock tolerance study. However, consistent with Nisbett and Borgida’s (1975) data for the shock tolerance study, there was no significant consensus effect when subjects were not given information about random sampling.

The results of Study 1 also show that consensus information significantly affected the subjects’ predictions regarding the behavior of all 18- to 22-year-old males primarily when they had information about the random selection procedure. These data suggest that the information about random selection was necessary to reflect adequately the essential theoretical connotation of consensus, that is, that many people do or would engage in a particular action (Kelley, 1967).

**Study 2**

A second study was conducted to address some important issues that were not resolved by the first study. First, we were concerned with why the trait scales, which on the surface appeared relevant to the behavior in question, were not affected by consensus information. Second, in Study 1 we did not investigate Nisbett and Borgida’s (1975) other major finding that subjects ignored consensus information in their predictions of a target person’s behavior. This finding does not relate to the random-sampling problem, since the target person was randomly drawn from the participants, but it is an important finding to consider when evaluating evidence concerning people’s use of consensus information. In addition to examination of these issues, an attempt was made to test in a more direct way subjects’ beliefs about the representativeness of a given sample of participants.

A possible explanation for the nonsignificant results for the trait ratings in the first study was that the traits chosen were not generally perceived by subjects as being di-
rectly relevant to the behavior. For example, the trait "masochistic" might have had a stronger perceived relevance to the behavior of taking shock. Nisbett and Borgida (1975) reported having chosen two of the traits used in their research from interviews with subjects. This procedure, however, failed to lead to clearly significant results for any of the trait ratings that the subjects had been asked to make.

In Study 2, therefore, a more systematic approach to the trait–behavior relationship was adopted. The procedure for this study involved a description of the helping study (Darley & Latane, 1968) that had been used in Nisbett and Borgida's (1975) research. The trait–behavior issue was approached by first adding new traits to the list that had been used in Study 1. Also, and more important, the no-consensus-information subjects were asked to estimate the percentage of participants who possessed each trait both for those participants in the experiment who had helped as quickly as possible and for those who had not helped at all. In effect, these percentage estimates made by subjects in the no-consensus-information condition were employed as estimates of the perceived strength of the relationship between the trait dimensions and the specific behavior. Accordingly, it was predicted that if engaging in the behavior in question versus not engaging in it produced significant differences in estimates regarding particular traits, those traits would be affected by consensus information.1

Nisbett and Borgida (1975) have pointed out that consensus information is essentially base-rate information similar to that involved in Kahneman and Tversky's (1973) well-known research, a principal finding of which was that people do not use base-rate information optimally in making various kinds of judgments and predictions. Nisbett and Borgida have suggested that their evidence for no significant effect of consensus on attributions has a strong parallel in Kahneman and Tversky's research and shows that people generally are not influenced by certain types of logically compelling information. However some questions can be raised about the cogency of this parallel. First, Nisbett and Borgida summarized Kahneman and Tversky's findings with the statement that people "ignore category base rates" (p. 934), and they are not alone in making this type of summary statement (e.g., Fischhoff, 1976; Simon, 1976). However, Kahneman and Tversky actually found a significant overall effect for base-rate information at $p < .01$ (1973, p. 242). Perhaps the basis for the frequently and imprecisely reported finding deriving from their research is Kahneman and Tversky's unfortunate use of the term "ignore." In actuality, their data indicate that although people underutilize base-rate information, they nevertheless do make use of this information.

If Kahneman and Tversky (1973) found significant effects for base-rate information, then why did Nisbett and Borgida's (1975) subjects ignore base-rate information in their predictions? It appears that Nisbett and Borgida did not control for "individuating evidence." Whereas Kahneman and Tversky attempted to ensure that some of the descriptions of their target cases were relatively uninformative, Nisbett and Borgida apparently took no such precautions, and it seems that their 150-word descriptions of target cases provided considerable personal, individuating information (e.g., "Physically, he was short and somewhat stocky . . . somewhat quiet and a bit aloof in manner," 1975, p. 937). In addition, when subjects were asked the prediction questions, Nisbett and Borgida's instructions reminded them that they had received detailed personal descriptions of the participants and directed them to "Please go back and refresh your memory about them so that you can answer the fol-

---

1 Fishbein and Ajzen (1974) have proposed the use of this type of index to assess attitude–behavior relatedness. They argue that a "behavior is relevant or valid if performance of that behavior indicates a favorable attitude and its nonperformance indicates an unfavorable attitude" (p. 70). Our formulation substituted traits for attitudes. In line with the argument concerning trait dimensions in the present study, Fishbein and Ajzen noted that "most investigators select behavioral criteria that they assume have these characteristics. Unfortunately, it appears that more often than not, their intuition is misleading" (p. 68).
following questions.” Therefore, the present study contained prediction questions that did not provide the type of individuating information provided by Nisbett and Borgida, and we expected that consensus information would significantly affect target predictions.

Finally, measures were taken in Study 2 to provide more direct evidence about the relationship between the subjects’ knowledge of random sampling and their perception of representativeness of the sample in the experiment described to them. Study 1 provided indirect evidence about representativeness by asking the subjects to predict the behavior of all 18- to 22-year-old males. In Study 2, the subjects were asked about the extent to which the participants were “nonrepresentative” and “different” from introductory psychology students in general. It was expected that if the no-knowledge-of-random-sampling subjects were attributing the increasing levels of consensus information to nonrepresentative sampling, perceived nonrepresentativeness and differentness ratings would increase directly with consensus levels. On the other hand, it was expected that the knowledge-of-random-sampling subjects would not show significant increases on these measures as a function of consensus level.

**Method**

**Subjects**

The subjects were 96 male and female students of introductory psychology who participated in partial fulfillment of a course requirement. They participated in groups set up in such a way that communication among subjects could not occur. There were 12 subjects in each condition.

**Procedure**

The procedure was virtually identical to that used by Nisbett and Borgida (1975) in describing the helping experiment. As described below, additions to or modifications of this procedure were implemented to achieve the objectives of Study 2.

**Description of the helping experiment.** The subjects read a description of the helping experiment taken verbatim from Nisbett and Borgida (1975, p. 936). Included in this information was a description of the participants as 15 New York University students, ranging in age from 17 to 22 years. The information indicated that a group of six of the participants were seated in small individual rooms and told that they would discuss over an intercom system personal problems associated with college life.

One participant, actually a confederate, started to discuss his difficulties in adjusting to New York, mentioned that he was prone to epileptic seizures, and then, after other members of the group had spoken, began to have a seizure that could be clearly detected over the intercom by his stuttering.

**Manipulation of sampling knowledge.** For half of the subjects, the following information was added to the description of the experiment:

It is important to point out that these subjects were randomly selected from the introductory psychology students. That is, they were not allowed to choose whether they would sign up but were selected by drawing their names from an urn that had every student’s name in it. This procedure assured that the experiment was conducted on a representative sample, not just students who sign up for experiments.

The subjects who received this information constituted the knowledge-of-random-sampling condition, whereas those who did not receive this information constituted the no-knowledge-of-random-sampling condition.

**Manipulation of consensus.** The last page, after the description of the experiment, provided the results of the experiment that served as the consensus information. Table 4 presents the consensus information given to subjects according to condition. The no-consensus-information subjects did not receive any results.

**Dependent Measures**

**Assessment of the trait–behavior relationship.** The description of the experiment and the results were not available to the subjects while they answered the dependent variable questionnaire. The subjects in the no-consensus-information condition were asked to estimate what percentage of the persons who had not helped at all could be characterized by each of the following adjectives: considerate, helpful, rebellious, concerned, strong-minded, likable, kind, responsive, decisive, dependable, passive, cold, and ethical. In addition, the subjects were asked to estimate the percentage for each of the above traits for persons who had helped as soon as the accomplice began stuttering. The difference in the percentage estimates for each trait was considered indicative of the strength of the trait–behavior relationship; the greater the difference, the stronger the presumed relationship.

**Recall of consensus information.** Subjects in all conditions were asked to indicate what they thought the behavior of all 15 participants had been by placing a number to show distribution in each of the six behavior categories (as shown in Table 4). For the consensus-information subjects, this measure...
Table 4
Consensus Information Given to Subjects on Number of Participants in Each Behavior Category in Study 2

<table>
<thead>
<tr>
<th>Behavior category</th>
<th>High consensus</th>
<th>Moderate consensus*</th>
<th>Low consensus</th>
<th>No consensus-information*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Helped as soon as accomplice began stuttering</td>
<td>0</td>
<td>0</td>
<td>6</td>
<td>—</td>
</tr>
<tr>
<td>Helped when accomplice asked for help</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>—</td>
</tr>
<tr>
<td>Helped when accomplice stammered that he had a seizure coming on</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>—</td>
</tr>
<tr>
<td>Helped by end of speech when accomplice was choking</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td>—</td>
</tr>
<tr>
<td>Helped within 4 minutes after end of accomplice’s speech</td>
<td>3</td>
<td>5</td>
<td>0</td>
<td>—</td>
</tr>
<tr>
<td>Never helped</td>
<td>11</td>
<td>6</td>
<td>1</td>
<td>—</td>
</tr>
</tbody>
</table>

* These data are identical to those presented by Nisbett and Borgida (1975) in their consensus and no-consensus conditions.

served as a manipulation check; for the no-consensus-information subjects, it served as an estimation of naive base-rate expectancies.

**Attribution question.** The next item on the questionnaire was a brief description of one of the participants, Greg R., a freshman majoring in psychology. He had never helped the accomplice, although when the mock seizure began, he had been heard to exclaim “Jesus Christ!” into his microphone, and when the experimenter came to his cubicle, Greg had asked nervously, “Is that guy all right?” The subjects were asked to rate, on a 7-point scale, the extent to which Greg’s personality had been responsible for his not having helped versus the extent to which the situation had been responsible.

**Trait questions.** In the next set of questions, the subjects were asked to rate Greg R. on 13 personality traits. The traits were assessed on 9-point continua anchored by the following adjectives: incon siderate–considerate, nonhelpful–helpful, conforming–rebelling, apathetic–concerned, weak willed–strong minded, likable–not likable, cruel–kind, nonresponsive–responsive, indecisive–decisive, dependable–unde pendable, passive–active, ethical–nonethical, cold–warm.

**Prediction questions.** All subjects were told that the names of 3 participants (Allan V., 22 years old; Mark N., 19 years old; and Charles H., 20 years old) had been randomly drawn from the total of 15 participants. Subjects were asked to predict how quickly these participants would have helped by placing them in one of the six behavior categories. In addition, subjects were asked to check one of the six categories in response to the question, “How quickly do you think you would have helped?”

**Estimates of representativeness and differentness.** Finally, subjects were asked to indicate on 11-point scales to what extent these participants were generally “representative” of and “different” from introductory psychology students in general.

**Results**

**Recall of Consensus Information**

A 4 (Consensus Level) × 2 (Knowledge of Random Sampling) analysis of variance on recall of the number of participants (converted to percentages) who had never helped revealed a significant effect for consensus information, \( F(3, 88) = 193.5, \ p < .001, \) no effect for knowledge of random sampling, and no interaction. Newman-Keuls analyses showed that the low-consensus and no-consensus-information conditions (\( M_s = .7 \) and 0, respectively) were significantly lower (\( ps < .05 \)) than the moderate-consensus condition (\( M = 33.7 \)), which in turn was lower than the high-consensus condition (\( M = 63.0 \)).

**Prediction Questions**

The subjects’ predictions about how quickly each of the three target persons would have helped were scored from zero (helped as soon as the accomplice began stuttering) to five (never helped) and averaged for each subject. A 4 (Consensus Level) × 2 (Knowledge of Random Sampling) analysis of variance
revealed a strong effect for consensus information, $F(3, 88) = 40.9, p < .001$, but no effect for knowledge of random sampling and no interaction. A Newman-Keuls analysis revealed that the high-base-rate, or high-consensus, conditions ($M = 3.2$) produced significantly higher ($p < .05$) predictions of the length of time it would have taken target subjects to respond than did the moderate-consensus conditions ($M = 2.3$), which in turn were higher ($ps < .05$) than the low-consensus and no-consensus-information conditions ($Ms = 1.5$ and $1.4$, respectively). There was no effect for consensus information on the subjects' estimates of how quickly they thought they would have helped if they had been subjects in the experiment.

**Attribution Question**

Responses to the question on situation-versus-personality attribution for the behavior of Greg R. produced significant main effects for consensus, $F(3, 88) = 18.2, p < .01$, and for knowledge of random sampling, $F(1, 88) = 5.5, p < .05$, and a significant Consensus $\times$ Knowledge of Random Sampling interaction, $F(3, 88) = 3.94, p < .05$. As in Study 1, a test was conducted for linear trend interaction in the $2 \times 3$ design that omits the no-consensus-information condition. Because of equal cell size and equal spacing of consensus information (for participants who had never helped), the numerical values $1, 0, -1$ were used as the linear coefficients. This analysis resulted in a significant linear trend interaction, $F(1, 66) = 16.5, p < .01$. The means and Newman-Keuls analysis are presented in Table 5. This pattern suggests that, as predicted, these effects were primarily due to differences across levels of consensus for subjects who had knowledge of random sampling.

**Trait Questions**

The results for the differences between the no-consensus-information subjects' estimates of what percentage of the participants possessed each of the 13 traits for those who had helped as soon as the "subject began stuttering" and their estimates for those who had never helped were analyzed as a $2$ (Knowledge of Random Sampling, No Knowledge of Random Sampling) $\times 2$ (Helped Immediately, Never Helped) mixed analysis of variance (with the latter factor being the within-subject factor) for each of the 13 traits. The results revealed that for 4 of the 13 traits, the behavior in question significantly affected estimates of the number of people having the trait. As expected, there were no significant effects for knowledge of sampling procedure. The four trait ratings for which there were significant differences were helpful, $F(1, 22) = 4.8, p < .05$; responsive, $F(1, 22) = 5.7, p < .05$; decisive, $F(1, 22) = 6.3, p < .05$; and dependable, $F(1, 22) = 4.3, p < .05$.

The subjects' trait ratings of Greg R. were analyzed by a $4$ (Consensus Level) $\times 2$ (Knowledge of Random Sampling) analysis of variance. The results revealed that of the nine traits that the no-consensus-information subjects perceived as being unrelated to the helping behavior, none was affected by consensus information. However, three of the four traits that the no-consensus-information subjects perceived as related to the helping behavior were significantly affected by consensus. Consensus main effects were obtained for ratings of responsive, $F(3, 88) = 7.9, p < .01$; decisive, $F(3, 88) = 16.4, p < .01$; and dependable, $F(3, 88) = 3.5, p < .05$. The fourth trait, helpfulness, was unaffected by consensus information. In addition, decisive-

<table>
<thead>
<tr>
<th>Consensus</th>
<th>No knowledge of random sampling</th>
<th>Knowledge of random sampling</th>
</tr>
</thead>
<tbody>
<tr>
<td>High consensus</td>
<td>3.51&lt;sup&gt;b&lt;/sup&gt;</td>
<td>4.75&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>Moderate consensus</td>
<td>2.91&lt;sup&gt;b&lt;/sup&gt;</td>
<td>3.50&lt;sup&gt;b&lt;/sup&gt;</td>
</tr>
<tr>
<td>Low consensus</td>
<td>2.58&lt;sup&gt;b&lt;/sup&gt;</td>
<td>2.42&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
<tr>
<td>No consensus-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>information</td>
<td>2.66&lt;sup&gt;b&lt;/sup&gt;</td>
<td>2.50&lt;sup&gt;be&lt;/sup&gt;</td>
</tr>
</tbody>
</table>

*Note.* Means not sharing a common subscript differ at the .05 level by a Newman-Keuls test. Higher numbers indicate more situational and less dispositional attribution of causality on a 7-point scale.
ness produced a significant Consensus × Knowledge of Random Sampling interaction, $F(3, 88) = 3.2, p < .05$. The test of linear trend interaction for the $3 \times 2$ design yielded a significant linear interaction for decisiveness, $F(1, 66) = 12.6, p < .01$. No other effects for any of the 13 traits were significant. Table 6 presents the means and a Newman-Keuls analysis of the three traits significantly affected by consensus information.

### Representativeness and Differentness Questions

A 4 (Consensus Level) × 2 (Knowledge of Random Sampling) analysis of variance on the questions regarding the extent to which the participants were perceived as representative of and different from introductory psychology students in general revealed significant main effects for consensus information, $F$s$(3, 88) = 19.2$ and $5.1$, respectively, $ps < .01$; for knowledge of random sampling, there was a significant main effect on the representativeness question, $F(1, 88) = 46.8, p < .01$, and a marginal main effect on the differentness question, $F(1, 88) = 3.9, p < .06$. Also, a marginally significant Consensus × Knowledge of Random Sampling interaction was obtained for the representativeness question, $F(3, 88) = 2.26, p = .086$. The test for linear trend interaction in the $3 \times 2$ design yielded significant linear interactions both for the representativeness question, $F(1, 66)$

### Table 6

<table>
<thead>
<tr>
<th>Sampling knowledge</th>
<th>High consensus</th>
<th>Moderate consensus</th>
<th>Low consensus</th>
<th>No consensus-information</th>
</tr>
</thead>
<tbody>
<tr>
<td>No knowledge of random sampling</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nonresponsive-responsive</td>
<td>4.00$\dagger$</td>
<td>3.25$\dagger$</td>
<td>3.00$\dagger$</td>
<td>2.92$\dagger$</td>
</tr>
<tr>
<td>Indecisive-decisive</td>
<td>3.92$\dagger$</td>
<td>3.42$\dagger$</td>
<td>3.00$\dagger$</td>
<td>3.00$\dagger$</td>
</tr>
<tr>
<td>Dependable-undependable</td>
<td>5.58$\dagger$</td>
<td>7.00$\dagger$</td>
<td>7.00$\dagger$</td>
<td>7.08$\dagger$</td>
</tr>
</tbody>
</table>

| Knowledge of random sampling |                |                    |               |                         |
| Nonresponsive-responsive | 4.92$\dagger$ | 3.58$\dagger$ | 2.92$\dagger$ | 2.58$\dagger$ |
| Indecisive-decisive | 5.00$\dagger$ | 4.08$\dagger$ | 2.76$\dagger$ | 2.75$\dagger$ |
| Dependable-undependable | 5.50$\dagger$ | 6.75$\dagger$ | 7.33$\dagger$ | 7.17$\dagger$ |

Note. Means without common subscripts within a trait category differ at the .05 level by a Newman-Keuls test. Higher numbers indicate the target person was perceived as more responsive, more decisive, and less dependable on a 9-point scale.

### Table 7

<table>
<thead>
<tr>
<th>Sampling knowledge</th>
<th>High consensus</th>
<th>Moderate consensus</th>
<th>Low consensus</th>
<th>No consensus-information</th>
</tr>
</thead>
<tbody>
<tr>
<td>No knowledge of random sampling</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Representativeness</td>
<td>5.3$\dagger$</td>
<td>7.4$\dagger$</td>
<td>8.4$\dagger$</td>
<td>9.1$\dagger$</td>
</tr>
<tr>
<td>Differentness</td>
<td>6.6$\dagger$</td>
<td>5.2$\dagger$</td>
<td>4.0$\dagger$</td>
<td>4.2$\dagger$</td>
</tr>
</tbody>
</table>

| Knowledge of random sampling |                |                    |               |                         |
| Representativeness | 8.8$\dagger$ | 9.0$\dagger$ | 9.4$\dagger$ | 10.4$\dagger$ |
| Differentness | 4.8$\dagger$ | 4.3$\dagger$ | 4.1$\dagger$ | 4.3$\dagger$ |

Note. Means within a perception category without a common subscript differ at $p < .05$ by the Newman-Keuls procedure. Higher numbers mean that the participants were perceived as more representative of and more different from introductory psychology students in general on 11-point scales.
CONSENSUS AND ATTRIBUTIONS

291

= 12.4, \( p < .01 \) and for the differentness question, \( F(1, 66) = 9.5, p < .01 \). The means and Newman-Keuls results are presented in Table 7. These data revealed that consensus information had an effect primarily in the no-knowledge-of-random-sampling conditions.

Discussion

The main finding of these two studies is that consensus information significantly affected attributions in a manner consistent with attribution theory predictions. For consensus information to produce effects of significant magnitude, however, knowledge of random sampling was required to prevent subjects from dismissing the increasing levels of consensus information as being due to a non-representative sample, and in Study 2, more extreme levels of consensus were required than Nisbett and Borgida (1975) employed. Behavior was attributed to the situation to a greater degree under high consensus and when the subjects were told that the participants were a random, representative sample of students. This finding is in accord with the derivation from Kelley's (1967) covariation principle concerning consensus, specifically: Behavior that is clearly in line with what most other people do or would do is attributed to the situation.

The subjects not informed of random, representative sampling, however, appear to have made attributions to the nature of the sample. There are two lines of evidence in support of this interpretation. First, in Study 1, the no-knowledge-of-random-sampling subjects' predictions of the behavior of all 18- to 22-year-old males were not significantly influenced by consensus information, whereas the predictions of subjects told of random sampling were significantly affected by consensus information. Second, the results of Study 2 reveal that the subjects not told of random sampling significantly increased their ratings of how different and nonrepresentative the participants were relative to introductory psychology students in general as a function of increasing levels of consensus. On the other hand, the subjects told of random sampling did not perceive the participants as significantly more different and nonrepresentative as consensus increased. Furthermore, there were main effects that indicate that the subjects without knowledge of random sampling perceived the participants to have been more different and nonrepresentative than did the subjects with knowledge of random sampling. Finally, it does not appear that the knowledge of sampling procedure simply made the consensus information more salient, since recall of consensus information in Study 2 was unaffected by this manipulation; rather, it appears that the sampling knowledge made the consensus information more meaningful.

In addition to the significant consensus effects on the relatively gross person-situation dimension on which attribution theory makes direct predictions, there were significant effects in Study 2 for consensus on specific traits. Although Kelley's (1967) covariation formulation does not specifically state how predictions can be made for a given trait, a procedure was adopted for empirical estimation of trait-behavior relatedness that allowed for a high degree of specification of which traits would be affected by consensus information. With this procedure, in Study 2 it was found that of the nine traits not perceived as related to the behavior in question, none was affected by consensus information. But of the four traits perceived as related, three were affected by consensus information, resulting in an overall prediction accuracy of over 90%. It is not clear why one of these four traits, helpfulness, was not significantly influenced by consensus information. As with the person-situation question, consensus information had a greater effect on trait ratings when subjects were told of random, representative sampling than when they did not receive sampling information. However, it should be noted that despite this difference, the pattern of data in the no-knowledge-of-random-sampling conditions was similar to that in the knowledge-of-random-sampling conditions.

Finally, in Study 2 it was found that subjects did not ignore base-rate (consensus) information in making predictions about the behavior of target cases. We have pointed out that, contrary to several secondary references,
such a finding is not inconsistent with the results of Kahneman and Tversky (1973). This finding is, however, in contradiction to Nisbett and Borgida's (1975) results. Perhaps Nisbett and Borgida's lengthy description of the target cases were simply too informative regarding personal information about the participants to make the consensus information a meaningful basis for category predictions. But it should be noted that the subjects in Study 2 did not make optimal use of the consensus information; in predicting behavior category, the most logical problem-solving strategy would have been to predict the modal category. For example, the high-consensus subjects should have predicted the target cases to have never helped (i.e., Category 5 in Table 4), but in fact the high-consensus subjects predicted helping behavior at a level between Category 3 and Category 4. Nevertheless, consensus, or base-rate, information did significantly affect predictions regarding the behavior of target subjects.

We should emphasize in conclusion that giving subjects knowledge of random sampling is not per se a crucial variable in the generic sense. In many cases, subjects might assume random sampling. However, Nisbett and Borgida (1975) used stimuli for which high levels of consensus are probably unexpected given a typical population. For other types of behavior, people probably do not have prior expectations about the distribution of behavior and are willing to accept various levels of consensus without giving much thought to the unusualness of the sample. A principal point brought out by our research is that consensus must be operationalized in line with the underlying concept of covariation between behavior and the person-situation attribution dimension. If there are plausible alternative covariates in a situation (such as the possible nonrepresentativeness of stimulus persons), these covariates must be taken into account in predicting the relationship between consensus and person-situation attributions.

References


Erratum to Langer and Saegert

In the article, “Crowding and Cognitive Control,” by Ellen J. Langer and Susan Saegert (Journal of Personality and Social Psychology, 1977, Vol. 35, No. 3, pp. 175–182), the scores in the bottom row of Table 1 (p. 180) are incorrect. The bottom row reads as follows:

| Number of correctly completed items | 6.81 | 3.66 | 2.61 | 3.09 |

It should be changed to read:

| Number of correctly completed items | 6.45 | 10.40 | 12.00 | 17.10 |