

The dilution effect: judgmental bias, conversational convention, or a bit of both?

PHILIP E. TETLOCK
The Ohio State University

JENNIFER S. LERNER and RICHARD BOETTGER
University of California, Berkeley, U.S.A.

Abstract

This study explored competing normative interpretations of the dilution effect: the tendency for people to underutilize diagnostic evidence in prediction tasks when that evidence is accompanied by irrelevant information. From the normative vantage point of the intuitive statistician, the dilution effect is a judgmental bias that arises from the representativeness heuristic (similarity-matching of causes and effects). From the normative perspective of the intuitive politician, however, the dilution effect is a rational response to evidence presented in a setting in which Gricean norms of conversation are assumed to hold. The current experiment factorially manipulated conversational norms, the degree to which diagnostic evidence was diluted by irrelevant evidence, and the accountability of subjects for their judgments. Accountable subjects demonstrated a robust dilution effect when conversational norms were explicitly primed as well as in the no-priming control condition, but no dilution when conversational norms were explicitly deactivated. Non-accountable subjects demonstrated the dilution effect across norm activation conditions, with the strongest effect under the activation of conversational norms. Although the results generally support the conversational-norm interpretation of dilution, the significant dilution effect among non-accountable subjects in the norm-deactivated condition is more consistent with the judgmental-bias interpretation.

Addressee for correspondence: Dr Philip E. Tetlock, Department of Psychology, 1885 Neil Ave., The Ohio State University, Columbus, OH, Y3210.

A National Science Foundation Grant (SRB 9505680) to the first author and a National Science Foundation Graduate Fellowship to the second author supported the conduct of this research.

The authors would like to acknowledge the research assistance of Lily Lee and Tricia Spinks, the helpful comments of Denis Hilton, Norbert Schwarz and the anonymous reviewers of this manuscript, as well as the administrative support from the Haas School of Business Administration and the Institute for Personality and Social Research.

CCC 0046-2772/96/060915-20
©1996 by John Wiley & Sons, Ltd.

*Received 16 August 1995
Accepted 10 December 1995*

INTRODUCTION

In the last 15 years, social psychologists have shown substantial interest not only in documenting how people do think, but also in prescribing how people should think. The research literature now abounds with references to biases, errors, judgmental shortcomings and other assorted deviations from normative models. Sometimes, people stand accused of failing to use information that they should have used; other times, they stand accused of using information that they should have ignored (for reviews, see Fiske & Taylor, 1991; Markus & Zajonc, 1985).

Normative judgments of this sort raise complex psychological, even epistemological, issues. Whether a particular effect qualifies as an error or bias depends critically on one's assumptions about what subjects in the experimental paradigm were trying to accomplish when they reached their judgments and whether they had reasonable grounds for supposing that reaching those judgments would help them to achieve their goals (Hamilton, 1980; Tetlock, 1991).

The present study explored the normative status of the 'dilution effect'. In a series of ingenious studies, Nisbett, Zukier and Lemley (1981) and Zukier (1982) demonstrated that linking diagnostic with nondiagnostic evidence produced more regressive predictions than people would otherwise have made. For example, people drew strong inferences about the likely grade point average (GPA) of a student when they learned that he studied either 3 hours or 31 hours per week. However, people moderated their predictions, sometimes dramatically, on receiving additional information that pretest subjects felt was completely irrelevant (e.g. plays tennis three or four times a month).

Nisbett *et al.* (1981) and Zukier (1982) 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 what an individual will do 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 hours per week is strongly associated with the outcome of a low GPA. Unique 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, dating habits).

From this theoretical perspective, the dilution effect is indeed a bias, or even an error. Assessed against the logical standards of multiple regression (the intuitive statistician), it seems unreasonable to reduce dramatically one's confidence in the potency of a valid predictor merely because that predictor is embedded in an array of irrelevant information. It is useful, however, to look at the dilution effect from a very different normative perspective. Far from representing an error, the dilution effect may constitute a rational response to the interpersonal and institutional demands that impinge on individual perceivers. 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 that the experimenter (their conversational partner) is following widely accepted norms of social discourse (Grice, 1975; Sperber & Wilson, 1986). These norms include truthfulness (do not make false or inadequately substantiated statements), relevance (do not make statements that are irrelevant to the aims of the interaction), and informativeness (be as specific as required but not more so).¹ In most conversations, people refrain from making statements that are utterly irrelevant or grossly deceptive (indeed, lay observers view the tendency to interject irrelevant comments as a sign of thought disorder and the tendency to lie as a serious character defect). 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 being relevant to the task at hand. The dilution effect 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.

This normative reinterpretation is not, moreover, entirely speculative. Several recent studies have shown that manipulations of conversational norms are powerful moderators of a number of experimental effects, including the 'underuse' of base-rate information, the 'overuse' of recently presented information, and the 'overestimation' of the likelihood of conjunctions of events (Dulany & Hilton, 1991; Hilton, 1990, 1995; Krosnick, Li, & Lehman, 1990; Schwarz, Strack, Hilton, & Naderer, 1991; Strack, Schwarz, & Wanke, 1991; Strack & Schwarz, 1992). Schwarz *et al.* (1991) state their theoretical position forcefully: many purported demonstrations of cognitive shortcomings occur, in their view, 'because basic assumptions which underlie all social discourse in everyday settings, are routinely violated in studies on judgmental biases' (p. 69).

In a previous study, Tetlock, Skitka and Boettger (1989) found that accountability magnified the dilution effect. They argued that accountable subjects were more motivated than unaccountable subjects to use all of the available information in making predictions in an effort to pre-empt potential criticism from the anticipated audience. Consistent with this interpretation, Tetlock *et al.* also found that accountable subjects did indeed form more integratively complex impressions of the persons whose conduct they were asked to predict. Moreover, this tendency to form more complex impressions was associated with greater susceptibility to the dilution effect. By motivating subjects to become more vigilant or complex thinkers, the accountability manipulation appeared to send subjects off on inferential wild goose chases in which they tried to weave together the disparate pieces of diluted

¹From a Gricean perspective, conversation is possible only if the participants follow the rules of relevance and truthfulness, assume that other parties are also following these rules, and make active, sometimes downright heroic, cognitive efforts to search for meaning in each other's utterances. To be sure, people often break these rules. We become suspicious, however, only under special circumstances (e.g. the other person has a motive to deceive us). In most conversational encounters, the default assumption is that people are indeed acting in good faith—that they are trying to be informative, truthful, and relevant. If one grants that subjects in dilution experiments walk into the laboratory with these generalized expectations, it is reasonable of them to assume that all of the information the experimenter provides them is somehow relevant to the task at hand and to try to use that information.

information into a coherent story. One subject, for example, concluded from the fact the student had never dated anyone longer than 3 months that the student must be emotionally unstable—a contra-indicator of academic success; another subject drew the opposite implication. The net result of all this cognitive activity was not, however, random. Rather, subjects appeared to follow Grice's norm of informativeness and assumed the experimenter must have had a good reason for linking a clearly diagnostic cue (studies 3 or 31 hours) with apparently irrelevant cues (dating and tennis habits) and thus thought especially carefully about ways in which the 'nondiagnostic' cues might qualify or attenuate the power of the diagnostic cue to predict the outcome of interest (GPA). The experimental outcome was, of course, an even stronger dilution effect.

The Tetlock *et al.* (1989) study shed little light, however, on the normative status of the dilution effect. On the one hand, one could argue that motivating subjects to be more integratively complex is counterproductive in environments with unfavourable signal-to-noise ratios. In such environments, complex thinkers are easily distracted by irrelevant information. They attempt to assimilate all information into their overall impressions of the target individual, and, as a result, are more likely to perceive the target individual as an unrepresentative cause of the outcome to be predicted. On the other hand, one could argue that accountable subjects acted quite rationally. These subjects expected to discuss their predictions later on with the same people who provided them with the information on which those predictions should be based. It is reasonable to posit that accountable subjects were especially willing to give their prospective conversational partners the benefit of the doubt and to assume partners were following the Gricean norms of truthfulness, relevance, and informativeness. Accountability, by this argument, simply altered the costs of making one of two possible cognitive errors: searching for relevance in others' remarks when it does not exist and failing to search for relevance when it does exist. Accountable subjects were willing to expend considerable mental effort to ensure they did not make the second type of error—the error which would prove more embarrassing in the anticipated justification session.

The current experiment disentangles these rival perspectives on dilution by systematically manipulating whether subjects believed conversational norms held in the experimental setting. Some subjects were explicitly told that the axioms of conversation did indeed apply and that the experimenter had carefully screened the information provided to subjects to ensure its relevance for the prediction task. Other subjects were explicitly told that the information may well *not* be relevant to the prediction task. Still other subjects were not given any explicit guidance one way or the other concerning the relevance of the information. In addition to activating or deactivating norms of conversation, the experiment manipulated whether subjects felt accountable for their predictions, whether diagnostic evidence was accompanied by nondiagnostic or irrelevant evidence, and the directionality (positive or negative) of the diagnostic evidence.

If the dilution effect is best thought of as a rational response to conversational norms, we should expect a dramatic reduction, perhaps even disappearance of the dilution effect when conversational norms have been explicitly deactivated. There should also be no tendency for accountability to magnify the dilution effect when conversational norms have been explicitly deactivated. By contrast, if the dilution effect is the product of a judgmental heuristic that involves automatic similarity-

matching of causes and effects, there is less reason to expect explicit deactivation of conversational norms to affect either the magnitude of the dilution effect or the tendency for the dilution effect to be greater among accountable subjects.

Finally, the potential role of integrative complexity as a mediating variable is likely to be complex. The conversational-norm hypothesis predicts that as long as subjects believe that Gricean norms hold (explicitly hold in the norm-activated condition and implicitly hold in the no-priming control condition), accountability will motivate subjects to think about even irrelevant evidence in integratively complex ways, thereby amplifying dilution. In the norm-deactivation condition, however, accountable subjects should generate less complex thoughts and become virtually immune to the dilution effect. When accountable subjects no longer believe conversational norms require them to search for relevance, they should have no difficulty selectively responding to diagnostic and ignoring nondiagnostic evidence.

METHOD

Subjects

A total of 250 college students participated in the study. Subjects were randomly assigned to conditions in a 2 (accountability) \times 2 (direction of diagnostic evidence) \times 2 (diagnostic or diluted) \times 3 (activation, deactivation or no-priming of conversational norms) between-subjects design, with 10 or 11 subjects in each condition. Eight subjects were dropped for failing to follow experimental instructions and were replaced.

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 materials described a student (Robert) and included a thought-list page on which subjects reported 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 predictions.

Subjects in the control conditions received only information about the number of hours per week the student studied, either 3 or 31 hours. Subjects in the diluted condition received the diagnostic information plus four to five pieces of nondiagnostic information. Examples of nondiagnostic information in the scenario included 'Robert is widely regarded by his friends as being honest', 'Robert plays tennis or racquetball 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'. After reading the vignette, subjects were given 5 minutes to think about the evidence and to report their findings. After reporting their thoughts about Robert, subjects predicted his GPA and assessed their confidence in their prediction. Subjects predicted GPA on a scale that ranged from 2.0 to 4.0, with 0.1 gradations.

The midpoint of the scale, 3.0, was specified to be the average student GPA at the institution, and subjects were told that, in the absence of any other useful information, the midpoint value of 3.0 would be their 'best guess' of the student's GPA.

Pretest procedures for selecting stimulus materials

The pretest procedures were based on those of Nisbett *et al.* (1981) and were reported in detail in Tetlock *et al.* (1989). Forty-five undergraduate students in a social psychology course rated the usefulness of information for predicting 'characteristics of other people'. We screened 50 items of information for the GPA prediction 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 likely to have a high GPA), counterdiagnostic (– means that the information, taken by itself, suggests a low GPA) or nondiagnostic (a zero means that the information taken by itself, is of no value in deciding whether the person has a low or high GPA).

For the experiment, we selected those items that were most consistently 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 per cent 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 per cent of the pretest subjects had to judge the items as having predictive value in the specified direction. All other things 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. The results were clear-cut. Items rated as non-diagnostic by themselves continued to be rated nondiagnostic together. The dilution effect reported later cannot be attributed to subtle interactive causal relations among the items selected here.

As a final precaution, we asked an independent group of 40 subjects to rate GPA 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). Again, the results were clear-cut. Most subjects (more than 80 per cent) continued to rate each of the nondiagnostic items as irrelevant to, 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). 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 out of 50 items to be diagnostic of the mean.

Accountability manipulation

Subjects received one of two sets of accountability instructions. Subjects in the no-accountability conditions were assured prior to receiving any information that their impressions of the target persons would be completely confidential and not traceable to them personally. Even the experimenter, it was emphasized, would not know how they had responded (the rationale for assigning each subject a five-digit identification code). Subjects in the accountability condition learned prior to receiving any evidence that the researchers would later conduct interviews to explore the types of information that subjects 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. After responding to the major dependent measures for each scenario, we asked two questions to check on whether subjects understood and believed the accountability manipulation: 'Do you expect that the responses you have made will be confidential or will be the subject of an interview?' (7-point scale, completely confidential to the subject of an interview) and 'Did you think about how you would justify your judgments while you were working on the task?' (7-point scale, 'never', to 'frequently').

Manipulations of conversational norms

Subjects received one of three sets of instructions designed to activate or deactivate the Gricean norm of 'relevance'. All instructions were given prior to reading any information on the target person whose behaviour was to be predicted. Subjects assigned to the 'activation-of-conversational-norm' condition were told:

The research team has carefully screened the information you have been given for its relevance to the task of predicting the student's GPA. Even if at first glance it seems to you that you haven't been given enough information or haven't been given the right type of information, please give careful thought to the task before you and try to think through all of the logical implications of the evidence provided.

Subjects assigned to the 'deactivation-of-conversational-norms' condition were told:

The research team cannot vouch one way or the other for the usefulness of the information that you will be given in predicting the student's GPA. The information you will be given has been randomly selected from a computer data file on the person in question. We are interested in how well people can make predictions from randomly selected facts about other people's lives.

Subjects assigned to the 'no-priming-of-norms' condition were told:

The research team will now provide you with some information about the student (whose GPA we will ask you to predict). You should feel free to use as much or as little information as you consider appropriate in making your predictions.

This latter set of instructions almost exactly replicated the procedures employed in the Tetlock *et al.* (1989) experiments. We included two manipulation check questions at the end of the experimental protocol to determine whether subjects interpreted this manipulation in the intended way: 'Do the researchers believe that the information you were given about the student is useful in making the predictions you were asked to make?' (7-point scale: 'definitely not' to 'definitely yes') and 'Did the researchers select the information on which you based your predictions in a careful, systematic way or in a random, haphazard way?' (7-point scale, 'systematic' to 'random').

Integrative complexity

Two trained coders rated the integrative complexity of the free-response personality sketches that subjects provided of Robert. 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 personality and situational determinants of complexity of information processing (see Schroder, Driver, & Streufert, 1967; Streufert & Streufert, 1978; Tetlock, 1986, 1989).

Integrative complexity is defined in terms of both conceptual differentiation and integration. Differentiation refers to the number of alternative interpretations that a person considers in analysing 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/her 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 behaviour (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 ($r = 0.89$), one of whom was unaware of both the experimental design and the hypotheses. Coders also counted the number of thoughts they judged relevant or irrelevant to the diagnostic evidence presented. Intercoder agreement was again satisfactory ($r = 0.82$).

RESULTS

Manipulation checks

To ensure that subjects interpreted the manipulations of accountability and conversational norms in the intended ways, we conducted analyses of variance of the manipulation checks. Accountable subjects were significantly more likely to believe their responses would not be private ($M_s = 5.6$ versus 3.3) and more likely to report thinking about how they would justify their judgments later on ($M_s = 5.9$ versus 3.1). Relative to norm-activated and control subjects, subjects in the norm-deactivated condition were less likely to believe the information they had been given was relevant ($M_s = 2.6$ versus 5.0) and that the researchers had selected the information in a careful as opposed to random manner ($M_s = 5.8$ versus 3.0). All differences were significant at the 0.001 level.

Creating the 'extremity-of-GPA-prediction' dependent variable

The number of hours studied had a powerful effect on GPA predictions (overall $M_s = 3.54$ versus 2.62; $F(1,248) = 1265$, $p < 0.0001$). There was also a significant asymmetry in the predictions. Learning that the student studied 31 hours a week resulted in a greater differential over the directed average GPA of 3.0 than learning that he studied only 3 hours a week which resulted in a differential below 3.0 (difference $M_s = 0.54$ versus 0.38, $F(1,248) = 38.66$, $p = 0.0001$). Although this asymmetry was substantial, it was spread relatively evenly across the 12 conditions, with no interactions between the directionality variable and the other independent variables. To simplify further analysis, therefore, and in keeping with prior procedures, we created a new 'extremity-of-prediction' dependent variable by computing the difference between the GPA predictions that subjects made and the midpoint value that subjects were directed to use if they felt they had no useful information (3.0). We then merged the directionality conditions.

Hypothesis-testing

An omnibus ANOVA revealed a second-order interaction among the three remaining independent variables: diagnosticity, accountability and norm activation, $F(2,238) = 6.26$, $p < 0.05$. Consistent with the Gricean hypotheses, the dilution effect depends not only on accountability (as shown in Tetlock *et al.* 1989), but also on the activation of conversational norms. Our analysis of this interaction revolves around questions concerning the replicability of past results as well as hitherto unexamined questions concerning the role of conversational norms in moderating dilution effects.

Were the dilution and accountability-by-dilution effects replicated in the no-norm-priming control condition?

Figure 1 shows extremity of GPA predictions in the control condition as a function of the diagnosticity and accountability manipulations. The results revealed

significant dilution effects in the baseline condition (no-priming-of-conversational-norms) (M diagnostic = 0.59 versus diluted = 0.26, $F(1,82) = 46.10$, $p < 0.05$). GPA estimates were more regressive when subjects received irrelevant information in addition to diagnostic information. The results also revealed an accountability-by-dilution interaction similar to that found in the no-norm-priming experiment by Tetlock *et al.* (1989): $F(1,80) = 7.14$, $p < 0.05$. Unaccountable subjects displayed significant dilution ($M_s = 0.56$ in the diagnostic versus 0.35 in the diluted conditions, $F(1,238) = 9.96$, $p < 0.05$), but accountable subjects were even more regressive in their estimates ($M_s = 0.63$ in the diagnostic versus 0.16 in the diluted conditions, $F(1,238) = 49.9$, $p < 0.05$).

Did activating conversational norms amplify dilution?

Activating conversational norms (Figure 1b) yielded results strikingly similar to the control condition (Figure 1a). Collapsing across accountability conditions, subjects given only diagnostic evidence produced estimates almost identical to those in the no-norm-primed control condition ($M_s = 0.65$ in the diagnostic norm-primed versus 0.59 in the no-norm primed condition, $F(1,238) = 1.33$, ns) as did subjects given diluted predictive information ($M_s = 0.25$ in the norm-primed versus 0.26 in the no-norm-primed condition, $F < 1$).

Consistent with a conversational-norm interpretation, the dilution effect was particularly powerful in the norm-activated condition, $F(1,238) = 70.04$, $p < 0.05$.

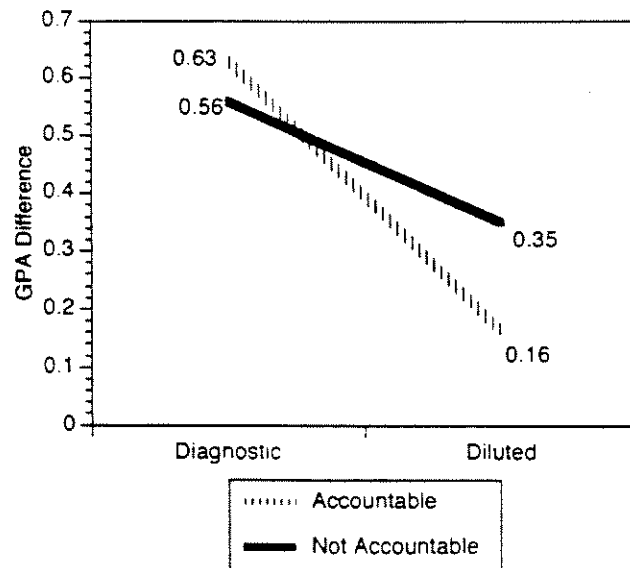


Figure 1a. Mean GPA difference as a function of diagnosticity and accountability under no-norm priming

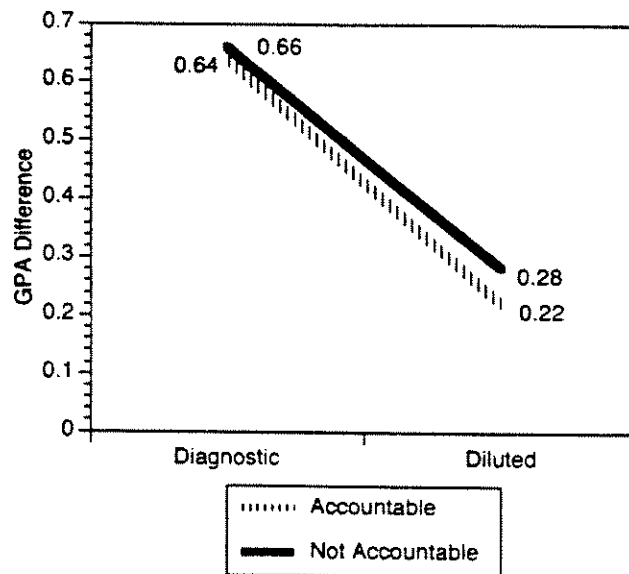


Figure 1b. Mean GPA difference as a function of diagnosticity and accountability under norm-activation

Indeed, the only major difference between the control and norm-activated conditions was that in the latter, the accountability-by-diagnosticity interaction fell far short of significance, (M_s diagnostic = 0.66 versus diluted = 0.28 < M_s diagnostic = 0.64 versus diluted = 0.22, $F < 1$)—a result traceable, in part, to the tendency of norm-primed unaccountable subjects to dilute their judgments more than they did in the no-norm primed condition. In the norm-primed condition, accountable and unaccountable subjects alike produced dramatically less regressive estimates than did subjects given diluted information (M_s diagnostic = 0.65 versus diluted = 0.25, average dilution = 0.40).

Did deactivating conversational norms attenuate dilution?

Norm deactivation attenuated dilution, especially among accountable subjects. The accountability-by-diagnosticity interaction within the norm-deactivated condition was significant ($F(1,78) = 4.89$, $p < 0.05$), but in the opposite direction from the interaction in the no-norm-primed baseline condition. As the conversational-norm interpretation predicted, accountability now attenuated dilution rather than amplified it (see Figure 1c). Simple main effects analyses indicated that the dilution effect completely disappeared among accountable subjects ($M_s = 0.63$ in the diagnostic versus 0.64 in the diluted conditions, $F < 1$).

One aspect of the results was, however, inconsistent with the conversational norm hypothesis and consistent with the representativeness-heuristic interpretation. Unaccountable subjects in the norm-deactivation condition still evidenced the dilution effect and made more regressive estimates given the irrelevant information ($M_s=0.52$ in the diagnostic condition versus 0.32 in the diluted condition, $F(1,238)=9.77$, $p=0.05$) despite being warned that the information had been randomly selected from a computer database.

Thought protocol analyses

Figures 2a, b and c present the mean integrative complexity of subjects' impressions of the student. As the conversational-norm interpretation predicted, subjects tried especially hard to weave irrelevant evidence into integratively complex impressions of the student when they believed Gricean norms held and when they felt accountable for their GPA predictions. Two first-order interactions were significant: norm-activation-by-diagnosticity ($F(2,238)=4.50$, $p<0.05$) and accountability-by-diagnosticity ($F(1,238)=6.13$, $p<0.05$). In the no-norm-primed (Figure 2a) and norm-activated conditions (Figure 2b), subjects given only diagnostic information were less complex than those in the diluted condition ($M_s=1.32$ versus 1.82 , $F(1,238)=30.68$), and accountable subjects were more

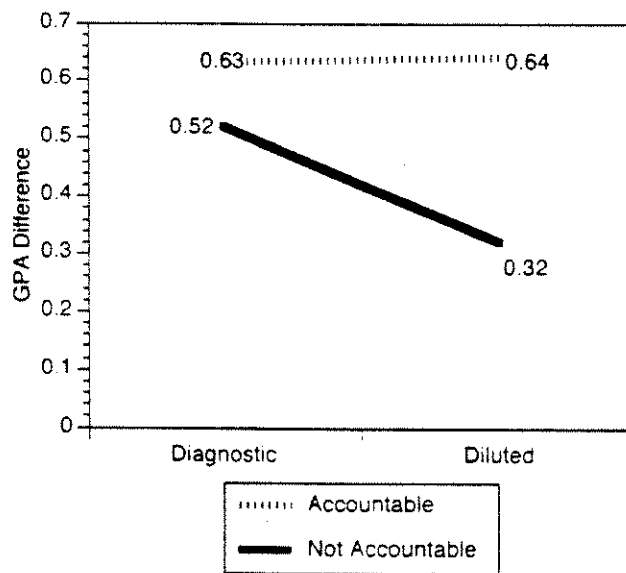


Figure 1c. Mean GPA difference as a function of diagnosticity and accountability under norm-deactivation

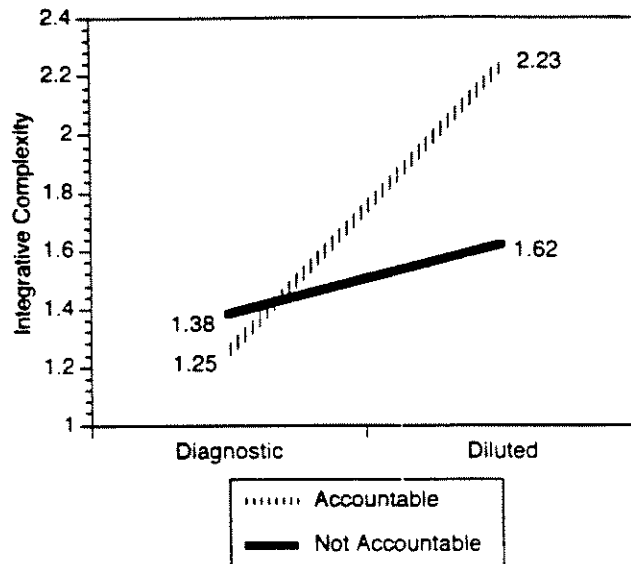


Figure 2a. Integrative complexity as a function of diagnosticity and accountability in the no-norm-primed control condition

complex than unaccountable subjects ($M_s = 1.70$ versus 1.45 , $F(1,238) = 7.78$, $p < 0.05$).

By contrast, neither the diagnosticity nor the accountability manipulations stimulated integratively complex thinking in the deactivated-norm condition (Figure 2c). Subjects in the norm-deactivated condition apparently saw little reason to engage in integratively complex thought. Average complexity was markedly lower than in both the control and norm-primed groups ($M_s = 1.34$ versus 1.63 and 1.51 , $F(2,238) = 4.76$, $p < 0.05$). Indeed, telling subjects that the evidence had been randomly sampled from a computer database brought average complexity down closer to the lowest possible score (1) than in any previous study of accountability and social cognition.

Finally, correlational analyses explored some potential mediational roles of integrative complexity.² One possibility is that there are two distinct pathways to

²The more complex the thought protocol, the more irrelevant thoughts subjects generated ($r = 0.70$; $F(1,249) = 232.08$, $p < 0.01$) and the less extreme the GPA predictions ($r = -0.25$; $F(1,249) = 17.06$, $p < 0.01$). Generally, the results for the mean number of irrelevant thoughts paralleled those for integrative complexity. Subjects in the diluted condition, not surprisingly, expressed more nondiagnostic thoughts than those given only diagnostic evidence, $M = 0.56$ versus 0.13 , $F(1,238) = 47.82$, $p < 0.00$. Accountable subjects had more nondiagnostic thoughts than did unaccountable subjects ($M = 0.41$ versus 0.29 , $F(1,238) = 3.76$, $p = 0.05$), and the group with deactivated conversational norms expressed fewer nondiagnostic thoughts than either of the other two conditions: $M = 0.22$ versus 0.38 and 0.44 , $F(2,238) = 4.42$, $p < 0.01$). The accountability-by-diagnosticity interaction was also significant, $F(1,238) = 4.72$, $p = 0.03$.

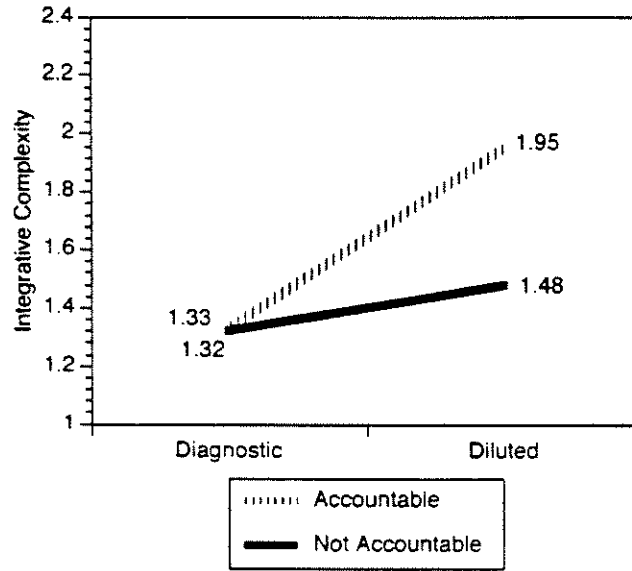


Figure 2b. Integrative complexity as a function of diagnosticity and accountability in the norm-activated condition

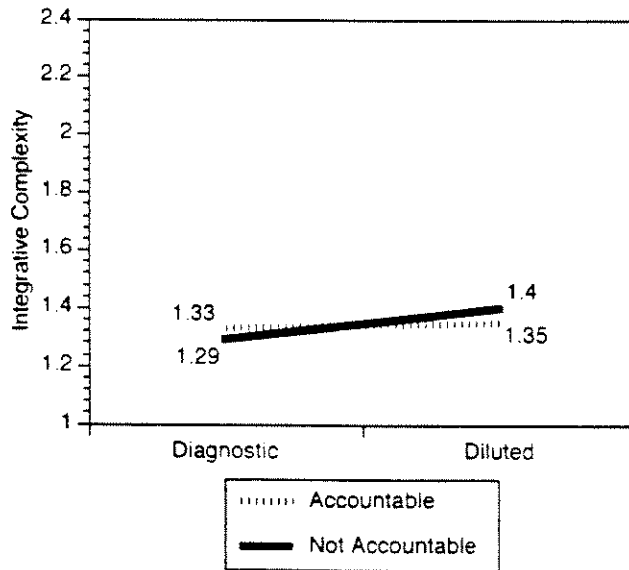
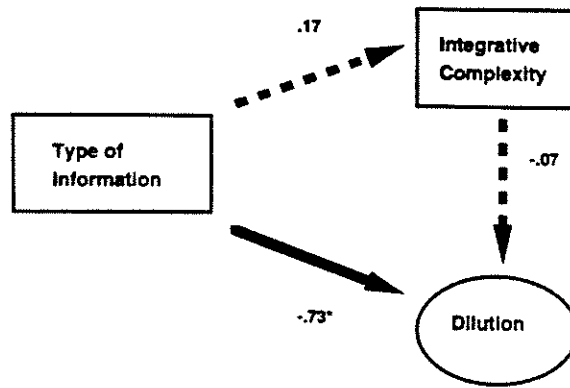
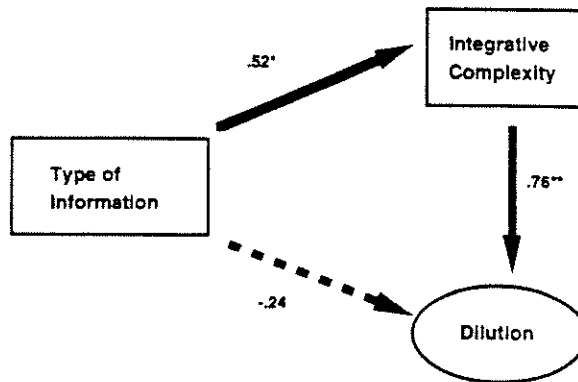


Figure 2c. Integrative complexity as a function of diagnosticity and accountability in the norm-deactivated condition



* $p < .01$ $N = 85$

Figure 3a. Path analysis: mediation of dilution for unaccountable subjects in the no-norm activated and norm-activated conditions



* $p < .01$ ** $p < .05$ $N = 83$

Figure 3b. Path analysis: mediation of dilution for accountable subjects in the no-norm activated and norm-activated conditions

dilution: an effortful, complex path that is activated when subjects are accountable and believe that conversational norms hold (either implicitly or explicitly) and a low-effort simple path that is activated when subjects are unaccountable and believe conversational norms hold. To investigate this possibility, we conducted separate path analyses for accountable and unaccountable subjects. In each analysis, we regressed subjects' predictions of student GPA on the set of potential determinants of those predictions, including type of information (diagnostic only or diagnostic plus irrelevant information) and integrative complexity of thought protocols. Each analysis included only subjects in the no-norm-primed and norm-activated conditions. Consistent with the two-pathway hypothesis, the inclusion of irrelevant information led directly to dilution among unaccountable subjects, with no mediating role for complexity of thought (see Figure 3a), but led only indirectly to dilution among accountable subjects, with a powerful mediational role for complexity of thought (see Figure 3b).

Additional regression analyses explored the power of all three experimentally manipulated variables plus integrative complexity to predict subjects' predictions of student GPA. The conversational norm analysis assigns a complex mediational role to integrative complexity. Dilution and accountability manipulations should jointly stimulate integrative complexity and thereby produce more regressive predictions, but only when conversational norms hold. Consistent with this hypothesis of conditional mediation, a multiple regression analysis—entering integrative complexity and the three experimentally manipulated independent variables—yielded a significant third-order interaction ($F(2,233)=4.90$, $p<0.01$) in which complexity of thought was associated with less extreme predictions in response to accountability and dilution in the no-norm and norm-primed conditions but not in the norm-deactivated conditions. It is also worth noting that when conversational norms held (either implicitly or explicitly), the diagnosticity manipulation and integrative complexity were positively related ($r=0.36$; $F=24.67$, $p<0.01$); as were the accountability manipulation and integrative complexity ($r=0.18$; $F=5.63$, $p=0.02$). When conversational norms did not hold, however, neither of the above relationships was significant (both correlations were weaker than 0.07).

DISCUSSION

This study assessed the relative merits of two conflicting interpretations of the dilution effect. According to one interpretation, dilution is a rational social response to conversational conventions. When someone provides you with information prior to a task, it is reasonable to presume its relevance and to make good faith efforts to use it. According to the other interpretation, dilution is an inappropriate judgment strategy rooted in a pervasive cognitive tendency to make predictions using the representativeness heuristic (in which people selectively search for causes that are representative or prototypical of the effects to be explained).

Defenders of the social rationality of the dilution effect can take special satisfaction from several findings. Most important, the dilution effect disappears among accountable subjects who were explicitly told that conversational norms did not apply because the information they had been given had been randomly selected

from a computer database. Accountability also ceased to motivate more integratively complex thought when conversational norms were explicitly deactivated. By contrast, strong dilution effects appear among subjects who have been explicitly told that conversational norms do apply and among subjects in the no-norm-primed control condition who were told nothing one way or the other about the relevance of the information provided. Moreover, accountability both amplified the dilution effect and increased the integrative complexity of thoughts in the no-norm-primed condition. Taken together, these results suggest that accountability motivated the cognitive work necessary for forming complex impressions of the evidence only when there was a reasonable presumption that the normal rules of the conversational game applied. When people were motivated to be integratively complex, they were more susceptible to the dilution effect.

The apparent strength of conversational norms in the no-priming condition adds further support to a Gricean reinterpretation of dilution. This condition most resembled experimental instructions in past dilution work and subjects assigned to it responded much more like subjects in the norm-primed than in the norm-deactivated condition. They assumed relevance (exactly as the Gricean analysis predicted), tried to weave nondiagnostic evidence into their impressions of the student, and moderated their predictions toward the overall undergraduate average. Telling subjects that the information was relevant was, in effect, largely redundant.

Contrary to the conversational-norm hypothesis, however, explicitly deactivating conversational norms was not sufficient to eliminate the dilution effect among unaccountable subjects. Although the effect was weakened (when compared to the norm-activated condition), it remained significant. It is difficult to reconcile these results with an interpretation of dilution that exclusively emphasizes conversational norms. The effect should, quite simply, have disappeared when subjects were told that the information provided had been randomly sampled from a computer database.

One theoretical possibility is that there are two distinct pathways to the dilution effect. As Figures 3a and b suggest, unaccountable subjects might rely on a simple similarity-matching heuristic that requires little mental effort and that is relatively unaffected by conversational context. They quickly try to gauge the degree to which the hypothesized cause (stimulus person) resembles the identified effect (low or high academic achievement). Accountable subjects, however, put considerable effort into processing the diluted predictive information—mental effort reflected in the greater integrative complexity of their thoughts when conversational norms had been primed either explicitly or implicitly. As observed in previous work, this increased mental effort is mostly directed to integrating the disparate strands of information at hand into a coherent story relevant to the prediction task assigned by the experimenter. From this dual-process perspective, dilution is the product of automatic information processing among unaccountable subjects and highly controlled information processing among accountable subjects.

This argument leads to testable hypotheses (*cf.* Bargh, 1994). Dilution among unaccountable subjects should occur without awareness, should be minimally affected by distractions because conscious monitoring is unnecessary, and should not require intentionality. Dilution among accountable subjects will be mediated largely by conscious mental processes in which they struggle with answering some version of the question: Why did this reasonable, even authoritative, person give me all this

apparently irrelevant information prior to asking me to make predictions? Accountable subjects will expend substantial cognitive effort trying to integrate irrelevant evidence into a coherent whole, their intentions will guide this effort, and the dilution effect among these subjects will be disrupted by distractions that require conscious monitoring and by time pressure that limits opportunities for integratively complex elaboration. Of course, people will follow this high-effort path to dilution only when they believe that the Gricean presumption of relevance holds; when we deactivate the presumption, dilution disappears among accountable subjects who now focus cognitive effort on distinguishing useful from useless information 'randomly selected' from the database. As a result, the same accountability manipulation that amplifies dilution when conversational norms hold (because subjects anticipate criticism for failure to use all available information) attenuates dilution when conversational norms cease to hold (because subjects now have social licence to be discriminating consumers of information).

Taken as a whole, the current results are consistent with other data on the impact of accountability on judgment and choice. Several studies have now documented that accountability can, under certain conditions, motivate people to form more differentiated impressions of others and to make decisions in more integratively complex ways. The current results also dovetail nicely with findings that accountability does not inevitably transform people into timid fence-sitters who are afraid to take a stand lest they offend somebody. Accountable subjects have been found in previous work to be willing to express confidence in their predictions of others' behaviour when confidence is justified (Tetlock & Kim, 1987) and to make dispositional attributions when possible situational explanations do not exist (Tetlock, 1985). This study demonstrates that accountable subjects are fully capable of disregarding irrelevant information when they believe conversational norms no longer require them to search for relevance in communications from others. Under norm-deactivation, accountable subjects generated fewer thoughts about the stimulus person whose conduct was to be predicted, less complex thoughts, and became virtually immune to the dilution effect. They also had no difficulty selectively responding to the diagnostic evidence that had been embedded in the array of information provided them. Accountable subjects who received only diagnostic information in the norm-deactivation condition made as extreme predictions as subjects who received only diagnostic information.

The experiment also explicitly links the literature on accountability to that on conversational norms. Advocates of conversational-norm interpretations have called into question a number of apparent biases in social judgment, including order effects, ignoring base-rates, and the conjunction fallacy (Hilton, 1990, 1995; Krosnick *et al.*, 1990; Schwarz *et al.*, 1991). We can now add the dilution effect to this list of normatively controversial effects. Before labelling an effect an error, we should carefully probe the assumptions that subjects make about the relevance of the information that the experimenter has provided them. By contrast, accountability researchers have been less inclined to question the normative standing of response tendencies and more inclined to identify social boundary conditions on the occurrence of certain effects. People, it is argued, can be motivated by the interpersonal or institutional context to become more vigilant information-processors who are therefore less prone to judgmental biases such as overconfidence, overattribution, and primacy. Although these theoretical

arguments are clearly distinct, taken together, they pose a serious double-pronged challenge to the still-influential characterization of the social thinker as an error-prone cognitive miser. One argument challenges the normative status of response tendencies as errors or biases; the other challenges the generalizability of the same response tendencies. With respect to dilution, several conclusions are currently in order: (a) under accountability, the dilution effect is most justifiable when conversational norms can reasonably be assumed to hold and least justifiable when these norms no longer hold; (b) accountability magnifies the dilution effect when it is most justifiable and attenuates the effect when it is least justifiable; (c) if one insists on characterizing dilution as a 'bias' (a debatable classification), the judgmental defect arises not only from cognitive miserliness, but also from excessive trust among subjects that the experimenters were playing the conversational game in the conventional way.

REFERENCES

- Bargh, J. A. (1994). Automaticity in social cognition. In R. S. Wyer, Jr., & T. K. Srull (Eds), *Handbook of social cognition* (Vol. 1, pp. 1-40). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Dulany, D. E., & Hilton, D. J. (1991). Conversational implicature, conscious representation, and the conjunction fallacy. *Social Cognition*, 9, 85-110.
- Fiske, S. T., & Taylor, S. (1991). *Social cognition*. Reading, MA: Addison-Wesley.
- 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, V. L. (1980). Intuitive psychologist or intuitive lawyer? Alternative models of the attribution process. *Journal of Personality and Social Psychology*, 39, 767-772.
- Hilton, D. J. (1990). Conversational processes and causal explanation. *Psychological Bulletin*, 107, 65-81.
- Hilton, D. J. (1995). The social context of reasoning: Conversational inference and rational judgment. *Psychological Bulletin*, 118, 248-271.
- 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.
- Krosnick, J., Li, F., & Lehman, D. (1990). Conversational conventions, order of information acquisition, and the effects of base rates and individuating information on social judgment. *Journal of Personality and Social Psychology*, 59, 1140-1152.
- 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.
- Nisbett, R. E., Zukier, H., & Lemley, R. (1981). The dilution effect: Nondiagnostic information. *Cognitive Psychology*, 13, 248-277.
- Payne, J., Bettman, E., & Johnson, P. (1992). Behavioral decision theory. *Annual Review of Psychology*. Palo Alto, CA: Annual Review Press.
- Schneider, D. J. (1991). Social cognition. *Annual Review of Psychology*, 42, 527-561.
- Schwarz, N., Strack, F., Hilton, D. J., & Naderer, G. (1991). Base rates, representativeness, and the logic of conversation: The contextual relevance of 'irrelevant' information. *Social Cognition*, 9, 67-84.
- Schroder, H. M., Driver, M., & Steufert, S. (1967). *Human information processing*. New York: Holt, Reinhart & Winston.
- Showers, C., & Cantor, N. (1985). Social cognition: A look at motivated strategies. *Annual Review of Psychology*, 36, 275-305.

- Strack, F., & Schwarz, N. (1992). Communicative influences in standardized question situations: The case of implicit collaboration. In G. R. Semin & K. Fiedler (Eds), *Language, interaction and social cognition* (pp. 173-193). London: Sage Publications.
- Strack, F., Schwarz, N., & Wanke, M. (1991). Semantic and pragmatic aspects of context effects in social and psychological research. *Social Cognition*, *9*, 111-125.
- Streufer, S., & Streufert, S. (1978). *Behavior in the complex environment*. Washington, DC: Winston.
- Tetlock, P. E. (1983a). Accountability and the 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. (1985a). Accountability: A social check on the fundamental attribution error. *Social Psychology Quarterly*, *48*, 227-236.
- Tetlock, P. E. (1985b). 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, F. E. (1986). A value pluralism model of ideological reasoning. *Journal of Personality and Social Psychology*, *50*, 819-827.
- Tetlock, F. E. (1991). An alternative metaphor in the study of judgment and choice: People as politicians. *Theory and Psychology*, *1*, 451-475.
- 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., Skitka, L., & Boettger, R. (1989). Social and cognitive strategies of coping with accountability: Conformity, complexity, and bolstering. *Journal of Personality and Social Psychology: Interpersonal Relations and Group Dynamics*, *57*, 632-641.
- 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 use of nondiagnostic information. *Journal of Personality and Social Psychology*, *43*, 1163-1174.