
The Smell of Bias: What Instigates Correction Processes in Social Judgments?

Diederik A. Stapel

University of Amsterdam

Leonard L. Martin

University of Georgia–Athens

Norbert Schwarz

University of Michigan–Ann Arbor

Participants were either informed that contextual influences bias their judgment and asked to correct for the unspecified influence (blatant warning) or they were instructed that they should correct for the unspecified influence if they felt that there may be biasing influences (conditional warning). Whereas blatantly warned participants corrected under all conditions (Study 2), conditionally warned participants corrected their judgments when the source of bias was salient but not when the source was subtle (Studies 1 to 3). Implications for models of theory-driven correction are discussed.

In the course of a normal day, people are likely to make numerous evaluations. For example, they may hire someone based on an evaluation of the candidate's research ability, teaching effectiveness, or interpersonal skills; they may decide on a dessert at lunch based on their relative liking of the options; and they may get married or divorced based on an evaluation of their own feelings or their partner's characteristics. Given the importance of many of these daily evaluations, one may expect that people would be very good at determining when they are making good, as compared to poor or biased, evaluations. It is somewhat surprising, therefore, to find that people are not especially good at making this discrimination. As Wilson and Brekke (1994) described it, "Human judgments—even very bad ones—do not smell" (p. 121). What they meant was that there is no phenomenal experience that reliably accompanies the making of a biased judgment as compared to an accurate judgment. People can feel just as confident about their bad judgments as their good judgments. On the other hand, we know that people sometimes do correct for perceived biases in

their judgments (e.g., Devine, 1989; Martin & Achee, 1992; Petty & Wegener, 1993; Schwarz, Strack, & Mai, 1991; Strack, Martin, & Schwarz, 1988). Together, these observations present us with a problem: If biased judgments do not smell, then what motivates people to correct for perceived biases in their judgments? This is the question we addressed in this research.

We begin by examining the role of people's naive, verbal theories in judgmental correction (Strack, 1992; Strack & Hannover, 1996; Wegener & Petty, 1995; Wilson & Brekke, 1994). Then, we discuss some evidence that people use theories in correction and we highlight some unanswered questions in this research. Finally, we report the results of three experiments that explored some instigators of correction.

THEORY-GUIDED CORRECTION

Wilson and Brekke (1994) suggested that people generally fail to appreciate the influence of contextual factors on their judgments because people generally have weak introspective abilities (Nisbett & Wilson, 1977).

Authors' Note: The reported research was conducted while the first author was visiting the University of Michigan as a Fulbright scholar. Further support was provided by Grant No. 575-70-074 from the Dutch Science Foundation (Nederlandse Organisatie voor Wetenschappelijk Onderzoek). We would like to acknowledge the help of Markus Kimmelmeier, Willem Koomen, and Piotr Winkielman during the research reported in this article. Correspondence should be addressed to Diederik A. Stapel, Department of Social Psychology, University of Amsterdam, Roetersstraat 15, 1018 WB Amsterdam, The Netherlands; fax: 31-20-639-1896; e-mail sp_stapel@macmail.psy.uva.nl.

PSPB, Vol. 24 No. 8, August 1998 797-806

© 1998 by the Society for Personality and Social Psychology, Inc.

Hence, they cannot reliably discriminate biased judgments from accurate judgments merely by turning inward. Being unable to detect bias through direct introspection, people must look elsewhere for guidance if they are to make their judgments more accurate.

A number of investigators (e.g., Bargh, 1992; Strack, 1992; Strack & Hannover, 1996; Wegener & Petty, 1995; Wilson & Brekke, 1994) have suggested that people get such guidance from their naive theories. Consider a person who rates a moderately attractive face as unattractive in the context of extremely attractive faces. Although this is a biased judgment, the person may fail to correct for this bias because the moderately attractive face may really appear unattractive in this context. In other words, the judgment may feel like the person's accurate judgment of the target. However, suppose that the person retrieved a theory that suggested that ratings of moderately attractive faces could be lowered by previous ratings of more attractive faces. Armed with this knowledge, the person would be in a position to correct for the biasing effects of the context. In this way, people's naive theories can inform them about a bias, even if their judgments do not smell.

After having detected a bias, people must then do something about it, but how do they know what to do? Should they make their ratings more positive or less positive? Should they shift their ratings a little or a lot? It has been suggested that the answers to these questions also come from people's naive theories (Strack, 1992; Strack & Hannover, 1996; Wegener & Petty, 1995; Wilson & Brekke, 1994). As Strack (1992) put it, "People can apply norms, rules, or theories to adjust their response for the effect of the pernicious influence. . . . It is important, however, that judges have such rules at the ready; otherwise, they would not know how to alter their responses" (p. 269).

Petty and Wegener (1993; Wegener & Petty, 1995) detailed the steps involved in this kind of theory-guided correction. They suggested that when people sense that their judgments are being biased, they consult their naive theories to determine the extent and direction of the bias. Then, they adjust their target ratings in a direction that is opposite to the theorized bias and in an extent that is commensurate with the theorized amount of the bias. To date, supporting evidence has been obtained in studies that explicitly inform participants that their judgments may be biased and ask them to correct for it.

SOME INITIAL EVIDENCE

Wegener and Petty (1995; see also Petty & Wegener, 1993) provided participants with a series of context-target configurations and asked them to indicate the likely effect of the context on their target ratings. They

found, for example, that most participants believed that ratings of a product would be biased toward desirability if the product was endorsed by attractive women as compared to unattractive women (an assimilation effect). Most participants also believed that ratings of moderately attractive women would be biased away from ratings of extremely attractive women (a contrast effect).

After establishing that there were sets of stimuli for which participants held theories of either assimilation or contrast, Wegener and Petty (1995) had participants rate these stimuli. Half of the participants were asked, without further elaboration, to rate the context and the target stimuli. The remaining participants rated the contextual stimuli and then were warned to keep their ratings of these stimuli from influencing their ratings of the target stimuli. The wording of this warning informed participants of a bias but it did not specify the direction or magnitude of this bias. Wegener and Petty assumed that the direction and magnitude would be gleaned by participants from their naive theories. The results were consistent with this hypothesis.

When participants made their ratings without being warned of a bias, their judgments indeed reflected assimilation under the conditions that they had earlier theorized would lead to assimilation but reflected contrast under the conditions that they had earlier theorized would lead to contrast. When participants had been warned of a bias, however, their judgments showed the opposite pattern. There was assimilation when participants rated stimuli that they had earlier theorized would lead to contrast but contrast when they rated stimuli that they had earlier theorized would lead to assimilation. Apparently, participants who had been warned of a bias shifted their ratings in a direction opposite to the direction of bias suggested by their naive theories.

WHAT IS INSTIGATED BY BLATANT WARNINGS?

Although Wegener and Petty's (1995) results are consistent with the hypothesis that participants used their theories to guide their corrections, their results left several questions unanswered. First, recall that Wegener and Petty's participants reported having theories of the way in which the context could bias their target judgments. Despite this, their participants did nothing to correct for this bias unless they were explicitly instructed to do so. Why would participants fail to correct for a bias that they knew existed? The most plausible explanation may be that these participants did not use their theories to detect a bias but rather used them to guide their corrections after they had been told that there was a bias. This interpretation raises another question: Did participants correct for a bias because the experimenter told them there was a bias or because they thought there was a bias?

This question cannot be answered on the basis of the procedures used by Wegener and Petty (1995) and in related work. Specifically, these procedures involved a blatant warning of the type, "Please try to make sure that your perceptions of the [context] do not influence your ratings of the [target]." These instructions convey that there is a bias in the first place; otherwise, providing this warning would violate the conversational maxim of relevance (cf. Bless, Strack, & Schwarz, 1993; Hilton, 1995; Schwarz, 1994, 1996, for conversational analyses of experimental procedures). Thus, blatant warnings bypass the first step of the assumed correction sequence (viz., bias detection). When given this type of warning, participants do not have to infer that the context may contaminate their target judgments.

Knowing that the judgment is biased, participants must then try to correct for it. As Wegener and Petty (1995) have suggested, people's naive theories may guide the correction processes that follow blatant warnings. It is important to note, however, that the retrieval of lay theories is not a necessary precondition for correction to occur. It is sufficient to consider that the judgment will infer the required correction: If the allegedly biased judgment is (very) positive, the unbiased one would presumably be (much) less positive. Conversely, if the allegedly biased judgment is negative, the unbiased one would presumably be less negative. Accordingly, people may have theories about a given source of bias, as Petty and Wegener (1993; Wegener & Petty, 1995) documented for the contexts employed in their studies, but they may not necessarily apply these theories when confronted with a blatant warning—after all, the warning, in combination with the judgment itself, provides sufficient information to arrive at a correction in a less demanding way. Thus, when blatant warnings are used to instigate judgmental correction processes, it is equivocal whether these correction processes are driven by lay theories that specify the conditions under which a bias is likely to occur.

Petty and Wegener (1993) were aware that blatant warnings may have an impact of their own when they noted that

asking subjects not to be influenced by their perceptions of the initial locations basically told them that the locations were capable of affecting their later ratings. . . . Perhaps our blatant correction instructions activate processes that are not usually active or that are unlikely to occur with more natural instigations of correction processes. (p. 150)

Petty and Wegener (1993) attempted to rule out this explanation by manipulating the blatancy of their warning. In one study (Study 4), they had participants rate either moderate or extremely positive contextual stimuli

and then rate moderate target stimuli. Some participants received no warning, some were given a subtle warning, and some were given a blatant warning. The subtle warning told participants that, following their ratings of the contextual stimuli, there would be "more vacation spots to consider." The blatant warning told participants that the next group of ratings would include a group of cities that were "quite different from the vacation spots just rated."

Petty and Wegener (1993) found that both the subtle and the blatant warnings produced corrections relative to the judgments of participants who were not warned to correct. They concluded, therefore, that it was "unlikely that the explicit nature of the requests in Experiments 2 and 3 were activating processes that were unnatural or somehow unlike corrections instigated by the more subtle procedures used in prior studies" (p. 155). However, an inspection of the means of their neutral context conditions—conditions with sets of stimuli for which participants expected no bias—calls this conclusion into question. In the neutral condition, participants who received the blatant warning corrected their judgments relative to the control condition, whereas participants who received the subtle warning did not. In contrast, participants who received the subtle warning did not correct their judgments relative to the control condition when they were exposed to neutral context stimuli. To explain this unexpected finding, Petty and Wegener (1993) suggested that the corrections

may have occurred because of the statement that the target locations were "quite different" from the contextual locations that had just been rated. . . . It may be that the wording of the blatant cue served to accentuate the initial differences between the target and neutral context locations. (p. 155)

Note, however, that the observed correction in the absence of any actual contextual bias is also consistent with the possibility that a blatant warning can elicit corrections without the retrieval of lay theories about the source, direction, and magnitude of a bias.

In sum, it is less than clear what a blatant warning does. Theoretically, a warning should invite people to access their naive theories of judgment and to scrutinize the context to detect any possible source of bias specified by these naive theories. If people detect a possible source of bias, they need to determine the likely direction and size of its influence to correct for it. The possibility that blatant warnings bypass the detection stage, and perhaps elicit corrections that are independent of people's lay theories of bias, suggests that this type of warning instigates judgmental processes that are different from more subtle types of warnings. In the present studies, we address this issue by comparing the impact of blatant warn-

ings (as used by Petty & Wegener, 1993) with the impact of conditional warnings that invite people to correct if they feel that their judgment may be unduly influenced. We predict, and find, that conditionally warned participants only engage in corrections when the source of bias is salient rather than subtle, whereas blatantly warned participants engage in corrections independent of the salience of the bias.

EXPERIMENT 1

Experiment 1 was an attempt to find a warning that would sensitize participants to a possible bias induced by exposure to the contextual stimuli without presupposing such a bias. This warning should induce participants to correct if they detect a bias but not if they do not detect a bias. We accomplished this in a very straightforward manner. After exposing participants to the contextual stimuli, some rated the target stimuli, whereas others received the following warning: "Please try to make sure that your ratings of the desirability of the [stimuli] below reflect your true response. When you feel there is something that may have an unwanted influence on your ratings, please try to adjust for that influence." This conditional warning carries no implication that there is a bias or that participants should adjust their ratings. It is up to the participants to determine whether their target judgments are accurate.

We manipulated the salience of the contextual bias by having half of the participants in these two groups (conditional warning, no warning) rate the contextual stimuli before they rated the target stimuli and by having half of the participants merely read about the contextual stimuli as part of the instructions for the target rating task. We assumed that rating the contextual stimuli would highlight the impact of these stimuli on the target ratings (Brown, 1953), whereas having the contextual stimuli embedded in the instructions would influence the participants' ratings without the participants knowing that they had been influenced (Kubovy, 1977). Hence, participants may be more likely to detect a bias after rating the contextual stimuli than after being exposed to the contextual stimuli in the instructions.

We predicted that participants who did not receive a warning would not correct. Hence, their judgments should reflect contrast (i.e., the uncorrected effect of the context-target configuration presented to them) regardless of the salience of the bias. Participants who received the conditional warning, on the other hand, may correct when they detect a bias but not when they do not detect a bias. Hence, their ratings should reflect assimilation when they rate the contextual stimuli prior to rating the target stimuli but should reflect contrast when they are merely exposed to the contextual stimuli as part of the instructions for the rating task.

Method

Participants and design. The study consisted of 98 University of Michigan undergraduates who participated in the experiment in partial fulfillment of a class requirement. They were randomly assigned to the conditions of a 2 (warning: none, conditional) \times 2 (bias: rate context, read about context) between-subjects design. Participants in a control condition rated the target stimuli without having rated or been exposed to the contextual stimuli.

Procedure. The procedure was similar to that employed by Wegener and Petty (1995, p. 41). Participants received a packet of stimulus materials, the first page of which contained the stimuli for this experiment. The packets continued with additional tasks for a number of other unrelated experiments. The rating task for this experiment was introduced as follows:

This study is part of a larger research project that looks at the different ways in which people evaluate and judge locations, such as parks, cities, and countries. In this research project, we use several questionnaires to assess people's perceptions and evaluations of all sorts of locations.

Following Wegener and Petty (1995), participants in the rate-context condition began by rating the desirability of the weather in the Bahamas, Hawaii, and Jamaica. They did this on 9-point scales that ranged from 1 (*not at all desirable*) to 9 (*very desirable*). Then, they rated the desirability of the weather in colder midwestern cities (i.e., Indianapolis, Kansas City, Des Moines), again using 9-point scales similar to those on which they had rated the contextual stimuli. This task has been shown to produce contrast as its uncorrected, default response (e.g., Petty & Wegener, 1993; Wegener & Petty, 1995).

Participants in the read-about-context condition did not rate the contextual stimuli. Rather, after the general introduction that all participants received, these participants read the following:

Thus, when we want to know something about the desirability of the weather in particular locations in the world, we assess people's perceptions in a questionnaire in which they are asked about the weather in, for example, Jamaica, the Bahamas, and Hawaii.

Half of the participants in each of these conditions (rate the context, read about the context) rated the target stimuli immediately after being exposed to the contextual stimuli. Half were given a warning between the two tasks. This warning was phrased in conditional terms. It allowed participants to correct if they detected a bias, but it carried no implications that there was a bias

or that participants should correct. Specifically, before rating the target stimuli, the warned participants read the following:

Please try to make sure that your ratings of the desirability of the weather in the locations below reflect your true response. When you feel there is something that may have an unwanted influence on your ratings, please try to adjust for that influence.

Participants in the control group neither rated nor read about the contextual stimuli. After completing the entire packet of questionnaires, all participants were thanked, debriefed, and dismissed.

Results and Discussion

Ratings of the three targets were summed to form the primary dependent measure (Cronbach's alpha = .82). These ratings were submitted to a 2 (warning: none, conditional) x 2 (bias: rate context, read about context) between-subjects analysis of variance (ANOVA). We hypothesized that participants who were not warned would render uncorrected judgments (i.e., show contrast) regardless of the salience of the contextual bias, whereas participants who had received a conditional warning would correct (i.e., show assimilation) when they had rated the context and the target but would show the default contrast when they had been exposed to the context as part of the instructions. As can be seen in Table 1, the results supported these hypotheses.

There was a main effect for warning— $F(1, 75) = 6.09, p < .05$ —indicating that judgments of warned participants were more likely than those of participants who were not warned to reflect assimilation (i.e., a correction relative to the default contrast effect). The analysis also revealed a main effect for context, $F(1, 75) = 4.73, p < .05$. Judgments were generally more likely to reflect assimilation (i.e., correction) when the contextual stimuli were rated than when they were merely presented in the instructions.

Both of these main effects were qualified, however, by the predicted interaction between warning and bias, $F(1, 75) = 10.14, p < .01$. Relative to participants in the control group, participants who were not warned showed contrast in their ratings (i.e., the default effect of this context-target configuration). Specifically, they judged the weather in the midwestern cities to be undesirable ($M = 4.2$) relative to the control participants ($M = 5.0$), and this was true regardless of whether they had previously rated the vacation spots or had merely read about these places, $t(56) = 2.33, p < .05$.

Participants who were given the conditional warning, on the other hand, showed contrast ($M = 4.2$) when they merely read about the contextual stimuli, but they

TABLE 1: Experiment 1: Mean Target Ratings (the desirability of the weather in midwestern cities in the context of the weather in vacation spots) as a Function of Warning (no, conditional) and Bias (read about context, rate context)

Bias	Warning	
	No	Conditional
Read about context	4.4	4.2
Rate context	4.1	5.7

NOTE: Scale range is from 1 (not at all desirable) to 9 (very desirable). The mean target rating of participants in the control group, who were not exposed to any of the context stimuli, was 5.0.

showed assimilation ($M = 5.7$) when they had previously rated the context, $F(1, 75) = 14.55, p < .01$. In other words, they showed an assimilative correction away from the default contrast when the contextual influence was salient but an uncorrected contrast effect when the contextual influence was not salient.

In summary, participants without a warning showed contrast effects under salient as well as subtle context conditions. Participants who received a conditional warning corrected for this influence when its source was apparent but not when its source was subtle. This suggests that the conditional warning induced participants to scrutinize the context to identify a possible bias. Their sophistication in identifying biases, however, is limited to relatively obvious sources, resulting in attempts to correct when the source is salient but not when it is subtle.

EXPERIMENT 2

Given the findings of Experiment 1, we are now able to test if a blatant warning is likely to elicit corrections under conditions in which conditionally warned participants do not perceive a bias. If so, blatantly warned participants should show correction under subtle as well as salient context conditions, whereas conditionally warned participants should only correct for salient influences, replicating Experiment 1.

To test this possibility, some participants of Experiment 2 received the conditional warning used earlier, others received the blatant warning used by Petty and Wegener (1993; Wegener & Petty, 1995), and a third group received no warning. Within each of these conditions, some participants were exposed to the contextual stimuli in a way that highlighted the potential influence of these stimuli on the target ratings, whereas some were exposed to the contextual stimuli in a way that disguised the influence. Specifically, half of the participants rated the context and target on the same dimension (i.e., desirability of weather in both cases), whereas half rated the context and target on different dimensions (job satisfaction vs. desirability of weather, respectively). We

assumed that the contextual bias would be more salient when participants rated the context and target on the same dimension than when they rated the two on different dimensions (Brown, 1953). Participants in the control condition rated the target stimuli without having rated the contextual stimuli.

We predicted that participants who were not given a warning would not correct. Hence, their ratings should reflect the default contrast effect of this context-target configuration. Participants who were given a conditional warning should correct when they rate the context and target on the same dimension (salient context) but not when they rate the context and target on different dimensions (subtle context). Finally, blatantly warned participants should correct under all conditions, regardless of the salience of the context.

Method

Participants and design. The study consisted of 163 University of Michigan undergraduates who participated in the experiment in partial fulfillment of a class requirement. They were randomly assigned to the conditions of a 3 (warning: none, conditional, blatant) \times 2 (bias: rate on same dimension, rate on different dimension) between-subjects design. In a control condition, participants rated the target stimuli without having rated the contextual stimuli.

Procedure. The procedure was similar to that used in Experiment 1. Students participated in large classroom sessions composed of between 30 to 60 people per session. Those in the experimental groups began by rating the contextual stimuli (Bahamas, Hawaii, and Jamaica) either in terms of the desirability of the weather in these locations or in terms of people's job satisfaction in these locations. As before, participants did this on 9-point scales that ranged from 1 (*not at all desirable*) to 9 (*very desirable*).

Following these ratings, some participants moved on to rate the target stimuli, whereas others read either the blatant or the conditional warning. The blatant warning was the same as that used by Wegener and Petty (1995). It asked participants to "Please try to make sure that your perceptions of the weather in the vacation spots above do not influence your ratings of the following places." The conditional warning was the same as that used in Experiment 1. It instructed participants to correct only if they detected a bias.

As the target task, all participants rated the desirability of the weather in midwestern U.S. cities (i.e., Indianapolis, Kansas City, Des Moines) on 9-point scales similar to those on which they rated the contextual stimuli. Participants in the control group rated the target stimuli without having rated the contextual stimuli. After complet-

ing the entire packet of questionnaires, participants were thanked, debriefed, and dismissed.

Results

Ratings of the three targets were summed to form the primary dependent measure (Cronbach's $\alpha = .84$). These ratings were submitted to a 3 (warning: none, conditional, blatant) \times 2 (bias: rate on same dimension, rate on different dimension) between-subjects ANOVA. We hypothesized that (a) ratings of the participants who were not warned would reflect uncorrected contrast regardless of the dimension on which the context and target were rated, (b) ratings of the blatantly warned participants would reflect assimilative correction regardless of the dimension on which the context and target were rated, and (c) ratings of participants who were given the conditional warning would reflect assimilation (i.e., correction) when they made their ratings on the same dimension but contrast (i.e., the default) when they made their ratings on different dimensions. As can be seen in Table 2, these hypotheses were supported.

The ANOVA revealed a main effect for warning— $F(1, 137) = 17.38, p < .01$ —indicating that judgments of warned participants were generally more likely than those of participants who were not warned to reflect assimilation (i.e., a correction relative to the default contrast effect). This main effect was qualified, however, by the anticipated Warning \times Bias interaction, $F(1, 137) = 4.39, p < .05$. Judgments of participants who were not given a warning showed the expected default contrast effect. Compared to participants in the control group ($M = 5.1$), participants who were not warned judged the weather in the midwestern cities to be relatively undesirable ($M = 4.2$), and this was true regardless of whether these participants judged the context and target on the same dimension or on different dimensions— $t(65) = 2.39, p < .05$ —comparing the control group with the no warning group.

Also consistent with expectations, when participants received a blatant warning, their judgments reflected assimilation (correction) regardless of the dimension on which the contextual stimuli had been judged. That is, they rated the weather in the midwestern cities as relatively desirable ($M = 5.8$) compared to participants who received no warning ($M = 4.2$), $F(1, 137) = 31.56, p < .01$. This was independent of whether they rated the context and targets on the same or different dimensions.

Finally, participants who received a conditional warning showed corrective assimilation when they rated the target and the context on the same dimension ($M = 5.9$), but they showed the default contrast when they rated the target and the context on different dimensions ($M = 4.6$), $F(1, 137) = 11.46, p < .01$. Compared to the control

TABLE 2: Experiment 2: Mean Target Ratings (the desirability of the weather in midwestern cities in the context of the weather in vacation spots) as a Function of Warning (no, conditional, blatant) and Bias (rate on same dimension, rate on different dimension)

Bias	Warning		
	No	Conditional	Blatant
Rate on same dimension	4.1	5.9	5.8
Rate on different dimension	4.3	4.6	5.7

NOTE: Scale range is from 1 (*not at all desirable*) to 9 (*very desirable*). The mean target rating of participants in the control group, who were not exposed to any of the context stimuli, was 5.1.

group ($M = 5.1$), conditionally warned participants judged the weather in the midwestern cities more favorably ($M = 5.9$) after having rated the weather in the vacation spots (an assimilative correction away from the default contrastive influence of the context)— $t(42) = 1.90$, $p = .065$ —but rated it less favorably ($M = 4.6$) after having rated people's job satisfaction in the vacation spots (the default contrast effect)— $t(41) = 1.15$, $p = .257$ —although these latter comparisons did not reach ordinary levels of significance.

Discussion

In summary, the present results replicate the findings of Experiment 1, with a new manipulation of context saliency. Again, they show that conditionally warned participants corrected when the source of bias was salient but not when it was subtle. Hence, we conclude that conditionally warned participants only correct when they are likely to identify a likely source of bias in the first place. In contrast, participants given the blatant warning corrected regardless of whether the contextual influence was subtle or salient. Of importance, these participants corrected under conditions in which conditionally warned participants apparently did not perceive any bias. A blatant warning thus makes participants perceive contaminating influences under conditions in which conditionally warned participants do not see these influences. This observation is consistent with Petty and Wegener's (1993) data that reflected an attempt to correct under neutral conditions (i.e., conditions under which no biasing context was presented to begin with).

How would such a correction be possible? Models of theory-guided correction (e.g., Strack, 1992; Strack & Hannover, 1996; Wegener & Petty, 1995; Wilson & Brekke, 1994) suggest that people draw on their theories of contextual influences to arrive at a correction. But how can people correct for sources of influences that are not very salient? Research suggests that people are better at identifying subtle sources of influence when they are

sufficiently motivated (see Thompson, Roman, Moskowitz, Chaiken, & Bargh, 1994). From this perspective, a conditional warning induces people to check the context for obvious sources of bias, and more subtle sources go unnoticed. In contrast, a blatant warning informs them that there is a bias and that they may not rest until they have identified it, thus increasing the likelihood that they discover more subtle influences. One may therefore argue that the blatancy of a warning operates in the same way as the saliency of a source of bias: It affects the likelihood that theory-guided correction processes are instigated. Although we cannot rule out this possibility on the basis of our own data, Petty and Wegener's (1993) findings render this interpretation unlikely. As noted earlier, these authors obtained corrections under neutral context conditions, for which participants did not have lay theories of bias. If blatant warnings resulted in a more accurate performance because highly motivated participants eventually detect the actual source of bias, this finding should not have been obtained.

As an alternative account, we propose that blatantly warned participants do not use an elaborate theory to arrive at a correction. Being told by the experimenter that their judgment will be biased, it is sufficient that they consider the judgment itself to infer the necessary correction: If their allegedly biased judgment is positive (negative), their unbiased judgment would presumably be less positive (negative). Hence, the biased judgment itself carries the directional information necessary for a correction, except under conditions in which the allegedly biased judgment is neutral, in which case corrections may take either direction, resulting in increased variance.

Future research needs to clarify under which conditions, and to what extent, participants draw on subjective theories of bias when they engage in corrections that are instigated by a blatant warning. We surmise that corrections may be theory driven under these conditions when a relevant theory is highly accessible, but we emphasize that such a theory is not a necessary prerequisite for correction under blatant warning conditions. In any case, the present findings highlight the importance of distinguishing between blatant and conditional warnings: The saliency of contextual bias affected the extent to which participants corrected when they were conditionally warned but not when they were blatantly warned. These findings thus suggest a less optimistic picture of lay theories of judgmental bias than has been provided by research that is solely based on blatant warnings. In terms of Wilson and Brekke's (1994) metaphor, our findings suggest that people's ability to identify smelling judgments may be limited to strong odors.

EXPERIMENT 3

In the first two experiments, we investigated the determinants of conditional correction effects under conditions in which participants held theories of contrast. To test whether similar effects may be observed under conditions in which participants hold theories of assimilation, we exposed participants of Experiment 3 to context-target pairs that produce assimilation when participants do not correct their judgments.

This extension is conceptually relevant because there are theoretical reasons to assume that people may be more likely to identify contextual influences that evoke contrast effects than to identify contextual influences that evoke assimilation effects. Unlike contrast, many assimilation effects reflect that the contextually primed information colors the target stimulus. Because the contextual bias and the judge's initial subjective experience of the target stimulus are congruent—of the same valence—it may be more difficult for people to detect an assimilative, as compared to a contrastive, bias. Given that our first two experiments indicated that people's ability to identify sources of bias is limited to relatively salient context conditions, it is therefore conceivable that they are unlikely to correct for assimilative influences, unless a blatant warning prompts them to do so.

Experiment 3 was designed to address this issue. In previous research, Wegener and Petty (1995) introduced a context-target configuration that evokes assimilation as its uncorrected reaction. Specifically, their participants rated how satisfied people were with their jobs in warm, exotic vacation spots. Some of the participants made these ratings after rating the desirability of the weather in these vacation spots, whereas others made these ratings after reading about the vacation spots in the introduction to the target-rating task. Using a blatant warning, Wegener and Petty (1995) observed correction under both conditions. We replicated their experiment with a conditional warning. We predicted that conditionally warned participants would not correct when the context manipulation is subtle (i.e., when the vacation spots are only mentioned in the introduction). It is less clear if participants would correct when the context manipulation is salient (i.e., when they rated the job satisfaction in vacation spots). On one hand, the findings of Experiments 1 and 2 suggest that correction is likely under these conditions, whereas on the other hand, assimilative influences may be less prominent than the contrastive influences addressed earlier and may hence escape participants' attention.

Method

Participants and design. The study consisted of 72 University of Michigan undergraduates who participated in the experiment in partial fulfillment of a class require-

ment. The participants were randomly assigned to the conditions of a 2 (warning: none, conditional) \times 2 (bias: rate context, read about context) between-subjects design. In a control condition, participants rated the target stimuli without reading about or rating the contextual stimuli.

Procedure. The procedure was similar to that used in Experiment 1, with the exception that the configuration of context and target stimuli gave rise to assimilation as its default, uncorrected effect (Wegener & Petty, 1995). Participants in the salient bias condition began by rating the desirability of the weather in the Bahamas, Hawaii, and Jamaica on 9-point scales ranging from 1 (*not at all desirable*) to 9 (*very desirable*). Participants in the subtle context condition read about the vacation spots in the instructions.

After being exposed to the contextual stimuli, participants either proceeded to rate the target stimuli or they received the subtle warning used in Experiments 1 and 2. The target task involved rating how satisfied people in Hawaii and the Bahamas were with their jobs. These ratings were made on 9-point scales ranging from 1 (*not at all satisfied*) to 9 (*very satisfied*). Participants in the control group rated the target stimuli without having previously rated or read about the contextual stimuli.

Results and Discussion

Ratings of the two targets were summed to form the primary dependent measure (Cronbach's alpha = .78). These ratings were submitted to a 2 (warning: none, conditional) \times 2 (bias: rate context, read about context) between-subjects ANOVA. Our hypotheses were that participants who were not warned would show the default effect of the context (assimilation) under subtle as well as salient context conditions. Moreover, we predicted that conditionally warned participants would show the default assimilation effect when they merely read about the contextual stimuli (subtle context). Finally, conditionally warned participants may (or may not) show a correction (i.e., contrast) when they rated the contextual stimuli (salient context), reflecting that an assimilative influence may be difficult to identify. As can be seen in Table 3, our results replicated the findings of Experiments 1 and 2 and reflect correction under salient context conditions.

Specifically, the ANOVA revealed a main effect for warning, $F(1, 68) = 4.76, p < .05$. Participants who received the warning were generally more likely than those who received no warning to show contrast in their judgments (i.e., a correction for the assimilative default effect). This main effect was qualified, however, by the predicted Warning \times Bias interaction, $F(1, 68) = 5.72, p < .05$. Participants who were not given a warning rated job satisfaction in the vacation spots more favorably ($M = 6.7$)

TABLE 3: Experiment 3: Mean Target Ratings (job satisfaction in Hawaii and the Bahamas in the context of the weather in these vacation spots) as a Function of Warning (conditional, blatant) and Bias (read about context, rate context)

Bias	Warning	
	No	Conditional
Read about context	6.5	6.6
Rate context	6.9	5.8

NOTE: Scale range is from 1 (*not at all satisfied*) to 9 (*very satisfied*). The mean target rating of participants in the control group, who were not exposed to any of the context stimuli, was 6.2.

than did the control participants ($M = 6.2$), and this was true regardless of whether they had rated the contextual stimuli or merely read about them, although this comparison did not reach ordinary levels of significance, $t(53) = 1.63, p = .11$.

Participants who were asked to correct if they detected a bias, however, showed contrast when they rated the context and target ($M = 5.8$) but showed assimilation when they only read about the contextual stimuli in the introduction ($M = 6.6$), $F(1, 68) = 5.11, p < .05$.

These results replicate and extend those of Experiments 1 and 2. Again, conditionally warned participants corrected when the source of bias was salient but not when it was subtle. Thus, the present findings indicate that people's sensitivity to contextual influences is not limited to contrastive influences. Instead, they identify assimilative influences as well, provided that their source is salient.¹

GENERAL DISCUSSION

In combination, the present studies confirm that people may identify contextual sources of bias and can correct their judgments in ways that are commensurate with the theorized direction and size of the bias. However, the present studies qualify the conclusions drawn from previous research by demonstrating that people's ability to identify sources of bias is rather limited. Previous research (e.g., Petty & Wegener, 1993; Wegener & Petty, 1995) was based on blatant warnings that informed participants that the context will bias their judgment, without specifying the direction of the bias. In these studies, participants corrected away from the default influences, suggesting that they correctly perceived how the context would influence them. Unfortunately, under these conditions it is unclear whether the nature of the warning or the nature of the context stimuli instigated participants' attempts to correct, as previously discussed in some detail earlier in this article.

We avoided this problem by introducing a conditional warning that asked participants to correct if they felt that

their judgment may be influenced by the context. Under this condition, participants corrected only when the source of bias was salient but not when it was subtle. In contrast, blatantly warned participants corrected under both conditions, including the subtle context conditions in which conditionally warned participants did not identify any bias to begin with. Moreover, Petty and Wegener's (1993) data showed that blatantly warned participants may even correct when a neutral context does not induce any bias in the first place (although these authors offer a different explanation).

In addition to bearing on the differential correction processes evoked by blatant and conditional warnings, the present findings bear on models of theory-driven correction. First, our results from the conditionally warned participants reflect that these participants did not engage in corrections unless they detected a bias (Martin & Achee, 1992; Wegener & Petty, 1995), thus confirming the relevance of the bias detection stage that had been bypassed in many previous studies. Second, these corrections were in the direction of contrast when participants perceived the bias to be assimilative and in the direction of assimilation when there was likely to be a contrastive bias (Petty & Wegener, 1993; Wegener & Petty, 1995), thus confirming that corrections are theory driven. Third, our results also suggest a qualification to Wilson and Brekke's (1994) conclusion that people have difficulty making their judgments accurate because judgments, even biased ones, do not smell. Perhaps they do not have to. As long as the context has a sufficiently strong odor associated with it, correction can be initiated. In our studies, it was the salience of the context that initiated correction among the conditionally warned participants.

Why would the salience of the context initiate correction? A number of researchers have addressed this question from a variety of perspectives and, surprisingly, have converged on a roughly similar answer. The general assumption is that people attempt to remove contextual reactions from their target judgments when it is clear that the contextually evoked reactions are not their spontaneous reactions to the target (e.g., Bargh, 1992; Martin & Achee, 1992; Schwarz & Bless, 1992; Schwarz & Clore, 1983; Strack, 1992). In other words, people have to be alerted—either by the nature of the context stimuli or by an explicit warning—to the possibility of a bias. For example, people may not use a contextually induced reaction in forming their target judgments if the context and the target belong to different categories (Schwarz & Bless, 1992; Seta, Martin, & Capehart, 1979). People also fail to use contextual reactions if those reactions have been attributed to a nontarget source (Kubovy, 1977; Schwarz & Clore, 1983) or if the judge has adopted a communication rule that suggests that the contextual

reaction be excluded from the target reaction (Schwarz et al., 1991; Strack et al., 1988). In addition, people attempt to avoid context-related reactions if those reactions violate social norms, as might be the case with stereotypes (Devine, 1989; Yzerbyt, Schadron, Leyens, & Rocher, 1994).

The present studies show, however, that such corrections for contextual contamination do not come easily. Even when people are pointed to the possible contamination of their judgments by the context in which these judgments are made, they will not by default attempt to correct for such unwanted influences. In our experiments, subtly warning participants that their judgments may smell only instigated correction when the relevance of the context stimuli for target judgments was salient and transparent (i.e., when context and target stimuli were rated on the same dimension). On the other hand, blatant warnings that explicitly informed participants that their judgments were likely to smell instigated correction independent of the relevance of the context stimuli. Of importance, such warnings may not necessarily require any insight into the source of the alleged bias to allow for a correction. When told that a given positive (or negative) judgment is biased, the unbiased judgment would presumably be less positive (or negative). To what extent such corrections without insight can be obtained provides a promising avenue for future research.

NOTE

1. As a reviewer pointed out to us, because in our studies treatment means do not always differ significantly from scores in the control conditions, strictly speaking it is difficult to tell which of the results we report reflect absolute assimilation and contrast and which reflect baseline effects. However, following classic studies of knowledge accessibility effects in which positive and negative (but not control) priming conditions are compared (e.g., Devine, 1989; Martin, 1986), we prefer to speak of assimilation when judgments are affected toward the valence of activated information and of contrast when judgments are affected away from that valence (see also Higgins, 1996; Schwarz & Bless, 1992).

REFERENCES

- Bargh, J. A. (1992). Does subliminality matter to social psychology? Awareness of the stimulus versus awareness of its influence. In R. Bornstein & T. Pittman (Eds.), *Perception without awareness* (pp. 236-255). New York: Guilford.
- Bless, H., Strack, F., & Schwarz, N. (1993). The informative functions of research procedures: Bias and the logic of conversation. *European Journal of Social Psychology*, *23*, 149-165.
- Brown, D. R. (1953). Stimulus-similarity and the anchoring of subjective scales. *American Journal of Psychology*, *66*, 199-214.
- Devine, P. G. (1989). Stereotypes and prejudice: Their automatic and controlled components. *Journal of Personality and Social Psychology*, *56*, 5-18.
- Higgins, E. T. (1996). Knowledge activation: Accessibility, applicability, and salience. In E. T. Higgins & A. W. Kruglanski (Eds.), *Social psychology: Handbook of basic principles* (pp. 133-168). New York: Guilford.
- Hilton, D. J. (1995). Conversational processes in causal explanation. *Psychological Bulletin*, *107*, 65-81.
- Kubovy, M. (1977). Response availability and the apparent spontaneity of numerical choices. *Journal of Experimental Psychology: Human Perception and Performance*, *3*, 359-364.
- Martin, L. L. (1986). Set/reset: Use and disuse of concepts in impression formation. *Journal of Personality and Social Psychology*, *51*, 493-504.
- Martin, L. L., & Achee, J. W. (1992). Beyond accessibility: The role of processing objectives in judgment. In L. L. Martin & A. Tesser (Eds.), *The construction of social judgments* (pp. 195-216). Hillsdale, NJ: Lawrence Erlbaum.
- Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, *84*, 231-259.
- Petty, R. E., & Wegener, D. T. (1993). Flexible correction processes in social judgment: Correcting for context-induced contrast. *Journal of Experimental Social Psychology*, *29*, 137-165.
- Schwarz, N. (1994). Judgment in a social context: Biases, shortcomings, and the logic of conversation. In M. Zanna (Ed.), *Advances in experimental social psychology* (pp. 123-162). San Diego, CA: Academic Press.
- Schwarz, N. (1996). *Cognition and communication: Judgmental biases, research methods and the logic of conversation*. Hillsdale, NJ: Lawrence Erlbaum.
- Schwarz, N., & Bless, H. (1992). Constructing reality and its alternatives: An inclusion/exclusion model of assimilation and contrast effects in social judgment. In L. L. Martin & A. Tesser (Eds.), *The construction of social judgments* (pp. 217-245). Hillsdale, NJ: Lawrence Erlbaum.
- Schwarz, N., & Clore, G. L. (1983). Mood, misattribution, and judgments of well-being: Informative and directive functions of affective states. *Journal of Personality and Social Psychology*, *45*, 513-523.
- Schwarz, N., Strack, F., & Mai, H. P. (1991). Assimilation and contrast effects in part-whole question sequences: A conversational-logic analysis. *Public Opinion Quarterly*, *55*, 3-23.
- Seta, J. J., Martin, L. L., & Capehart, G. (1979). The effects of contrast and generalization on the attitude similarity-attraction relationship. *Journal of Personality and Social Psychology*, *37*, 462-467.
- Strack, F. (1992). The different routes to social judgments: Experiential versus informational strategies. In L. L. Martin & A. Tesser (Eds.), *The construction of social judgments* (pp. 249-276). Hillsdale, NJ: Lawrence Erlbaum.
- Strack, F., & Hannover, B. (1996). Awareness of influence as a precondition for implementing correctional goals. In P. M. Gollwitzer & J. A. Bargh (Eds.), *The psychology of action* (pp. 579-596). New York: Guilford.
- Strack, F., Martin, L. L., & Schwarz, N. (1988). Priming and communication: Social determinants of information use in judgments of life satisfaction. *European Journal of Social Psychology*, *18*, 429-442.
- Thompson, E. P., Roman, R. J., Moskowitz, G. B., Chaiken, S., & Bargh, J. A. (1994). Accuracy motivation attenuates covert priming: The systematic reprocessing of social information. *Journal of Personality and Social Psychology*, *66*, 474-489.
- Wegener, D. T., & Petty, R. E. (1995). Flexible correction processes in social judgment: The role of naive theories in corrections for perceived bias. *Journal of Personality and Social Psychology*, *68*, 36-51.
- Wilson, T. D., & Brekke, N. (1994). Mental contamination and mental correction: Unwanted influences on judgments and evaluations. *Psychological Bulletin*, *116*, 117-142.
- Yzerbyt, V. Y., Schadron, G., Leyens, J.-P., & Rocher, S. (1994). Social judgability: The impact of meta-informational cues on the use of stereotypes. *Journal of Personality and Social Psychology*, *66*, 48-55.

Received November 25, 1996

Revision accepted July 22, 1997