The Focus of Judgment Effect: A Question Wording Effect Due to Hypothesis Confirmation Bias

Darrin R. Lehman
University of British Columbia

Jon A. Krosnick
Ohio State University

Robert L. West
University of British Columbia

Fan Li
Ohio State University

Three studies examined whether the focus of judgment of a rating scale question can influence responses through a hypothesis confirmation bias. Study 1 showed that focusing the question wording on one of two political parties led nonpartisan subjects who had little relevant knowledge to evaluate media coverage as more hostile to that party. Studies 2 and 3 demonstrated that asking for the likelihood that a specified person was an engineer stimulated more engineerlike ratings in response to a personal description than asking for the likelihood that the person was a lawyer. This effect occurred only when a large amount of hypothesis-compatible information was available to subjects, when they paid close attention to it, and when they had previously been successful at interpreting information as consistent with the hypothesis implied by the question. Together, these studies illustrate two focus of judgment effects and identify factors that regulate their magnitudes.

A number of studies have demonstrated that people's reports of their attitudes and beliefs can sometimes be altered dramatically by slight changes in question wording or format. For example, asking people whether they agree with an assertion supporting a public policy typically yields more positive responses than asking people directly whether they favor or oppose the policy (Schuman & Presser, 1981). This difference presumably results from acquiescence response bias among some respondents. Furthermore, varying the order in which response alternatives are read to respondents can significantly alter their selections in ranking questions (Krosnick & Alwin, 1987). The effects of such variations are referred to as response order effects.

One possible explanation for these two findings is a bias in respondents' information processing toward hypothesis confirmation (Skov & Sherman, 1986; Wason, 1960). For example, McClendon (1991) and Krosnick (1991) proposed that acquiescence response bias may result from respondents focusing their thinking primarily on generating reasons to agree with propositions offered by questions. Similarly, Krosnick and Alwin (1987) argued that response order effects occur because respondents approach their task by trying to generate as many reasons as they can to support selection of each alternative. These arguments are consistent with Ross's (1977) view that perseverance effects (Anderson, Lepper, &

Authors' Note: This research was supported by Grants 410-88-1338 and 410-90-1519 from the Social Sciences and Humanities Research Council of Canada and Grant 37-1715 from the Natural Sciences and Engineering Research Council of Canada to Darrin R. Lehman and by Grant BNS-8920430 from the National Science Foundation to Jon A. Krosnick. Robert L. West is now at the Department of Psychology, Simon Fraser University. We thank Rebecca Collins, Geoffrey Fong, Kenneth Hemphill, Delroy Paulhus, Peter Suedfeld, Norbert Schwarz, and two reviewers for their comments on an earlier draft and James McCloskey for his help with data collection. Correspondence concerning this article should be addressed to Darrin R. Lehman, Department of Psychology, University of British Columbia, 2136 West Mall, Vancouver, B.C., Canada V6T 1Z4.

PSPB, Vol. 18 No. 6, December 1992 690-699

Ross, 1980; Ross, Lepper, & Hubbard, 1975) occur because people are biased toward generating beliefs that support a description of themselves that they have been told is accurate. These arguments are also consistent with evidence illustrating hypothesis confirmation biases in social perception (Snyder, 1981; Snyder & Campbell, 1980; Snyder & Swann, 1978; for reviews of the hypothesis confirmation bias literature more generally, see Fischhoff & Beyth-Maron, 1983; Klayman & Ha, 1987) and in covariation assessment (Chapman & Chapman, 1969; Hamilton, 1979; Jennings, Amabile, & Ross, 1980).

The hypothesis confirmation account suggests that the wording of a question, by conveying an implicit hypothesis, biases respondents to think about the question in a particular way. Although this notion is in line with much of the psycholinguistic research on the nature of presuppositions (see, e.g., Clark, 1985; Clark & Clark, 1977; Levinson, 1983) and, in particular, with research on rules of conversations, or conversational norms (Grice, 1975; Hilton, 1990), there is one important difference. The hypothesis confirmation account is purely cognitive, or intrapsychic, whereas the rules of conversation account is more interpersonal. That is, if a subject shows a focus of judgment effect, according to a rules of conversation explanation, this effect would be viewed as originating from the subject's sense of what the experimenter wanted him or her to focus on, by dint of the way the experimenter asked the question. The hypothesis confirmation account does not hinge on this interpersonal element; subjects can confirm an implicit hypothesis offered by a question without any consideration of what the experimenter was trying to communicate or any awareness of the experimenter's wishes or goals.

In this article, we explore whether bias toward hypothesis confirmation may affect people's answers to questions that ask respondents to indicate the degree to which some specific condition obtains. For example, a question might inquire about the degree to which a television news story about the Arab/Israeli conflict was biased against Israel (Vallone, Ross, & Lepper, 1985). Such a question might ask respondents to select a point on a scale ranging from *not at all* to *a great deal.*

Implicitly, this question offers a hypothesis: that the broadcast was biased against Israel. If, in answering the question, respondents are biased toward hypothesis confirmation, the choice of object to serve as the focus of judgment might be consequential. That is, hypothesis confirmation bias would lead subjects to rate the broadcast as biased against Israel, a pro-Israeli sentiment. If instead the question were to ask how much the broadcast was biased against Arabs, a pro-Arab sentiment might be expressed. This pattern implies a *focus of judgment* effect, whereby ratings are biased toward supporting any implicit assertion or hypothesis that is the focus of a question stem.

The same confirmation bias may also operate in more neutrally worded questions, but in a less obvious manner. That is, a focus of judgment effect could occur even when a question stem does not offer a direct hypothesis. Rather, a question could simply focus respondents' attention on a person or group. A large literature now indicates that evaluations of others are subject to a *positivity bias,* whereby people are inclined toward positive evaluations of individuals or groups (Sears, 1983). Therefore, by focusing a respondent's attention on a person or group, a question may instigate automatic generation of a sympathetic implicit hypothesis, yielding favorable information processing regarding that person or group. This process would then enhance the positivity of evaluations of that person or group.

The primary goals of the present studies were to test these hypotheses and to examine the conditions under which focus of judgment effects are most likely to occur. Surprisingly, social psychological research has not yet offered much insight regarding the conditions under which hypothesis confirmation bias is most likely to appear. The reasoning offered above about focus of judgment effects suggests two factors that may regulate their magnitude: *a priori bias* and *level of knowledge.* Among respondents who have an a priori opinion on the matter in question (and who are thus biased), that opinion may be the hypothesis that they are inclined to confirm. In relation to the Arab/Israeli example, individuals who support Israel may wish to express pro-Israeli sentiment regardless of the phrasing of a question (see, e.g., Sanitioso, Kunda, & Fong, 1990). For individuals who do not have an a priori opinion on the matter in question (and are thus neutral), level of knowledge may become the differentiating variable. A respondent who has little knowledge about the topic will be forced to generate beliefs, whereas a respondent who has a great deal of relevant knowledge can simply retrieve these beliefs in order to make a judgment. The former individual, thus, may be more susceptible to a confirmatory bias because he or she is less likely and able to attempt to disconfirm the hypothesis. In sum, then, the wording of a question may offer unknowledgeable individuals who have no a priori opinion (i.e., neutrals) a hypothesis that they are then inclined to confirm. Therefore, the focus of judgment effect may appear only among such respondents.

To test these ideas, we first designed an experiment based on the study that led us to generate these hypotheses: the demonstration by Vallone et al. (1985) of the hostile media phenomenon. In their study, pro-Arab, pro-Israel, and neutral subjects evaluated the degree of bias in several television news stories about the Arab/

Israeli conflict. As expected, the pro-Arab subjects consistently reported a pro-Israel bias in the stories, while the pro-Israel subjects perceived an anti-Israel bias instead.

Our interest was not in these two groups, but, rather, in the neutral subjects. These subjects' judgments of bias closely resembled those of the pro-Israel subjects, a finding that Vallone et al. speculated was the result of a generally pro-Israel bias among Americans. However, because their subjects were asked about how well they thought Israel had been treated in the broadcasts, our focus of judgment hypothesis could also explain this finding. Given a positivity bias (Sears, 1983) and a hypothesis confirmation set, we would expect that focusing neutral subjects' thinking on Israel would lead these individuals to resemble pro-Israel subjects. Furthermore, the correspondence Vallone et al. observed between the neutral subjects and the pro-Israel subjects was greatest among those neutral subjects who reported knowing little about the Arab/Israeli conflict. This finding is consistent with our claim that neutral subjects with little relevant knowledge are more susceptible to focus of judgment effects than subjects with greater amounts of knowledge. To gain confidence that this interpretation of the findings is correct, however, it was necessary to replicate the procedure Vallone et al. used including an experimental variation of question wording.

The setting for our first study was the province of British Columbia, Canada. British Columbia is ideal for a study of this kind because it has two polarized main political parties: the Social Credit party (SCP), a highly conservative group in power during data collection, and the New Democratic party (NDP), a highly liberal group. Our stimulus materials were newspaper clippings concerning two controversial SCP initiatives: privatization (the selling off of government-run services to the private sector) and decentralization (the dividing of the province into semiautonomous regions). On ideological grounds, the NDP was strongly opposed to these SCP initiatives. Therefore, we expected to replicate the hostile media phenomenon using media coverage of these two issues and partisans from these two groups. Furthermore, we expected that these issues would constitute vehicles to examine whether question wording can produce a focus of judgment effect.

STUDY 1

Method

Subjects. Subjects were 165 volunteers from the University of British Columbia (UBC) and Simon Fraser University (SFU). The SCP partisan group was composed of 39 undergraduate members of SCP student associations at UBC and SFU. The NDP partisan group was composed of 21 undergraduate members of NDP student associations at UBC and SFU. Subjects from both partisan groups were invited to participate at general meetings, and all members who were asked agreed to participate. The neutral group comprised 105 undergraduates at UBC (94% of those asked) who were not members of either student association and who did not belong to either the SCP or the NDP.

Procedure. Subjects completed a pretest questionnaire, read a set of newspaper clippings, and completed a posttest questionnaire in groups ranging from 9 to 71.

Newspaper clippings. Excerpts from 10 articles dealing with the issues of privatization and decentralization were selected from local newspapers. Although a few authors merely reported the initiatives in an objective manner, most sided with one viewpoint or the other. Occasionally, a single author presented both opinions, often in the form of criticisms by the NDP and rebuttals by the SCP leaders. Both sides' points of view were clearly represented, and the 10 excerpts were equated as much as possible in length, clarity, and the quality of SCP and NDP arguments.

Pretest questionnaire. The pretest questionnaire assessed membership in provincial politics and knowledge about the parties and the issues. Knowledge was measured with two direct questions and two questions that assessed exposure to relevant information. Subjects were asked, on 5-point scales from *I know virtually nothing* to *I know almost everything*, "How much do you know about the ideological differences between the NDP and the SCP?" and "How much do you know about the recent conflict concerning the SCP's plans of decentralization and privatization?" Subjects were also asked, on 8-point scales from *less than ½ hour per week* to *more than 16 hours per week*, "How much time have you spent (in the last 3 months) gathering information (from newspapers, magazines, radio, and T.V.) about the government's policies of decentralization and privatization?" and "How much time have you spent (in the last 3 months) talking about these issues with friends, relatives, or strangers?"

Posttest questionnaire. The posttest questionnaire, adapted from Vallone et al. (1985, p. 580) and Ross (personal communication, January, 1988), measured perceptions of fairness and objectivity of the media reports. SCP, NDP, and neutral subjects were randomly assigned to receive posttest questionnaires that were anchored on either the SCP or the NDP. Both versions included five questions, all with 75-mm lines as response scales. Sub-

jects were asked to put a slash through each line to indicate their answer. Responses were scored from 1 to 75 on the basis of measurements.

On the NDP-anchored questionnaire, the first question was "What is your impression of how these particular newsclippings have treated the NDP?" (the endpoints were *strongly biased AGAINST the NDP* and *strongly biased IN FAVOR of the NDP*). The second question was "Consider a group of people who were undecided or ambivalent as to their feelings about the NDP and *the only information they had access to* was from these clippings. Would the information included in these clippings communicate *to these people* a positive (good) impression of the NDP or a negative (bad) impression?" (endpoints were *very NEGATIVE impression of the NDP* and *very POSITIVE impression of the NDP*). The third question was "Based upon your reading of these newsclippings, what do you think are the *personal views* of the editorial staff that put them together (the people who chose what information to include)?" (endpoints were *very PRO NDP* and *very ANTI NDP*). Fourth, subjects were asked, "With respect to the policies of decentralization and privatization, consider all media coverage that you are familiar with (T.V., radio, magazines, etc.). How would you summarize the way in which the media in general has treated the NDP?" (endpoints were *strongly biased AGAINST the NDP* and *strongly biased IN FAVOR of the NDP*). The fifth question was identical to Question 4, except that, instead of asking about "decentralization and privatization," it asked about "the conflict between the SCP and the NDP."

On the SCP-anchored questionnaire, all references to the NDP above were replaced by SCP. Answers were coded so that higher scores on the bias questions indicated more perceived bias against the NDP.

Scale construction. The five items measuring bias were all highly correlated with one another (*r*s ranged from .5 to .8), and item-total correlations ranged from .7 to .8, with an alpha of .88. Because the items appear to have been measuring the same construct, they were combined into one variable called perceived bias, which ranged from 1 (strongly biased against the SCP) to 75 (strongly biased against the NDP).

Results

The hostile media phenomenon. Partisan subjects' responses clearly replicated the hostile media phenomenon (Vallone et al., 1985). Both the SCP and the NDP partisans saw the newspaper clippings as biased against their own group (see Table 1). SCP supporters saw the clippings as biased against the SCP ($M = 26.8$), whereas NDP supporters saw the same clippings as biased against the NDP ($M = 44.2$). The difference between the two mean ratings was significant, $F(1, 58) = 30.85$, $p < .001$, and both the SCP and the NDP groups were significantly different from the neutral group ($M = 31.9$), $F(1, 142) = 5.17$, $p < .025$, and $F(1, 124) = 18.75$, $p < .001$, respectively. Each of the individual items revealed significant differences between SCP and NDP subjects, thus offering strong support to the hostile media phenomenon hypothesis.[1]

TABLE 1: Perceived Bias as a Function of Question Anchor and Group Membership, Study 1

	Group		
Question Anchor	*SCP*	*Neutral*	*NDP*
SCP	25.9 (21)	27.6 (54)	43.2 (10)
NDP	28.3 (18)	36.4 (51)	45.2 (11)

NOTE: SCP = Social Credit party; NDP = New Democratic party. Higher numbers indicate perceptions of greater bias against the NDP. Possible scores range from 1 to 75. The number of subjects in each group is given in parentheses below the mean.

Effects of question anchor and knowledge. As expected, questions anchored on the SCP garnered ratings of greater bias against the SCP ($M = 28.9$) than questions anchored on the NDP ($M = 35.8$), $F(1, 163) = 13.35$, $p < .001$. Although the Group × Question Anchor interaction fell short of significance, $F(2, 162) = 1.61$, $p = .10$ (one-tailed), planned comparisons revealed a significant effect of question anchor on ratings for the neutral group, $F(1, 103) = 16.65$, $p < .001$, but not for the pro-SCP group, $F(1, 37) = 0.45$, $p = .51$, or the pro-NDP group, $F(1, 19) = 0.13$, $p = .73$.

Next, we divided the subjects into groups above and below the median on knowledge about the issues (one neutral subject was lost because of missing data on the knowledge items). Among the neutral subjects, the expected Question Anchor × Knowledge interaction did indeed appear, $F(1, 103) = 6.74$, $p < .01$. Neutral subjects who were low in knowledge were affected by the question anchor, $F(1, 72) = 25.62$, $p < .001$, whereas highly knowledgeable neutrals were not, $F(1, 31) = 0.01$, $p = .93$ (see Figure 1).[2]

When questions were anchored on the SCP, the low-knowledge neutral subjects were significantly different from the NDP partisans, $F(1, 44) = 25.70$, $p < .001$, but not from the SCP partisans, $F(1, 55) = 0.01$, $p = .93$. But when the questions were anchored on the NDP, the low-knowledge neutral subjects were significantly different from the SCP partisans, $F(1, 52) = 9.43$, $p < .005$, but not from the NDP partisans, $F(1, 45) = 3.22$, $p = .08$. This replicates conceptually the finding by Vallone et al. (1985) of no difference between the low-knowledge neu-

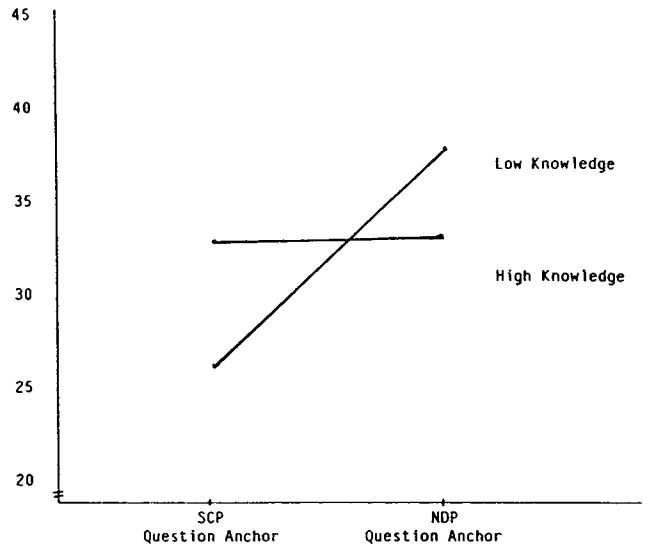

Figure 1 "Perceived bias" scores as a function of question anchor and knowledge for neutral group, Study 1.

trals and the pro-Israel partisans when their questions were anchored on Israel.

Not surprisingly, almost all the partisan subjects fell into the high-knowledge group in the overall sample median split. However, when the partisan subjects were divided into high- and low-knowledge groups according to their own knowledge median, the Question Anchor × Knowledge interaction for the two partisan groups combined was not significant, $F(1, 58) = 1.21$, $p = .28$.

Discussion

The results from Study 1 are consistent with our expectations. Among neutral subjects who were low in issue-relevant knowledge, whichever group was the focus of our bias questions was rated as having been treated more poorly in the newspaper clippings. Thus, we have identified a new question wording effect and two factors that regulate its magnitude, suggesting that the effect is due to a hypothesis confirmation bias.

These results suggest a different explanation than the one offered by Vallone et al. (1985) to account for the fact that their neutral subjects (especially those lowest in relevant knowledge) more closely resembled their pro-Israel subjects than their pro-Arab subjects. Aside from the possibility that their videotaped stimuli were actually biased against Israel, these authors reasoned that American sentiments may have been somewhat pro-Israel and that their "neutral" subjects may have favored Israel more than they realized and reported. Our data suggest that much of the unexpected similarity between their neutral subjects and their pro-Israel subjects was due to the wording of their dependent measures. Because Vallone et al.'s questions focused on Israel, they apparently led low-knowledge, neutral subjects to express pro-Israel sentiments (i.e., media bias against Israel).

Two alternative explanations for the focus of judgment finding in Study 1 deserve comment. First, although we have concentrated on an exclusively cognitive (or intrapsychic) process, it is possible that a more interpersonal process is triggering the effect. For example, because subjects were asked only to evaluate how much the news clippings were biased against *one* of the two political parties, perhaps they assumed that the experimenter was interested merely in this kind of bias. Consequently, in an attempt to respond in a cooperative manner (Grice, 1975), respondents may have based their judgments on a particular subset of the clippings (those particularly critical of the political party in question), thus altering their answers. This account suggests that, rather than confirming a hypothesis, subjects may have been responding to different judgmental tasks conveyed by the two different question wordings. However, if this were the case, there would be no reason to expect that the effect would appear only among low-knowledge, neutral subjects.

Second, although we reasoned that a hypothesis confirmation bias combined with a positivity bias to account for our findings, it is theoretically possible that the assumed positivity bias alone caused the differences in judgments between subjects responding to the NDP- and SCP-anchored questionnaires. However, a straight positivity bias explanation implies that, regardless of whether subjects have information compatible with their judgment of choice, they will endorse such a choice. The hypothesis confirmation bias perspective, however, suggests that the availability of hypothesis-compatible information is critical. That is, the focus of judgment effect would disappear if there were no evidence in the news clippings consistent with the "hypothesis" being tested. As well, the more such evidence there is, the stronger the effect should be. Studies 2 and 3 allow a more detailed examination of this issue.

Finally, it is also useful to note that our results replicated the hostile media phenomenon (Vallone et al., 1985) with a different issue and with a different set of stimuli. Vallone et al. (1985) reported that they (Vallone, Lepper, & Ross, 1981) failed to obtain the phenomenon using written statements, a result that raised questions about the generalizability of their finding. Our replication suggests that the effect is not limited to perceptions of dramatic videotaped news coverage of the Middle East conflict, with its long history and charged emotions. On the contrary, the hostile media phenomenon appears to be a robust and pervasive effect.

STUDY 2

In Study 2, we set out to replicate the focus of judgment effect and to further validate our explanation for it. If we are correct that the effect is the result of a hypothesis confirmation bias in reasoning, two additional factors should regulate its magnitude. First, the effect should be stronger if more information available to an individual is compatible with the hypothesis to be confirmed. That is, no matter how much one may be biased toward confirming a particular hypothesis, one will not be able to do so unless the relevant available information can be interpreted as being consistent with the hypothesis (see, e.g., Kunda, 1987). It seems unlikely, for example, that hypothesis confirmation bias would alter answers to the question "How likely is this object to be a person?" when an individual is looking at a brick. Thus, the first limiting condition involves the amount of hypothesis-compatible information available. More such information should yield a stronger focus of judgment effect.

The second limiting factor should be the degree to which an individual attends to hypothesis-compatible information. When a person makes any given judgment, he or she may have an array of available relevant information, only some of which is compatible with the hypothesis. The more the person's attention is focused on the compatible information, the stronger the focus of judgment effect should be. Conversely, the less attention he or she pays to such information, the weaker the effect should be.

To assess the validity of these hypotheses, we examined a second set of data. These data were originally collected for a study of the effect of information presentation order on social judgments (see Krosnick, Li, & Lehman, 1990, Study 1). After we completed our report of the study, however, we realized that its design allowed us to examine whether a focus of judgment effect occurred and whether its magnitude was regulated by (a) the amount of hypothesis-compatible information available and (b) the degree of attention subjects paid to the hypothesis-compatible information.

This study was originally intended to test for an order effect on base rate and individuating information use. Subjects were given one of various versions of the lawyer/engineer problem, adopted from Kahneman and Tversky (1973) with minor revision. This problem gave subjects individuating information about a person who was randomly selected from a group of 100 lawyers and engineers, and it gave subjects the proportions of lawyers and engineers in the group (i.e., the base rate). Subjects were than asked to assess the likelihood that the selected person was an engineer or a lawyer.

Our study was a $2 \times 2 \times 2 \times 2$ between-subjects design, and the manipulations involved (a) order of information presentation (base rate information first vs. individuating information first), (b) base rate level (70% engineers vs. 30% engineers), (c) individuating information (stereotypical engineer vs. stereotypical lawyer), and (d) question focus (an engineer vs. a lawyer). These manipulations allowed replication of the focus of judgment effect and examination of its limiting conditions, as we explain below.

Method

Subjects. The subjects were 210 undergraduates at the Ohio State University enrolled in an introductory psychology course who participated in the study for course credit. Because one subject did not complete the experimental booklet, 209 subjects were used in the analysis. Subjects were randomly assigned to the 16 conditions.

Materials. Subjects were introduced to the problem with the following sentence: "A panel of psychologists interviewed and administered personality tests to a group of 100 men, some engineers and some lawyers, all successful in their respective fields." The base rate information was "Of the 100 people interviewed in this study 70 (or 30) were engineers and 30 (or 70) were lawyers." The stereotypical lawyer target person description was:

> Tom W. is of high intelligence, is quite self-confident, and tends to be argumentative, even with people he doesn't know very well. He is very involved in his work, and tends to work long hours. He is generally well dressed, even when not at work. His writing is interesting and creative, and is usually very convincing. He has a strong drive for competence and is rather competitive with others in his field. He is interested in social issues and reads the newspaper daily. He drives a sports car and lives in a suburban upper-middle-class neighborhood. He is not particularly uncomfortable if he lies to someone.

The stereotypical engineer target person description was:

> Tom W. is a 45-year-old man. He is married and has four children. He is generally conservative, careful, and ambitious. He shows no interest in political and social issues and spends most of his free time on his many hobbies, which include home carpentry, sailing, and mathematical puzzles.

After exposure to the base rate and individuating information, subjects were asked: "On a scale from 0 to 100, how likely do you think it is that Tom W. is (an engineer/a lawyer)? _____%"[3]

Procedure. Subjects were run in groups of approximately 35. Each subject received a booklet that presented each piece of information on a separate page so as to

control order of information exposure. Subjects were given 30 s to read each page and were told not to turn each page until they were instructed to do so. They were also required not to turn back to previous pages. After reading the two pieces of information, subjects estimated the probability that Tom W. was a lawyer or an engineer.

Predictions

According to the focus of judgment hypothesis, subjects asked how likely Tom W. is to be an engineer should give a higher mean probability than the implied mean probability that he is an engineer provided by subjects who were asked how likely he is to be a lawyer. This main effect of question wording would support our focus of judgment hypothesis.

The manipulation of the individuating information allowed us to examine whether the magnitude of the focus of judgment effect varies with the amount of hypothesis-compatible information available. The stereotypical lawyer paragraph contains many details (approximately 15) that could be viewed as consistent with being a lawyer, whereas the stereotypical engineer paragraph is shorter and contains fewer pieces of information (approximately 7) that could be interpreted as sounding like an engineer. Consequently, we predicted that it would be easier for subjects to confirm the lawyer hypothesis with the stereotypical lawyer paragraph than to confirm the engineer hypothesis with the stereotypical engineer paragraph. Therefore, the question wording effect should be stronger when the stereotypical lawyer paragraph is presented.

We were also able to test our hypothesis regarding attention to hypothesis-compatible information. Our prior research indicated that whichever piece of information (base rate or individuating information) is presented last carries special weight in subjects' thinking and judgments (Krosnick et al., 1990) Furthermore, it seems reasonable to assume that the individuating information is highly susceptible to hypothesis-confirmation-biased evaluation whereas the base rate is not. That is, it is difficult to imagine how a base rate stating, for example, that a sample included 30% lawyers could be subjected to biased interpretation as easily as a paragraph of individuating characteristics. We therefore expected that the focus of judgment effect might be stronger when the individuating information is presented last than when the base rate is presented last. If this prediction is supported, it would bolster the claim that hypothesis confirmation bias is stronger when an individual pays close attention to hypothesis-compatible information than when he or she does not.

In sum, we expected to observe three effects: (a) a main effect of question wording on judgments, indicating a focus of judgment effect, (b) an interaction between question wording and individuating information, with a larger question wording effect for the stereotypical lawyer description than for the stereotypical engineer description, and (c) an interaction between question wording and presentation order, with a larger question wording effect when the individuating information was presented last.

Results

Consistent with the focus of judgment hypothesis, the manipulation of question wording had a significant effect on subjects' judgments, $F(1, 193) = 7.70$, $p < .007$. Subjects' mean estimate of the probability that the target person was an engineer was 55.7% when the question asked about whether he was an engineer. After subtraction from 100, the same probability was 45.6% when the question asked about whether he was a lawyer.

Also as expected, the Individuating Information × Question Wording interaction was significant, $F(1, 193) = 5.67$, $p < .02$. The effect of question wording was significant only when the stereotypical lawyer paragraph was used (M difference = 19.48), $t(106) = 3.88$, $p = .0002$, and not when the stereotypical engineer paragraph was used (M difference = 1.30), $t(99) = 0.23$, n.s. The Presentation Order × Question Wording interaction was also significant, $F(1, 193) = 4.25$, $p < .05$. The question wording effect was significant only when the individuating information was presented last (M difference = 16.55), $t(103) = 2.80$, $p = .006$, and not when the base rate information was presented last (M difference = 3.57), $t(102) = 0.60$, n.s.

The Individuating Information × Presentation Order × Question Wording interaction was not significant, $F(1, 193) = 0.15$, n.s., suggesting that individuating information and presentation order combined additively to determine the magnitude of the question wording effect. When the lawyer description was used and it was presented last, the question wording effect was large (M difference = 24.97), $t(54) = 3.83$, $p = .0003$. When the engineer description was used and it was presented first, the question wording effect disappeared completely and even reversed slightly (M difference = –6.43), $t(50) = 1.75$, $p = .086$. And when either the lawyer description was presented first or the engineer description was presented last, the question wording effect was moderate in size (M differences = 10.23, $t[47] = 1.33$, $p = .19$, and 13.58, $t[50] = 0.80$, n.s., respectively).

STUDY 3

Although the results of Study 2 conformed to our expectations, the interaction pattern predicted and obtained seems complex enough to warrant replication. A third study was therefore conducted using a reduced

version of the Study 2 design with a within-subjects component added.

In this study, no base rates were presented. All subjects read both the stereotypical lawyer paragraph and the stereotypical engineer paragraph used in Study 2. Half the subjects read the lawyer paragraph first, and half read the engineer paragraph first. After reading each paragraph, half the subjects were asked how likely that individual was to be a lawyer, and the other half were asked how likely the individual was to be an engineer. Thus, this study was a 2 (Individuating Information, within-subjects) × 2 Presentation Order, between-subjects) × 2 (Question Wording, between-subjects) design.

Method

Subjects. Subjects were 135 undergraduates at the Ohio State University enrolled in an introductory psychology course who participated in the study for course credit. Subjects were randomly assigned to the four between-subjects conditions.

Materials. Subjects were simply told to read each paragraph and to make the required judgment about each. The paragraphs were identical to those used in Study 2, except that the stereotypical lawyer was always referred to as Bill F. and the stereotypical engineer was always referred to as Tom W. After each paragraph, subjects were asked: "How likely is (Bill F./Tom W.) to be a (lawyer/engineer)?" Answers were made on 10-point scales ranging from *not likely at all* to *extremely likely.*

Procedure. Subjects were run individually. Each subject received a booklet that presented all paragraphs and questions on the same page. Subjects were given up to 5 min to complete the questionnaire.

Predictions

According to the focus of judgment hypothesis, subjects who are asked how likely it is that Tom W. and Bill F. are engineers should give higher ratings than the implied mean rating (after scale reversal) provided by subjects asked how likely they are to be lawyers.[4] This main effect of question wording would further strengthen our focus of judgment hypothesis. Moreover, given that the longer stereotypical lawyer paragraph provides more opportunity for hypothesis confirmation through biased information processing, the question wording effect should be stronger for that paragraph than for the stereotypical engineer paragraph.

Results

As expected, the question wording manipulation had a substantial effect on ratings. The mean rating made by subjects asked how likely Bill F. and Tom W. were to be engineers was 5.79, whereas the mean rating made by subjects asked how likely they were to be lawyers (after reversing the coding of the scale) was 3.28. Thus, the effect of question wording was 2.51 scale units, $F(1, 131) = 145.77$, $p < .0001$. Also as expected, this effect of question wording interacted with the content of the paragraph, $F(1, 131) = 37.21$, $p < .0001$. The effect was larger for the stereotypical lawyer paragraph, 4.29 units, $F(1, 134) = 132.87$, $p < .0001$, than for the stereotypical engineer paragraph, 0.72 unit, $F(1, 134) = 4.32$, $p = .04$. It again appears that the question wording effect was stronger when more information was available that could be interpreted in a hypothesis-consistent fashion.

Unexpectedly, one other significant effect appeared in our ANOVA: an interaction of question wording with presentation order, $F(1, 131) = 4.69$, $p = .03$. The question wording effect was larger when the stereotypical lawyer paragraph was judged before the stereotypical engineer paragraph (2.97 units) than when the stereotypical engineer paragraph was judged before the stereotypical lawyer paragraph (2.07 units). Presumably, successfully interpreting the first paragraph as confirming a hypothesis enhanced the likelihood that subjects would attempt to confirm that same hypothesis in evaluating the second paragraph. In contrast, finding the first paragraph more difficult to interpret in a way consistent with the hypothesis presumably decreased the likelihood that subjects would attempt to confirm that same hypothesis when confronting the second paragraph.

Limitation of Studies 2 and 3

Although the differential length of the lawyer and engineer descriptions in Studies 2 and 3 predicted the strength of the focus of judgment effect, it is important to note that length was not independently manipulated. Consequently, other factors, such as differential prototypicality or specific wording subsets, may have caused (or at least contributed to) the magnitude of the focus of judgment effect. As a result, the findings on length should be treated as preliminary; more direct tests would be useful in the future.

GENERAL DISCUSSION

The findings of these three studies contribute to the questionnaire design literature in several ways. First, the focus of judgment effect that we identified expands a growing list of other context effects that appear to be due to a bias toward hypothesis confirmation. Along with acquiescence response bias (Schuman & Presser, 1981) and response order effects (Krosnick & Alwin, 1987), the focus of judgment effect is apparently due to a tendency

for subjects to confirm any hypothesis offered directly by a question or stimulated indirectly (for example, by way of a positivity bias).

Second, we demonstrated that knowledge level determines susceptibility to the effect. Lack of knowledge alone, however, was not sufficient for the focus of judgment effect to appear in our first study. Rather, it was also necessary for subjects low in knowledge to be neutral with regard to the controversy in question. Low-knowledge subjects who had a priori biases in one directior or the other showed no focus of judgment effect. Thus, knowledge regulates the magnitude of the effect by interacting with a priori bias.

Our results contribute to the questionnaire design literature in a third way as well: by documenting an interesting question order effect. In Study 3, our question wording manipulation was stronger for the second set of individuating information when the first set was more compatible with the hypothesis suggested by the question. Although the literature on question order effects is quite extensive (see Tourangeau & Rasinski, 1988), it has not yet included any indication that successful hypothesis confirmation in answering one question is associated with a higher likelihood that respondents will be biased toward that same hypothesis in answering subsequent questions. Thus, our evidence expands this body of empirical findings.

As we indicated above, there are two plausible explanations for the focus of judgment effects we observed: one based on hypothesis confirmation bias and the other based on beliefs about conversational conventions. We designed our experiments to test the validity of the first explanation by identifying limiting conditions of the effect. As we expected, the effect was indeed most likely to appear among low-knowledge, neutral subjects when they paid close attention to a large amount of hypothesis-compatible information. We find it difficult to see how the conversational conventions explanation would predict these limiting conditions. Therefore, we view our evidence as offering more support for the hypothesis confirmation bias explanation than for the conversational conventions explanation. Nonetheless, we look forward to more direct tests of both explanations.

If our results are indeed due to hypothesis confirmation bias, they contribute to the literature on that phenomenon (see Fischhoff & Beyth-Maron, 1983; Klayman & Ha, 1987) in a number of ways. Most of the prior literature on hypothesis confirmation bias has focused either on hypotheses that subjects bring to a judgment task on the basis of motivations such as self-enhancement or self-protection or on hypotheses provided in a very compelling manner by an experimenter (see, e.g., Sanitioso et al., 1990). Thus, these past studies suggest that powerful and emotion-laden hypotheses can produce observable effects on reasoning. Our work adds to the existing psycholinguistic literature on the nature of presuppositions by suggesting a more subtle effect: Hypotheses provided simply by the phrasing of a question may direct cognitive processing. Our evidence also suggests that the effect is observed only when subjects lack relevant knowledge and a priori bias, when a large amount of hypothesis-compatible information is available to them, when they pay close attention to hypothesis-compatible information, and when they have already had a successful experience confirming the hypothesis.

NOTES

1. The mean rating of perceived bias by the neutral subjects indicates that the package of newspaper clippings was most likely slightly biased against the SCP: The neutral group's mean rating (31.9) was significantly different from the scale midpoint (37.5), $t(104) = 3.57$, $p < .001$.

2. We also performed a median split on knowledge *within* the neutral group and ran the identical Question Anchor by Knowledge analysis, which revealed the same interaction, $F(1, 100) = 4.18$, $p < .05$.

3. Estimates made by subjects whose prediction question asked about the probability that Tom W. was a lawyer were subtracted from 100 so that they would be on the same scale as the predictions made by subjects whose prediction question asked about the probability that Tom W. was an engineer.

4. In contrast to Study 2 (where subjects were told the target person was either a lawyer or an engineer), the mean rating after scale reversal does not refer to the likelihood that the target person is an engineer. Rather, it refers to the likelihood that he is an engineer or any other occupation except a lawyer. Thus, it represents an upper bound on the likelihood that he is an engineer. To the extent that the actual value is less than the upper bound, the focus of judgment effect of interest here is even stronger than it appears in the analyses we report.

REFERENCES

Anderson, C. A., Lepper, M. R., & Ross, L. (1980). Perseverance of social theories: The role of explanation in the persistence of discredited information. *Journal of Personality and Social Psychology, 39,* 1037-1049.

Chapman, L. J., & Chapman, J. P. (1969). Illusory correlation as an obstacle to the use of valid psychodiagnostic signs. *Journal of Abnormal Psychology, 74,* 271-280.

Clark, H. H. (1985). Language use and language users. In G. Lindzey & E. Aronson (Eds.), *Handbook of social psychology* (pp. 179-231). New York: Random House.

Clark, H. H., & Clark, E. V. (1977). *Psychology and language.* San Diego, CA: Harcourt Brace Jovanovich.

Fischhoff, B., & Beyth-Maron, R. (1983). Hypothesis evaluation from a Bayesian perspective. *Psychological Review, 90,* 239-260.

Grice, H. P. (1975). Logic and conversation. In P. Cole & J. L. Morgan (Eds.), *Syntax and semantics 3: Speech acts* (pp. 41-58). New York: Academic Press.

Hamilton, D. L. (1979). A cognitive attributional analysis of stereotyping. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 12, pp. 53-84). New York: Academic Press.

Hilton, D. J. (1990). Conversational processes and causal explanation. *Psychological Bulletin, 107,* 65-81.

Jennings, D., Amabile, T. M., & Ross, L. (1980). Informal covariation assessment: Data-based vs. theory-based judgments. In A. Tversky, D. Kahneman, & P. Slovic (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 211-230). New York: Cambridge University Press.

Kahneman, D., & Tversky, A. (1973). On the psychology of prediction. *Psychological Review, 80,* 237-251.

Klayman, J., & Ha, Y. (1987). Confirmation, disconfirmation, and information in hypothesis testing. *Psychological Review, 94*, 211-228.

Krosnick, J. A. (1991). Response strategies for coping with the cognitive demands of attitude measures in surveys. *Applied Cognitive Psychology, 5*, 213-236.

Krosnick, J. A., & Alwin, D. F. (1987). An evaluation of a cognitive theory of response-order effects in survey measurement. *Public Opinion Quarterly, 51*, 201-219.

Krosnick, J. A., Li, F., & Lehman, D. R. (1990). Conversational conventions, order of information acquisition, and the effect of base rates and individuating information on social judgments. *Journal of Personality and Social Psychology, 59*, 1140-1152.

Kunda, Z. (1987). Motivated inference: Self-serving generation and evaluation of causal theories. *Journal of Personality and Social Psychology, 53*, 636-647.

Levinson, S. C. (1983). *Pragmatics*. Cambridge: Cambridge University Press.

McClendon, M. J. (1991). Acquiescence and recency response-order effects in interview surveys. *Sociological Methods and Research, 20*, 60-103.

Ross, L. (1977). The intuitive psychologist and his shortcomings: Distortions in the attribution process. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 10, pp. 173-220). New York: Academic Press.

Ross, L., Lepper, M. R., & Hubbard, M. (1975). Perseverance in self perception and social perception: Biased attributional processes in the debriefing paradigm. *Journal of Personality and Social Psychology, 32*, 880-892.

Sanitioso, R., Kunda, Z., & Fong, G. T. (1990). Motivated recruitment of autobiographical memories. *Journal of Personality and Social Psychology, 59*, 229-241.

Schuman, H., & Presser, S. (1981). *Questions and answers in attitude surveys*. New York: Academic Press.

Sears, D. O. (1983). The person-positivity bias. *Journal of Personality and Social Psychology, 44*, 233-250.

Skov, R. B., & Sherman, S. J. (1986). Information-gathering processes: Diagnosticity, hypothesis-confirmatory strategies, and perceived hypothesis confirmation. *Journal of Experimental Social Psychology, 22*, 93-121.

Snyder, M. (1981). Seek and ye shall find: Testing hypotheses about other people. In E. T. Higgins, C. P. Herman, & M. P. Zanna (Eds.), *Social cognition: The Ontario Symposium on Personality and Social Psychology* (pp. 277-303). Hillsdale, NJ: Lawrence Erlbaum.

Snyder, M., & Campbell, B. H. (1980). Testing hypotheses about other people: The role of the hypothesis. *Personality and Social Psychology Bulletin, 6*, 421-426.

Snyder, M., & Swann, W. B. (1978). Hypothesis-testing in social interaction. *Journal of Personality and Social Psychology, 36*, 1202-1212.

Tourangeau, R., & Rasinski, K. A. (1988). Cognitive processes underlying context effects in attitude measurement. *Psychological Bulletin, 103*, 299-314.

Vallone, R. P., Lepper, M. R., & Ross, L. (1981). *Perceptions of media bias in the 1980 presidential election*. Unpublished manuscript, Stanford University.

Vallone, R. P., Ross, L., & Lepper, M. R. (1985). The hostile media phenomenon: Biased perception and perceptions of media bias in coverage of the Beirut massacre. *Journal of Personality and Social Psychology, 49*, 577-585.

Wason, P. C. (1960). On the failure to eliminate hypotheses in a conceptual task. *Quarterly Journal of Experimental Psychology, 12*, 129-140.

Received March 11, 1991
Revision received November 3, 1991
Accepted November 6, 1991