

## Accountability: A Social Magnifier of the Dilution Effect

Philip E. Tetlock  
University of California, Berkeley

Richard Boettger  
Texas Christian University

This research demonstrated that accountability can not only reduce judgmental bias, but also exacerbate it—in this case, the dilution effect. Ss made predictions from either diagnostic information alone or diagnostic information plus mixtures of additional data (nondiagnostic information, additional diagnostic data pointing to either the same conclusion or the opposite conclusion). Relative to unaccountable Ss, accountable Ss (a) diluted their predictions in response to nondiagnostic information and (b) were more responsive to additional diagnostic information. The accountability manipulation motivated subjects to use a wide range of information in making judgments, but did not make them more discriminating judges of the usefulness of that information.

Cognitive social psychologists have painted numerous portraits of the person as information processor. Early work emphasized the rigorous rationality with which people analyzed and drew inferences from evidence: the correspondent inference model, the causal schemata model, the covariation model, and the Bayesian model. Later work emphasized people's judgmental shortcomings. People were depicted as cognitive misers whose preference for simple, easy-to-execute heuristics rendered them vulnerable to a variety of errors and biases (Abelson & Levi, 1985; Fiske & Taylor, 1984; Markus & Zajonc, 1985). People in this view are too quick to draw conclusions about the personalities of others from fragmentary evidence (Jones, 1979), too slow to revise their impressions in response to contradictory evidence (Nisbett & Ross, 1980), and too confident in their factual judgments and predictions (Fischhoff, 1982).

Although the cognitive-miser metaphor highlights important facts about social judgment, there is now growing recognition within the field that the metaphor has also proven misleading. People often use different information-processing strategies in different situations. Much depends on the perceiver's goals and the normative context (cf. Borgida & Howard-Pitney, 1983; Chaiken, 1980; Erber & Fiske, 1984; Kraut & Higgins, 1984; Payne, 1982; Petty, Cacioppo, & Goldman, 1981; Showers & Cantor, 1985; Tetlock, 1985a). To paraphrase Jenkins's (1981) contextualist critique of cognitive psychology, the appropriate question is not "What kind of machine is the human information processor?" but rather "What kinds of machines do people become when confronted with particular tasks in particular environments?" People are cognitive misers much of the time, but not all of the time. The cognitive-miser metaphor needs to be qualified to take into account wide situational and individual-difference variation in patterns or styles of social reasoning.

Recent research points to one such set of qualifications. Under certain conditions, accountability—pressure to justify one's views to others—motivates people to process information in

more complex and multidimensional ways. When people expect to justify their views to an audience with unknown opinions and learn of being so accountable prior to exposure to the relevant evidence, they often engage in "preemptive self-criticism" (Tetlock, 1983a; Tetlock, Skitka, & Boettger, in press). They analyze the evidence especially carefully and pay particular attention to inconsistent information. In effect, people try to prepare themselves for potential objections that a well-informed audience might raise to the stands that they have taken.

A series of studies has shown not only that predecisional accountability to unknown audiences leads to more complex thought (Tetlock, 1983a), but also that this increased complexity confers resistance to well-known judgmental biases. Tetlock (1983b) found, for example, that accountability both improved recall of evidence and eliminated primacy effects in a legal decision-making task. Tetlock (1985b) found that accountability reduced the overattribution effect in an essay-attribution paradigm. Tetlock and Kim (1987) found that accountability led to more realistic levels of confidence in subjects' predictions of how people had responded to a personality test. They also discovered that accountability increased the accuracy of predictions and that increased complexity of thought mediated improvements in both accuracy and appropriateness-of-confidence ratings.

The "de-biasing" studies to date have all focused on experimental paradigms in which more complex styles of thinking are likely to reduce inferential bias. More complex information processors are less likely to allow their initial impressions of events to dominate how they interpret later evidence, less likely to ignore situational constraints on an actor whose behavior they expect to explain, and less likely to ignore inconsistencies in the responses of people whose behavior they expect to predict. The same underlying processes that improved judgment in the de-biasing studies may, however, exacerbate bias in other settings. Encouraging people to take more information into account is not always a good idea. In part, this is so because of the limited time and effort people can devote to understanding specific events or issues (e.g., Simon, 1957). This is also so, in part, because people sometimes confront environments in which there are only one or two valid cues for making judgments. These diagnostic cues are, however, embedded in a bewildering array of irrelevant variables. Experimental manipula-

---

This research was supported by National Institute of Mental Health Grant 39942 to Philip E. Tetlock.

Correspondence concerning this article should be addressed to Philip E. Tetlock, Institute of Personality Assessment and Research, 3657 Tolman Hall, University of California, Berkeley, California 94720.

tions that induce complex thinking will dilute the impact of the truly diagnostic variable by encouraging people to integrate information from multiple sources prior to making a judgment.

The present study systematically explores this possibility. We examine the impact of accountability on the dilution effect: the tendency for nondiagnostic information to dilute the extremity of predictions that people make when presented only with diagnostic information. In a compelling series of studies, Nisbett, Zukier, and Lemley (1981) and Zukier (1982) demonstrated that linking diagnostic with nondiagnostic information produced more regressive predictions than people would otherwise have made (for an important nonsocial precursor to this work, see Troutman & Shanteau, 1977). For example, people drew strong inferences about the likely grade point average (GPA) of a student given the knowledge that he studied either 3 hr or 31 hr per week. However, people moderated their predictions, sometimes dramatically, on receiving information that pretest subjects felt was completely irrelevant (e.g., the number of plants the student kept).

Nisbett et al. (1981) and Zukier (1982) have established the replicability of the dilution effect and explained the phenomenon by invoking the representativeness heuristic of Kahneman and Tversky (1973). People, in this view, judge whether an individual will perform an action by comparing key features of the individual with key features of the possible outcomes and predicting the outcome most similar to or representative of the individual. Mathematically, similarity between the individual and the outcome is a positive function of the number of common features and a negative function of the number of unique features (cf. Tversky, 1977). Common features are attributes of the individual that people frequently associate with the outcome. For instance, the image of a student who studies only 3 hr per week is strongly associated with the outcome of a low GPA. Noncommon features are attributes of the target that people rarely associate with the outcome. Thus, one reduces the similarity—and hence the perceived predictive link—between the indolent student and poor grades by including irrelevant details in the description of the student (e.g., tennis player, keeper of plants).

The present study builds on this analytical framework. The central hypothesis is that accountability—far from functioning as a de-biasing manipulation—will magnify the dilution effect. A second, key hypothesis is that the same mechanism that explains the de-biasing effectiveness of accountability in other contexts will be responsible for the magnification of the dilution effect here, namely increased complexity of thought. The more complex people try to be in the dilution conditions of the study, the more regressive their predictions will be.

To test this latter hypothesis, we measured not only the predictions subjects made, but also the impressions that subjects formed of the underlying evidence. As a process measure, we obtained thought protocols by asking subjects to report their thoughts about the evidence and by then subjecting those thoughts to systematic content and structural analysis (cf. Ericsson & Simon, 1980; Tetlock & Kim, 1987). In particular, we subjected the thoughts reported to integrative complexity coding, a technique that has proven useful in previous laboratory and archival research and that is described in detail in the Method section. Integrative complexity is measured on a 7-point scale in which low scores reflect the absence of both con-

ceptual differentiation and integration; moderate scores reflect the presence of conceptual differentiation, but the absence of conceptual integration; and high scores reflect the presence of both differentiation and integration. We expected that integratively complex thinkers would be more susceptible to the dilution effect and that statistically controlling for the complexity of the thought protocols would attenuate both the dilution effect and the tendency for accountability to exacerbate that effect.

The present study also allowed us to test whether accountability exacerbates the dilution effect through a very different psychological mechanism. Accountability demands may not so much affect how people think as what they are willing to say. Accountability may induce a cautious response set—a generalized unwillingness to make extreme predictions that might prove difficult to justify. Accountability, in this view, turns people into fence-sitters who rarely stray from the safe midpoints of the scales on which they express judgments.

The study reported here allowed us to disentangle the cognitive-motivational and response-bias interpretations. The cognitive-motivational explanation implies greater selectivity in the effects of accountability. If accountability simply induces a generalized reluctance to make extreme predictions about future states of the world, we should observe more regressive predictions in both control conditions (where only diagnostic information is presented) and dilution conditions. If, on the other hand, accountability motivates subjects to process carefully all information at their disposal and to integrate that information into a defensible judgment, we should expect more regressive predictions among accountable subjects in the dilution than in the control conditions (for it is only in the dilution conditions that subjects receive nondiagnostic information). In brief, the cognitive-motivational interpretation predicts an interaction in which dilution effects are particularly pronounced in the accountability conditions. The response-bias interpretation predicts a main-effect tendency for accountability to produce less extreme predictions.

These two interpretations also lead to different predictions in some additional experimental conditions included in the design. For instance, in one set of conditions, we augmented the diagnostic information and gave additional reasons for expecting a particular outcome (augmented-information conditions). From a response-bias perspective, accountable subjects in these conditions should make predictions closer to the midpoints of the scale than should unaccountable subjects. From a cognitive-motivational perspective, accountable subjects should make predictions as extreme as do their unaccountable counterparts. Indeed, accountable subjects may be even more extreme because they are more likely to have carefully processed the additional reasons supportive of an extreme prediction. In Tesser's (1978) terms, accountable subjects may be more prone to polarize their predictions in response to multiple items of diagnostic evidence for exactly the same reason that they are more prone to dilute their predictions in response to multiple items of nondiagnostic evidence: namely, their tendency to give more thought to the evidence.

In a final set of conditions, we gave people reasons for expecting opposite outcomes (contradictory-information conditions). Here uncertainty is normatively justifiable. The cognitive-motivational and response-bias interpretations thus converge on a

common prediction: moderation of judgment. We included these conditions for a different theoretical purpose. We were interested in whether accountable subjects would dilute their predictions to the same degree regardless of the diagnosticity of the evidence. The hypothesis that accountability exacerbates the dilution effect assumes that accountable subjects become more integratively complex even when the evidence is of minimal probative value. There is no reason, from this standpoint, to expect large differences in the regressiveness of the predictions that subjects make in the dilution (irrelevant) and contradictory-information (relevant) conditions. It is possible, however, that accountability transforms subjects into more discriminating "consumers" of the evidence who display greater complexity only when the evidence appears relevant to the judgment task. If this interpretation is correct, accountability should both reduce the dilution effect and magnify the differences between the dilution and contradictory-information conditions. Accountable subjects should make more extreme predictions in the dilution conditions and less extreme predictions in the contradictory-information conditions.

## Method

### Subjects

A total of 160 undergraduate subjects participated in return for course credit or money (\$4). Subjects were randomly assigned to conditions in two 2 (accountability)  $\times$  2 (direction of diagnostic evidence)  $\times$  4 (amount or types of evidence) between-subjects designs (one experiment on the prediction of GPA, the other on the prediction of child abuse), with 10 subjects in each condition. Subjects were randomly assigned to conditions in small group testing sessions.

### Procedure

The experimenter informed subjects that they would be participating in a study of person perception processes, that is, a study of how people both form impressions of others from various types of information and use those impressions in making predictions. The Scenario 1 materials described a student (Robert) and included a thought-list page on which subjects reported their impressions of the student as well as a forecasting page on which subjects predicted the GPA of the student and rated their confidence in their prediction. Scenario 2 materials included a description of a psychotherapy patient (David), a thought-list page on which subjects reported their impressions, and a forecasting page on which subjects predicted the likelihood that David was a child abuser and rated their confidence in their predictions.

Subjects in the control conditions of Scenario 1 received only information about the number of hours per week the student studied, either 3 or 31 hr. Subjects in the control conditions of Scenario 2 received only two items of information. They learned either that the patient was sexually aroused by violent sadomasochistic fantasies and had a serious drinking problem or that the patient wanted to adopt a second child and did volunteer work at a neighborhood school to promote good race relations.

Subjects in the dilution conditions received the control-diagnostic information plus a variety of nondiagnostic information. Examples of nondiagnostic information for Scenario 1 included "Robert is widely regarded by his friends as being honest," "Robert plays tennis or racquetball about three or four times a month," "Robert describes himself as a cheerful person," and "Two months is the longest period of time Robert has dated one person"; for Scenario 2, nondiagnostic information included "David has an IQ of 110," "He injured his back in a skiing

accident," "David is strong-minded and rarely willing to back off on an issue of principle," and "David likes to tell jokes."

Subjects in the augmented-information conditions of Scenario 1 received additional reasons for believing that the student Robert would do well or poorly in school. Subjects who learned that Robert studied 31 hr per week also learned that "Robert finds most of the college courses he has taken to be extremely interesting" and that "Robert derives great enjoyment from intellectual activities." Subjects who learned that Robert studied only 3 hr per week learned that "Robert finds most of the college courses he has taken boring" and "Robert does not want a high-powered career if it means giving up a relaxed, pleasant life-style." Subjects in the augmented-information conditions of Scenario 2 received additional reasons for believing that David was or was not likely to be a child abuser. Subjects who received evidence suggestive of child abuse also learned that David had been sexually assaulted by his stepfather and had no friends. Subjects who had received counterdiagnostic evidence also learned that "David has many close friends" and "David has a relaxed and easy-going style of dealing with other people."

Subjects in the contradictory-information conditions received a mixture of diagnostic and counterdiagnostic information. Thus, those in Scenario 1 who believed that Robert studied 31 hrs per week also learned that he found most college courses boring and disliked intellectual activities; those who believed Robert studied only 3 hr per week learned that he found most courses extremely interesting and very much enjoyed intellectual activities. Subjects in Scenario 2 who initially received evidence diagnostic of child abuse also received some counterdiagnostic evidence (has many close friends, does volunteer work at the neighborhood school to promote good race relations). Conversely, subjects in Scenario 2 who initially received counterdiagnostic evidence also received diagnostic evidence (sexually assaulted by his stepfather, has no friends).

After reading each vignette, subjects were given 5 min to think about the evidence and to report their thoughts. After reporting their thoughts about Robert (or about David), subjects predicted his GPA (or likelihood of being a child abuser) and assessed their confidence in their prediction.

Subjects were explicitly given the option of not using the information provided to make predictions. Subjects were told that they might or might not find the information useful in making their predictions. We simply wanted subjects to provide us with their best guess about Robert's academic performance and the likelihood of David's being a child abuser. In the case of Robert, subjects were told that if they felt they had no useful information, they should simply predict an average GPA for Berkeley, approximately 3.0. In the case of David, subjects were told that if they felt they had no useful information, they should simply check the midpoint of the 11-point likelihood scale (6, no greater or less likelihood of this person being a child abuser than anyone else).

### Pretest Procedures for Selecting Stimulus Materials

The pretest procedures were based on those of Nisbett et al. (1981). Forty-five undergraduate students in a social psychology course rated the usefulness of information for predicting "certain characteristics of other people." We screened 50 items of information for the GPA prediction task and 54 items for the child-abuse task. In a blank next to each item, subjects rated the items as diagnostic (+ means that the information, taken by itself, suggests that the person is either likely to have a very high GPA in one task or to be a child abuser in the other task), counterdiagnostic (– means that the information, taken by itself, suggests the person is likely to have a very low GPA or not to be a child abuser), or nondiagnostic (0 means that the information, taken by itself, is of no value in deciding whether the person has a low or high GPA or whether the person is a child abuser).

For use in the experiment, we selected six sets of four items, those most consistently rated as diagnostic of child abuse, counterdiagnostic of child abuse, and nondiagnostic of child abuse, and those most consis-

tently rated as diagnostic of a high GPA, counterdiagnostic of a high GPA, and nondiagnostic of GPA. To be included in the nondiagnostic item set, at least 80% of the pretest subjects had to judge the items as being of no value. To be included in the diagnostic or counterdiagnostic item set, at least 80% of the pretest subjects had to judge the items as having predictive value in the specified direction. When all other things were approximately equal, we selected items used by Nisbett et al. (1981) and Zukier (1982).

We also took additional precautions. Following Zukier (1982), we checked whether individual items rated as nondiagnostic were still perceived that way when presented together. We asked 15 subjects whether a person described by the combined set of nondiagnostic items selected for the experiment was likely to have a GPA higher or lower than the Berkeley average or to be more or less likely than others to be a child abuser. The results were clear-cut. Items rated as nondiagnostic by themselves continued to be rated as nondiagnostic together. The dilution effects reported later cannot be attributed to subtle interactive causal relations among the items selected here.

Finally, we checked that the items classified as nondiagnostic were truly seen as nondiagnostic and not as diagnostic of a midpoint value on the prediction scales to which our experimental subjects responded. An independent group of 40 subjects rated 25 GPA items and 25 child-abuse items (sets that included all the experimental items), not only for whether the items were diagnostic, counterdiagnostic, or nondiagnostic, but also for whether they were diagnostic of the mean (indicative of an average GPA or likelihood of child abuse). Again, the results were clear-cut. Most subjects (more than 80%) continued to rate each of the nondiagnostic items as nondiagnostic as opposed to diagnostic of the average. This was true, moreover, regardless of whether we used a forced-choice rating procedure (choose one of four categories) or continuous rating scales identical to those used in the experiments (rate GPA on a magnitude estimation scale ranging from 2.0 to 4.0, with 3.0 as the average value, or likelihood of child abuse on a scale ranging from 1 to 11, with 6 as the average value). Finally, the reluctance to rate nondiagnostic items as diagnostic of midpoint values cannot be attributed to a general reluctance to use the midpoint values. More than half of the raters judged 15 of the 50 items to be diagnostic of the mean.

### *Accountability Manipulation*

Subjects received one of two sets of accountability instructions. Subjects in the no-accountability control conditions were assured prior to receiving any information that their impressions of the test-takers would be completely confidential and not traceable to them personally. Even the experimenter, it was emphasized, would not know how they had responded. Subjects in the accountability conditions learned prior to receiving any evidence that the researchers would later conduct interviews with subjects to explore the types of information that people use to form impressions of others. These subjects expected to participate in such an interview after they had completed the experimental tasks and were asked to sign a form granting permission to audiotape the interview for future data-analytic purposes.

### *Integrative Complexity*

Two trained coders rated the integrative complexity of the free-response personality sketches that subjects provided of both Robert and David. The coding system used for this purpose has demonstrated reliability and construct validity. It has been successfully applied in numerous research contexts to test hypotheses concerning both personality and situational determinants of complexity of information processing (see Schroder, Driver, & Streufert, 1967; Streufert & Streufert, 1978; Tetlock, 1979, 1981, 1983a, 1984, 1986a, 1986b; Tetlock, Hannum, & Micheletti, 1984).

Integrative complexity is defined in terms of both conceptual differentiation and integration. *Differentiation* refers to the number of alter-

native interpretations that a person considers in analyzing an event or issue. For instance, a subject might take an undifferentiated view of a stimulus person by focusing on only one major theme running through the evidence (e.g., this person does nothing but study). A more differentiated statement would recognize either contradictory evidence on the person's standing on a trait dimension (e.g., this person is hardworking, but only when a subject catches his interest) or the difficulty of capturing the complexity of personality with a single trait label (e.g., on the one hand, this person seems ambitious; on the other hand, the person lacks an overall purpose in life). *Integration* refers to the development of complex connections among differentiated characteristics. The complexity of integration depends on whether the person perceives the differentiated characteristics as existing in isolation (low integration), in simple interactions (moderate integration), or in multiple, contingent patterns (high integration). For example, moderate-integration statements might specify how two traits interact to shape behavior (e.g., This person is both talented and enjoys his work. When you have this combination, the person is likely to be a star performer).

Integrative complexity scores ranged from 1 to 7 (1 = low differentiation and integration; 3 = moderate-high differentiation, low integration; 5 = moderate-high differentiation, moderate integration; 7 = high differentiation and high integration). Scores of 2, 4, and 6 represent transition levels that are assigned when there is evidence of implicit differentiation (e.g., use of qualifiers, recognition of uncertainty, information seeking) or implicit integration (e.g., allusions to interactions between different personality attributes). High interrater agreement existed between the two coders ( $r = .89$ ), one of whom was unaware of both the experimental design and the hypotheses.

## Results

For each scenario, there were three dependent measures: (a) the predicted GPA or likelihood of child abuse, (b) the integrative complexity of the thoughts reported, and (c) subjects' confidence in their predictions.

Figures 1–4 show the results for Scenario 1. Figure 1 reports mean GPA predictions, with the horizontal reference line at 3.0 indicating the perceived average GPA. Not surprisingly, the number of hours studied had an enormous effect on GPA predictions (overall  $M_s = 3.42$  vs.  $2.59$ ),  $F(1, 144) = 489.80$ ,  $p < .001$ . Note, however, that the predictions above 3.0 and below 3.0 are almost mirror images of each other. The directionality of the evidence did not influence the regressiveness of the predictions ( $M_s = 0.41$  vs.  $0.42$ ),  $F(1, 144) = 0.13$ , *ns*. There were also no significant interactions between the directionality variable and the other independent variables in the design. To simplify further analysis, therefore, we created a new dependent variable by computing the difference between the actual GPA predictions and the midpoint value subjects were directed to use if they felt they had no useful information. We also collapsed the low- and high-value-of-diagnostic-evidence conditions into each other.

Figure 2 shows this extremity-of-prediction variable as a function of both accountability and information-set conditions. Looking first at just the unaccountable subjects, we found the anticipated dilution effects: Adding nondiagnostic information resulted in a predicted GPA closer to the mean than in the control condition ( $M_s = 0.315$  vs.  $0.575$ ),  $F(1, 38) = 12.07$ ,  $p < .01$ . We also found that subjects moderated their predictions even more in the contradictory-information conditions ( $M_s = 0.175$  vs.  $0.575$ ),  $F(1, 38) = 28.57$ ,  $p < .001$ . Finally, as expected, subjects in the augmented-information conditions made more ex-

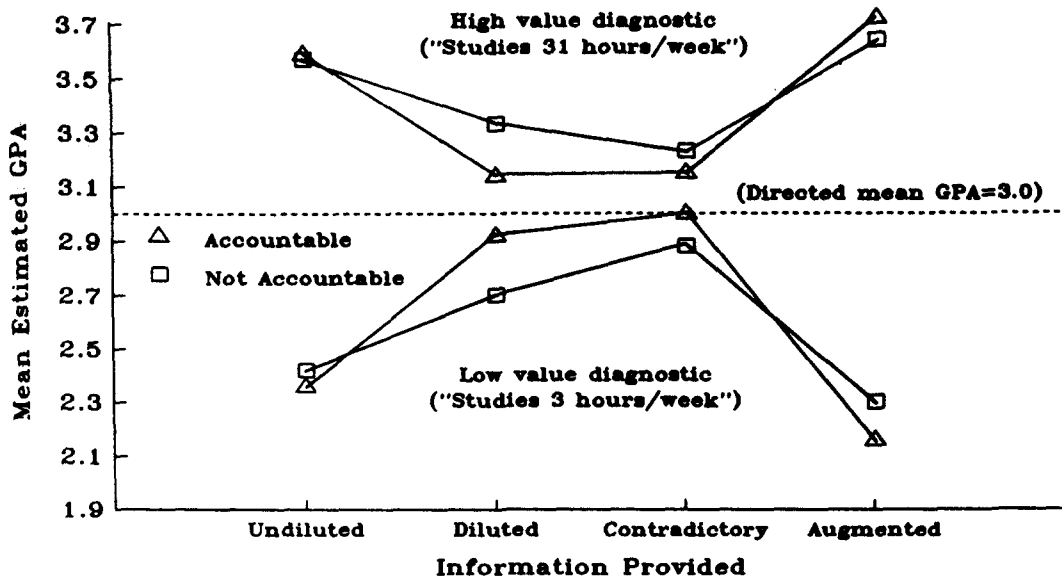


Figure 1. Mean grade point average (GPA) estimates as a function of accountability, information provided, and the value diagnostic.

treme predictions than did those in the control group ( $M_s = 0.67$  vs.  $0.575$ ), but not significantly so,  $F(1, 144) = 1.61, ns$ .

All of these effects were more pronounced among accountable subjects,  $F(3, 144) = 3.55, p < .02$ . The dilution effect was significantly greater in the accountable condition than in the nonaccountable condition (dilution effect under accountability condition,  $M_s = 0.62$  vs.  $0.11$ ; under no accountability condition,  $M_s = 0.58$  and  $0.32$ ),  $F(1, 76) = 5.36, p < .05$ . Accountable subjects were also more responsive to the additional diagnostic evidence in the contradictory-information and augmented-information conditions. Relative to unaccountable subjects, accountable subjects made more extreme judgments given augmenting information and less extreme judgments given contra-

dictory information ( $M_s = 0.78$  vs.  $0.67$  in the augmented-information conditions,  $M_s = 0.075$  vs.  $0.175$  in the contradictory-information conditions),  $F(1, 76) = 3.94, p < .05$ .

Figure 3 presents the mean integrative complexity of subjects' impressions of the student. Two trends merit note. First, subjects were more integratively complex when provided with more information. Subjects in the control conditions ( $M = 1.38$ ) were less complex than were subjects in the dilution conditions ( $M = 1.80$ ), in the contradictory-information conditions ( $M = 1.975$ ), and in the augmented-information conditions ( $M = 1.75$ ),  $F(1, 144) = 6.83, p < .01$ ;  $F(1, 144) = 13.61, p < .001$ ; and  $F(1, 144) = 5.32, p < .05$ , respectively.

Second, these trends were again more pronounced among ac-

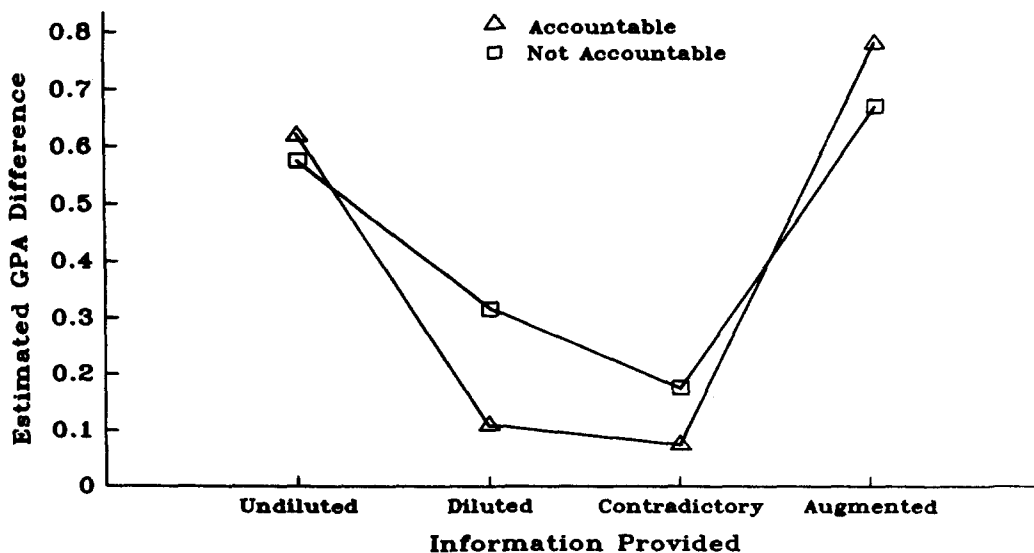


Figure 2. Grade point average (GPA) difference variable estimates as a function of accountability and the information provided.

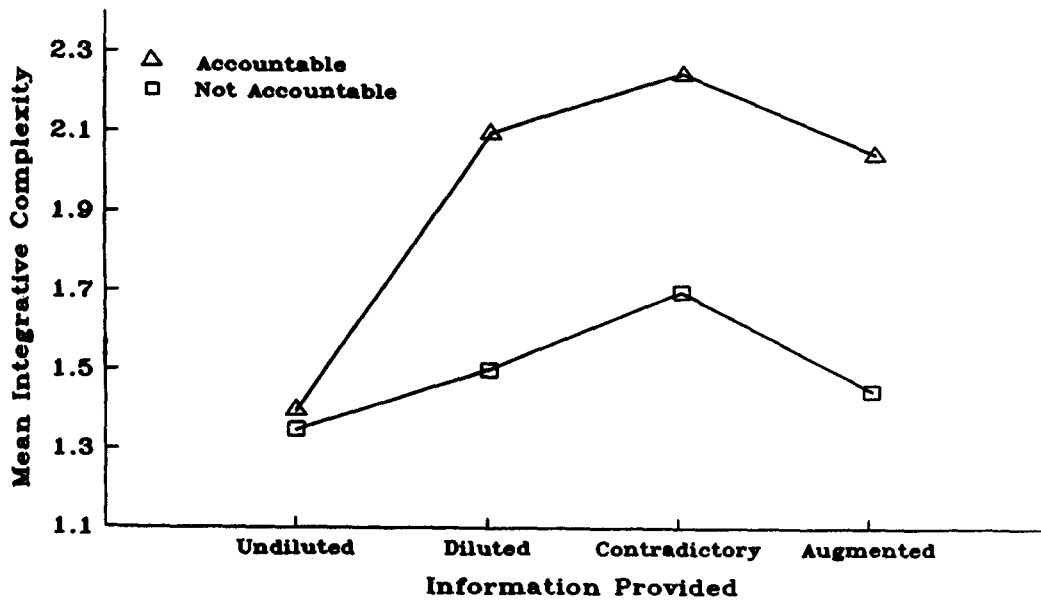


Figure 3. Integrative complexity of impressions of the student as a function of accountability and the information provided.

countable subjects ( $M_s = 1.95$  vs.  $1.50$ ),  $F(1, 144) = 15.31$ ,  $p < .001$ . Relative to nonaccountable subjects, accountable subjects became more complex in the dilution ( $M_s = 2.10$  vs.  $1.50$ ), contradictory-information ( $M_s = 2.25$  vs.  $1.70$ ), and augmented-information conditions ( $M_s = 2.05$  vs.  $1.45$ ),  $F(1, 38) = 6.81$ ,  $p < .01$ ;  $F(1, 38) = 5.72$ ,  $p < .05$ ; and  $F(1, 38) = 6.81$ ,  $p < .01$ , respectively. Accountability failed to increase complexity of thought only in the impoverished informational environment of the control conditions ( $M_s = 1.35$  vs.  $1.40$ ,  $F < 1$ ,  $ns$ ).

Figure 4 presents subjects' confidence in their predictions. Accountable subjects were less confident than were unaccountable subjects ( $M_s = 4.39$  vs.  $4.71$ ),  $F(1, 144) = 4.55$ ,  $p < .05$ , and subjects in the dilution and contradictory-information conditions were less confident than were subjects in the control and augmented-information conditions ( $M_s = 5.13$  vs.  $3.99$ ),  $F(1, 144) = 9.01$ ,  $p < .01$ . All other contrasts fell short of significance.

A remarkably similar pattern of results emerged when subjects judged the likelihood of child abuse. Figure 5 shows the predicted likelihood-of-child-abuse data, with the horizontal reference line at 6.0 representing an average likelihood of child abuse. Not surprisingly, whether subjects were told that David had violent fantasies and a drinking problem on the one hand or considered adopting a second child and did volunteer work on the other had a big effect (overall  $M_s = 4.18$  vs.  $7.55$ ),  $F(1, 144) = 283.29$ ,  $p < .0001$ . As before, the directionality of the diagnostic evidence did not significantly affect the divergence of the predictions from the mean ( $M_s = 1.82$  vs.  $1.55$ ),  $F(1, 144) = 1.88$ ,  $ns$ , and did not interact with other independent variables in the design. Accordingly, we once again derived a new dependent variable by computing the difference between subjects' actual predictions and the directed base-rate mean. We also again collapsed the low- and high-value-of-diagnostic-evidence conditions into each other.

Figure 6 shows this extremity-of-prediction variable as a

function of both accountability and information-set conditions. Looking first at just unaccountable subjects, we again replicated the basic dilution effect: Adding neutral information resulted in a predicted likelihood of child abuse closer to the mean than in the control conditions ( $M_s = 1.45$  vs.  $2.35$ ),  $F(1, 38) = 5.04$ ,  $p < .05$ . We also found that subjects' predictions virtually converged on the mean when they were given contradictory information, far closer than for the control group ( $M_s = 0.15$  vs.  $2.35$ ),  $F(1, 38) = 30.10$ ,  $p < .001$ , and significantly closer than for the dilution condition ( $M_s = 0.15$  vs.  $1.45$ ),  $F(1, 38) = 10.51$ ,  $p < .01$ . Finally, augmenting information again increased the extremity of predictions, but only to an insignificant degree compared with the control group ( $M_s = 2.65$  vs.  $2.35$ ),  $F(1, 38) < 1$ , in the nonaccountable condition.

These effects were all more pronounced among accountable subjects,  $F(3, 144) = 4.20$ ,  $p < .01$ . The dilution effect was significantly greater in the accountable condition than in the nonaccountable condition ( $M_s = 2.35$  down to  $1.45$  vs.  $2.55$  down to  $0.45$ ),  $F(1, 76) = 4.48$ ,  $p < .05$ . The tendency for accountable subjects to be more responsive to contradictory evidence than unaccountable subjects was in the predicted direction, but fell short of significance,  $F(1, 144) = 2.52$ ,  $p < .20$ . Here, an inspection of the means reveals a floor effect. The contradictory information undercut the initial diagnostic evidence so strongly that subjects' predictions almost reached the mean, diverging only 0.15 in the nonaccountable condition and 0.25 in the accountable condition. Neither mean is different from the other or from zero,  $F(1, 144) < 1$ ;  $t(18) = 0.45$ ,  $ns$ , and  $t(18) = 1.04$ ,  $ns$ , respectively. Finally, the tendency for accountable subjects to make more extreme predictions in response to augmented information was not obscured by floor effects. Accountable subjects did indeed make more extreme predictions than their nonaccountable counterparts ( $M_s = 3.65$  vs.  $2.65$ ),  $F(1, 38) = 6.22$ ,  $p < .05$ .

Figure 7 presents the mean integrative complexity of sub-

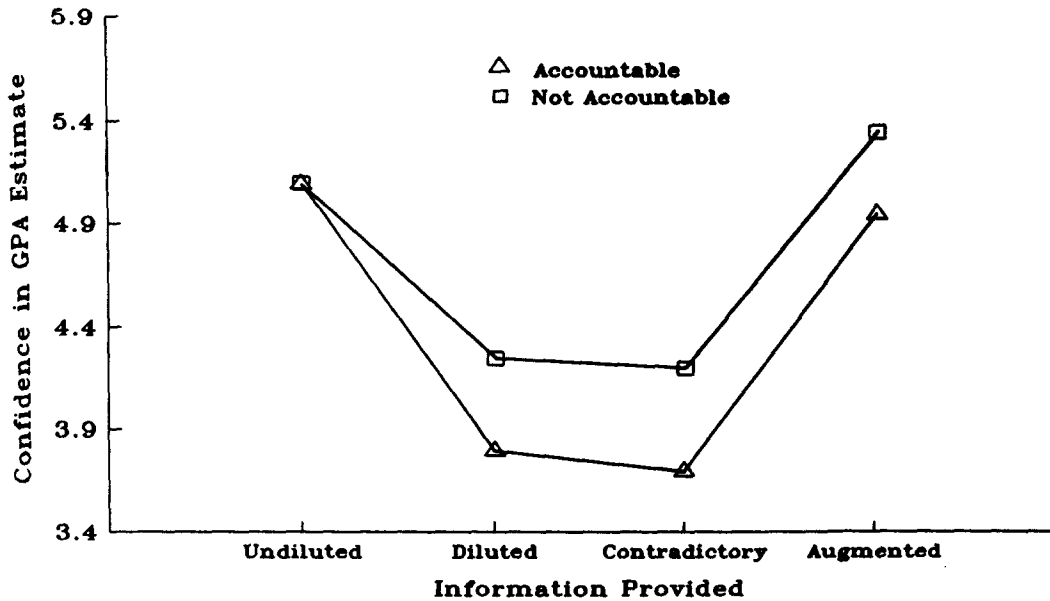


Figure 4. Confidence in grade point average (GPA) estimates as a function of accountability and the information provided.

jects' impressions of David. The same two trends emerged as in the GPA scenario. First, subjects were more integratively complex when provided with more information. Subjects in the control conditions ( $M = 1.375$ ) were less complex than were subjects in the dilution conditions ( $M = 1.85$ ) and the contradictory-information conditions ( $M = 1.90$ ),  $F(1, 76) = 8.55$ ,  $p < .01$ , and  $F(1, 76) = 19.32$ ,  $p < .001$ , respectively. Subjects in the augmented-information conditions ( $M = 1.65$ ) were also more complex than were control subjects, but not significantly so,  $F(1, 76) = 2.86$ ,  $p < .10$ .

Second, these trends were all more pronounced among accountable subjects ( $M_s = 1.875$  vs.  $1.517$ ),  $F(1, 144) = 9.95$ ,  $p < .01$ . Relative to nonaccountable subjects, these subjects were especially likely to become more complex in the dilution ( $M_s = 2.10$  vs.  $1.60$ ), contradictory-information ( $M_s = 2.10$  vs.  $1.70$ ), and augmented-information conditions ( $M_s = 1.90$  vs.  $1.40$ ),  $F(1, 76) = 4.73$ ,  $p < .05$ ;  $F(1, 76) = 3.03$ ,  $p < .10$ ; and  $F(1, 76) = 4.73$ ,  $p < .05$ , respectively.

Figure 8 presents subjects' confidence in their predictions. Once again, accountable subjects were less confident in their

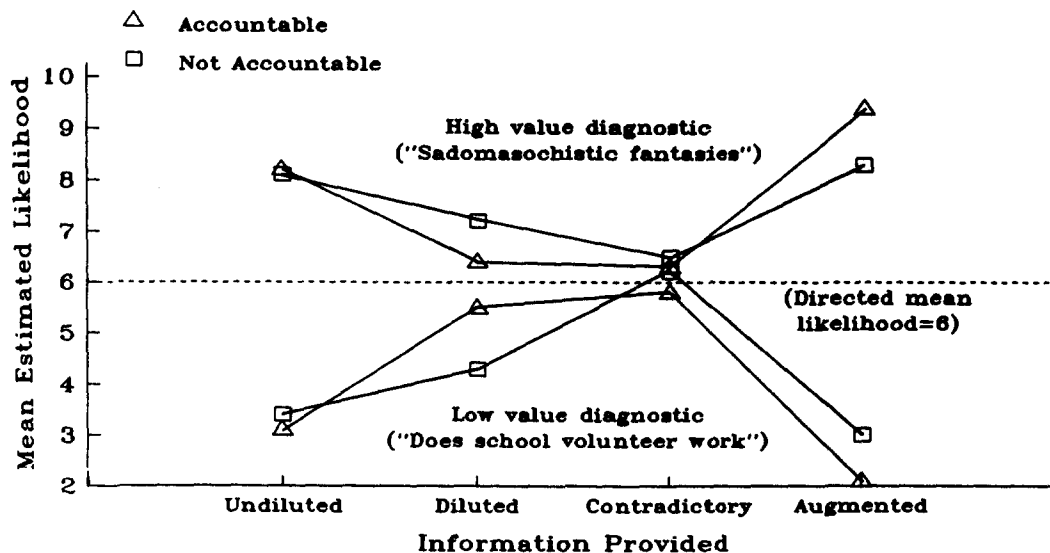


Figure 5. Mean estimates of the likelihood a patient is a child abuser as a function of accountability, information provided, and value diagnostic.

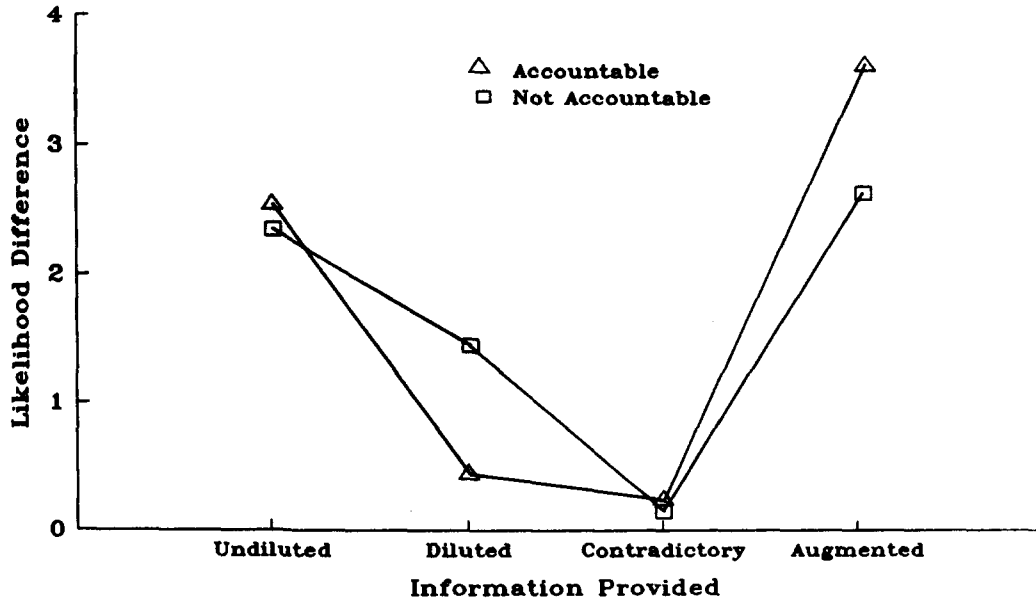


Figure 6. Likelihood difference variable estimates as a function of accountability and the information provided.

predictions than were unaccountable subjects ( $M_s = 4.42$  vs.  $4.71$ ),  $F(1, 144) = 3.81, p < .05$ , and subjects in the dilution and contradictory-information conditions were less confident than were subjects in the control and augmented-information conditions ( $M_s = 4.16$  vs.  $5.02$ ),  $F(1, 144) = 9.15, p < .01$ .

To examine the mediating role of integrative complexity, we compared the simple biserial correlations between accountability and extremity of predictions with the partial biserial correlation between those two variables controlling for integrative

complexity. We anticipated that in the dilution conditions, in which accountability was associated with both more moderate predictions and higher integrative complexity, the relation between accountability and the extremity of the prediction would disappear once we controlled for integrative complexity. For the GPA prediction, the point-biserial correlation between accountability and extremity in the diluted conditions was  $-.36, p = .01$ . After controlling for integrative complexity, the correlation fell to nonsignificance, partial point-biserial coefficient,

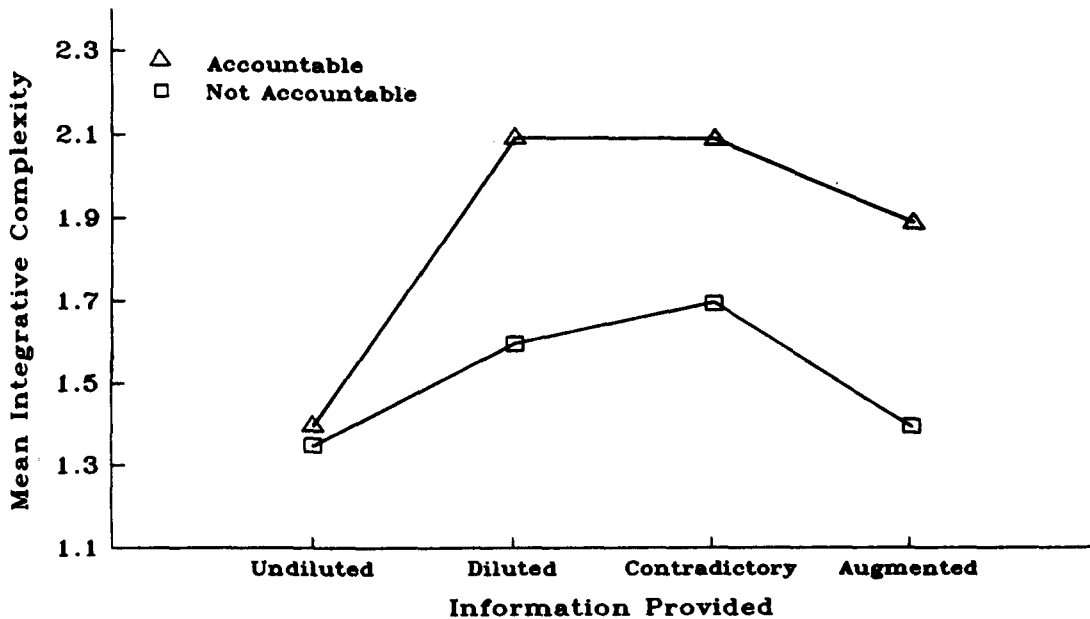


Figure 7. Integrative complexity of impressions of the patient as a function of accountability and the information provided.



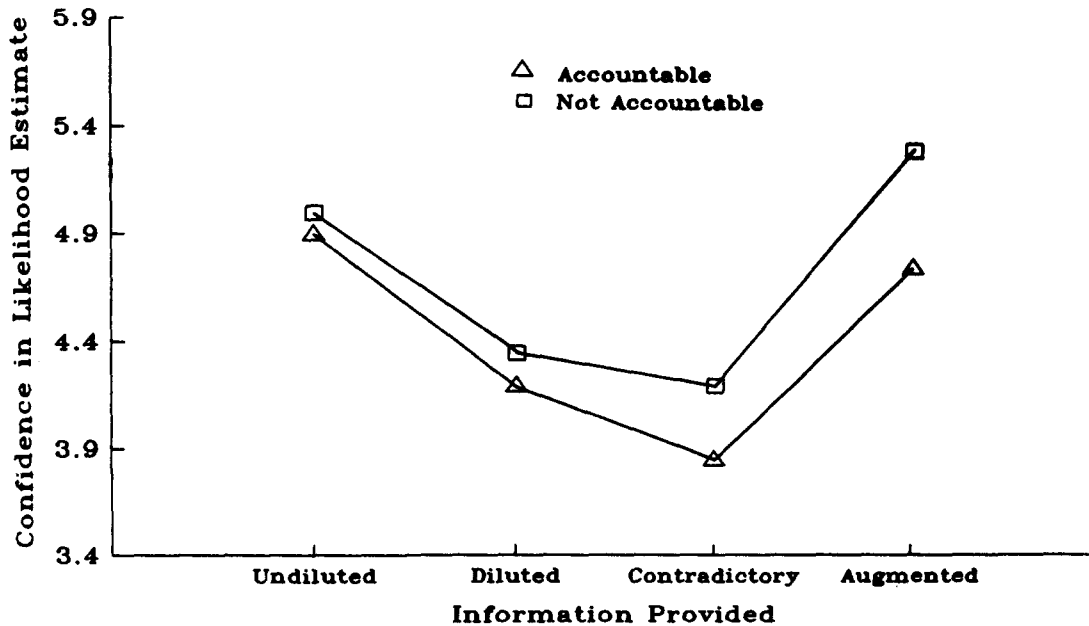


Figure 8. Confidence in likelihood estimates as a function of accountability and the information provided.

$r_{pb}(37) = -.21, p = .10$ , although it was still fairly substantial. Similarly, for the likelihood-of-child-abuse prediction, the correlation between accountability and extremity was  $-.35, p = .01$ . Controlling for integrative complexity also reduced this correlation to nonsignificance,  $r_{pb}(37) = -.23, p = .08$ , although again it was still fairly substantial.

We also computed within-cell correlations between integrative complexity and the extremity-of-prediction variable in both the dilution and augmented-information conditions. In the GPA dilution conditions, more integratively complex subjects made more moderate predictions,  $r(38) = -.46, p < .01$ . However, in the augmented-information conditions, the opposite relationship emerged,  $r(38) = .25, p = .11$ ; more complex subjects made more extreme predictions when extremism was warranted. Although the positive correlation was not significant, we found that the difference between the correlations in the dilution and augmented-information conditions was highly significant by using a Fisher's  $r$  to  $z$  transformation ( $Z = 3.24, p < .0001$ ). The analyses for the second scenario yielded similar results,  $r(38) = -.46, p < .01$ , in the dilution condition, and  $r(38) = .23, p = .15$ , in the augmented-information condition. Again, the  $r$  to  $z$  transformation revealed a highly significant difference ( $Z = 3.15, p < .001$ ).

We conducted analyses of covariance to explore whether the Accountability  $\times$  Dilution interaction could be explained by variation in integrative complexity. Focusing solely on the non-accountable conditions, integrative complexity was not a significant covariate,  $F(1, 75) = 1.46$ , and  $F < 1, ns$ , and did not alter any findings reported earlier. The dilution effect in these conditions was not strongly linked to integrative complexity. By contrast, under accountability, the complexity covariate was highly significant,  $F_s(1, 75) = 15.08$  and  $22.52, p < .001$ . The results yielded, however, at most moderate support for the mediational hypothesis. The Accountability  $\times$  Dilution interaction retained significance (although it was weakened) in Sce-

nario 1,  $F(1, 143) = 3.98, p < .05$ , and fell short of significance (although it was still substantial) in Scenario 2,  $F(1, 143) = 3.23, p = .08$ .

## Discussion

The results indicated that (a) accountability exacerbates the dilution effect, and (b) one likely mechanism underlying the impact of accountability is the tendency for accountable subjects to process evidence in more integratively complex ways than do their unaccountable counterparts. Three lines of evidence lend credibility to these claims. First, accountability did not lead to more moderate predictions in the control conditions in which only diagnostic evidence was presented. Accountability encouraged moderation only when the diagnostic evidence had been diluted by irrelevant evidence. These results held up, moreover, for both of the scenarios examined here. The effects of accountability were thus reliably more subtle and differentiated than one would have expected from the response-bias argument that people who expect to justify their views will simply stick to the safe midpoints of judgment scales. Accountability induced moderation only when subjects were given a combination of relevant and irrelevant information that they could integrate into their overall impression of the individual.

Second, far from merely encouraging moderation, accountability encouraged more extreme predictions in the augmented-information conditions in which subjects were given multiple independent reasons for expecting a particular outcome. When the evidence weighed in favor of an extreme prediction, accountable subjects were not afraid to use it.

Third, accountable subjects formed more integratively complex impressions of the evidence. Partial correlational analyses, moreover, indicated that once one controlled for the integrative complexity variable, the Accountability  $\times$  Dilution effects were either substantially reduced (Scenario 1) or nonsignificant (Sce-

nario 2). This result obviously does not prove the mediational hypothesis; nonetheless, it is consistent with the claim that integratively complex thinking is *part* of the causal mechanism by which accountability magnifies the dilution effect. The result also gains added force in view of previous studies that have documented both that accountability induces integrative complexity (Tetlock, 1983a) and that integrative complexity is associated with reduced susceptibility to the judgmental bias of overconfidence (Tetlock & Kim, 1987). The same cognitive mechanism that reduced the overconfidence bias in Tetlock and Kim (1987) was partly responsible for producing a form of underconfidence in the current study. There is apparently a fine normative line between when encouraging complex thought attenuates bias and when it creates bias in the opposite direction.

Direct examination of the thought protocols supplements the statistical analyses by revealing close conceptual connections between integrative complexity and the dilution effect. Reading the thought protocols, one cannot help but be struck by the determined efforts of accountable subjects in the dilution conditions to extract predictively useful information from the non-diagnostic information. One subject, for example, interpreted the fact that Robert had failed to date the same person for more than 2 months as a sign of emotional instability, a likely contraindicator of academic success. Another subject concluded that Robert's cheerfulness was a sign of superficiality, again a contraindicator of academic success. A third subject concluded that because Robert plays sports occasionally, he has a high energy level, a positive indicator of academic success. In brief, accountability motivated subjects to overinterpret, often to the point of implausibility, the meager evidence at their disposal in the dilution conditions. The more integratively complex the impressions subjects formed, the more likely they were to have engaged in this sort of interpretive overkill.

In closing, we raise some vexing normative issues. We have assumed up to this point that the dilution effect is indeed a bias or even error. There are good reasons for holding to this assumption. Assessed against the logical standards of multiple regression, it seems unreasonable for people to dramatically reduce their confidence in the potency of a valid predictor merely because that predictor is accompanied by irrelevant information. Moreover, in the context of "marble-bag" and "ball-in-urn" laboratory paradigms—in which there are well-defined accuracy criteria—the dilution effect appears to be a clear-cut error (cf. Troutman & Shanteau, 1977). From this standpoint, the current study—in conjunction with previous work—points to a straightforward, normative lesson: the importance of knowing when to pursue an integratively simple versus complex judgment strategy. Under some conditions, encouraging complex styles of reasoning reduces such well-documented biases as the primacy effect (Tetlock, 1983b), the overattribution effect (Tetlock, 1985b), and the overconfidence effect (Tetlock & Kim, 1987). Under other conditions, encouraging complex styles of reasoning exacerbates bias (in this case, the dilution effect). Whether the effects of accountability are normatively beneficial or not depends critically on the environment (cf. Neisser, 1976, on ecological realism). In some environments, even small increments in complexity of information processing can produce substantial gains. Considering an alternative interpretation of an event may forestall the fundamental attribution error, prevent first impressions from dominating the interpretation of

later evidence, or contribute to a realistic sense of the limits of one's knowledge. In other environments, complex efforts to make sense of events may turn out to be proverbial wild goose chases. The experimental environment set up in the dilution conditions of our study is a case in point. Subjects incorrectly assumed that the information that they had been given must be useful and made a valiant effort to figure out how it was useful. The harder accountable subjects tried to integrate irrelevant facts into their cognitive representation of the problem, the less representative the stimulus person became of the outcome to be predicted and the less confident subjects became in the predictive value of what valid information they did possess.<sup>1</sup>

It is possible, however, to look at the dilution effect from a very different normative perspective. Far from representing an error or bias, the dilution effect may constitute a rational response to the interpersonal and institutional demands that impinge on individual perceivers.<sup>2</sup> From this vantage point, the presentation of information in dilution experiments can be likened to a conversation between the researcher and the subject—an interaction in which subjects assume, following Grice's (1975) "axioms of conversation," that the information presented is indeed relevant to the task at hand. This assumption is hardly unreasonable. In most conversations, people refrain from making statements that are utterly irrelevant to the purposes of the interaction (indeed, the tendency to ramble on in disjointed fashion is taken as a sign of psychopathology). Given that the experimenters deemed it appropriate to include an assortment of evidence in their communications to subjects, one would expect a good Bayesian to attach a high prior probability to the evidence's being relevant to the task at hand. The dilution effect—and its magnification in the accountability conditions—may be as much an expression of demand characteristics ("the experimenter expects me to use all of the evidence in preparation for the conversation") as of judgmental heuristics.

We suspect that these divergent psychological interpretations of the dilution effect can eventually be disentangled by identifying ways of presenting information that are more or less likely to evoke a conversational "mental set." We doubt, however, that the normative issues will be so readily resolved. Patterns of thinking that appear profoundly flawed from one epistemologi-

<sup>1</sup> This interpretation of the Accountability  $\times$  Dilution interaction points to a boundary condition for the occurrence of the dilution effect. The effect should disappear when the diluting information is not only nondiagnostic, but also is so irrelevant to the judgment being made that even observers highly motivated to weave all the information together into a coherent narrative are unable to do so. One would, for example, have to be imaginative to the point of delusion to see conceptual connections between the student's GPA and recent outbreaks of communal violence in Burundi or a spate of methane storms on Jupiter. Moreover, the introduction of such information would also have interpersonal consequences: It would cast serious doubt on the presumed credibility, even sanity, of the source of the communication (cf. Grice, 1975).

<sup>2</sup> The dilution effect is by no means the only effect in the cognitive social literature that is open to such normative reinterpretation. One can also argue for the functional value of the overattribution effect or of belief perseverance (Tetlock & McGuire, 1986). Arguments of this sort are of more than epistemological interest; they sensitize us to the possibility that the cognitive strategies that subjects bring to the laboratory serve them quite well within more everyday contexts.

cal angle rarely look that bad, and sometimes even look pretty good, from other angles.

### References

- Abelson, R. P., & Levi, A. (1985). Decision-making and decision theory. In E. Aronson & G. Lindzey (Eds.), *Handbook of social psychology* (Vol. 2, pp. 231-309). Hillsdale, NJ: Erlbaum.
- Borgida, J. A., & Howard-Pitney, B. (1983). Personal involvement and the robustness of perceptual salience effects. *Journal of Personality and Social Psychology*, *34*, 978-989.
- Chaiken, S. (1980). Heuristic versus systematic information processing and the use of source versus message and cues in persuasion. *Journal of Personality and Social Psychology*, *39*, 752-766.
- Erber, R., & Fiske, S. T. (1984). Outcome dependency and attention to inconsistent information. *Journal of Personality and Social Psychology*, *47*, 709-726.
- Ericsson, K. A., & Simon, H. A. (1980). Verbal reports as data. *Psychological Review*, *87*, 169-196.
- Fischhoff, B. (1982). Debiasing. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty* (pp. 237-262). Cambridge, England: Cambridge University Press.
- Fiske, S. T., & Taylor, S. (1984). *Social cognition*. Reading, MA: Addison-Wesley.
- Grice, H. (1975). Logic and conversation. In P. Cole & J. Morgan (Eds.), *Syntax and semantics: Vol. 3. Speech acts* (pp. 161-192). New York: Academic Press.
- Jenkins, J. J. (1981). Can we have a fruitful cognitive psychology? In J. H. Flowers (Ed.), *Nebraska Symposium on Motivation* (Vol. 28, pp. 211-238). Lincoln: University of Nebraska Press.
- Jones, E. E. (1979). The rocky road from acts to dispositions. *American Psychologist*, *34*, 107-117.
- Kahneman, D., & Tversky, A. (1973). On the psychology of prediction. *Psychological Review*, *80*, 237-251.
- Kraut, R. E., & Higgins, E. T. (1984). Communication and social cognition. In R. S. Wyer & T. K. Srull (Eds.), *Handbook of social cognition* (Vol. 3, pp. 87-127). Hillsdale, NJ: Erlbaum.
- Markus, H., & Zajonc, R. B. (1985). Cognitive social psychology. In E. Aronson & G. Lindzey (Eds.), *Handbook of social psychology* (Vol. 1, pp. 137-230). Hillsdale, NJ: Erlbaum.
- Neisser, U. (1976). *Cognition and reality*. San Francisco: Freeman.
- Nisbett, R. E., & Ross, L. (1980). *Human inference: Strategies and shortcomings of social judgment*. New York: Prentice-Hall.
- Nisbett, R. E., Zukier, H., & Lemley, R. (1981). The dilution effect: Nondiagnostic information. *Cognitive Psychology*, *13*, 248-277.
- Payne, J. (1982). Contingent decision behavior. *Psychological Bulletin*, *92*, 382-402.
- Petty, R. E., Cacioppo, J., & Goldman, R. (1981). Personal involvement as a determinant of argument-based persuasion. *Journal of Personality and Social Psychology*, *41*, 847-855.
- Schroder, H. M., Driver, M., & Streufert, S. (1967). *Human information processing*. New York: Holt, Rinehart & Winston.
- Showers, C., & Cantor, N. (1985). Social cognition: A look at motivated strategies. *Annual Review of Psychology*, *36*, 275-305.
- Simon, H. A. (1957). *Models of man*. New York: Wiley.
- Streufert, S., & Streufert, S. (1978). *Behavior in the complex environment*. Washington, DC: Winston.
- Tesser, A. (1978). Self-generated attitude change. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 11, pp. 289-338). New York: Academic Press.
- Tetlock, P. E. (1979). Identifying victims of groupthink from public statements of decision makers. *Journal of Personality and Social Psychology*, *37*, 1314-1324.
- Tetlock, P. E. (1981). Pre- to post-election shifts in presidential rhetoric: Impression management or cognitive adjustment? *Journal of Personality and Social Psychology*, *41*, 207-212.
- Tetlock, P. E. (1983a). Accountability and complexity of thought. *Journal of Personality and Social Psychology*, *45*, 74-83.
- Tetlock, P. E. (1983b). Accountability and the perseverance of first impressions. *Social Psychology Quarterly*, *46*, 285-292.
- Tetlock, P. E. (1984). Cognitive style and political belief systems in the British House of Commons. *Journal of Personality and Social Psychology*, *46*, 365-375.
- Tetlock, P. E. (1985a). Accountability: The neglected social context of judgment and choice. In B. Staw & L. Cummings (Eds.), *Research in organizational behavior* (Vol. 1, pp. 297-332). Greenwich, CT: JAI Press.
- Tetlock, P. E. (1985b). Accountability: A social check on the fundamental attribution error. *Social Psychology Quarterly*, *48*, 227-236.
- Tetlock, P. E. (1986a). Integrative complexity of policy reasoning. In S. Kraus & R. Perloff (Eds.), *Mass media and political thought* (pp. 267-289). Beverly Hills, CA: Sage.
- Tetlock, P. E. (1986b). A value pluralism model of ideological reasoning. *Journal of Personality and Social Psychology*, *50*, 819-827.
- Tetlock, P. E., Hannum, K., & Micheletti, P. (1984). Stability and change in senatorial debate: Testing the cognitive versus rhetorical style hypotheses. *Journal of Personality and Social Psychology*, *46*, 979-990.
- Tetlock, P. E., & Kim, J. J. (1987). Accountability and judgment processes in a personality prediction task. *Journal of Personality and Social Psychology*, *52*, 700-709.
- Tetlock, P. E., & McGuire, C. (1986). Cognitive perspectives on foreign policy. In R. White (Ed.), *Psychology and the prevention of nuclear war* (pp. 255-274). New York: New York University Press.
- Tetlock, P. E., Skitka, L., & Boettger, R. (in press). Social and cognitive coping responses to accountability: Conformity, complexity, and bolstering. *Journal of Personality and Social Psychology*.
- Troutman, C. M., & Shanteau, J. (1977). Inferences based on nondiagnostic information. *Organizational Behavior and Human Performance*, *19*, 43-55.
- Tversky, A. (1977). Features of similarity. *Psychological Review*, *84*, 327-352.
- Zukier, H. (1982). The dilution effect: The role of the correlation and the dispersion of predictor variables in the use of nondiagnostic information. *Journal of Personality and Social Psychology*, *43*, 1163-1174.

Received August 8, 1988

Revision received April 3, 1989

Accepted April 4, 1989 ■