Flexible Correction Processes in Social Judgment: The Role of Naive Theories in Corrections for Perceived Bias

Duane T. Wegener and Richard E. Petty
Ohio State University

Unlike many models of bias correction, our flexible correction model posits that corrections occur when judges are motivated and able to adjust assessments of targets according to their naive theories of how the context affects judgments of the target(s). In the current research, people flexibly correct assessments of different targets within the same context according to the differing theories associated with the context–target pairs. In Study 1, shared theories of assimilation and contrast bias are identified. Corrections consistent with those theories are obtained in Studies 2 and 3. Study 4 shows that idiographic measures of theories of bias predict the direction and magnitude of corrections. Implications of this work for corrections of attributions and bias removal in general are discussed.

Social judgments take many forms. People continually make assessments of the qualities of other people, of places, and of objects in their environment. It is clear, however, that people attempting to assess the true qualities of a target object face a rather complex task. The reason is that their perceptions of the object potentially can be “contaminated” by any of a large number of personal and contextual factors often present in the judgment situation (Wilson & Brekke, 1994). For example, activities immediately preceding the judgment might have primed concepts that influence the judge’s perceptions of the target (e.g., Higgins, Rhose, & Jones, 1977; Martin, 1986; Slull & Wyer, 1980). Alternatively, the judge might have based initial perceptions of the target on information that is later found to be incorrect or somehow inappropriate (e.g., Golding, Fowler, Long, & Latta, 1990; W. C. Thompson, Fong, & Rosenhan, 1981; Wyer & Budesheim, 1987). Even conditions residing in the judge can unduly influence his or her assessments of the target. For instance, a judge’s current emotional state might bias his or her judgments of a target to be more positive if he or she feels good, or bias them to be more negative if he or she feels bad, than the judgments should be if they had been based on the merits of the target alone (e.g., Berkowitz & Troccoli, 1990; Forgas & Bower, 1987; Petty, Schumann, Richman, & Strathman, 1993; Schwarz & Clore, 1983). Because of these potential biasing agents and others, accurate assessments of the target often require some adjustment or correction in judgments of the target that counteract the biasing factor or factors. Corrections for these kinds of biasing agents have been identified by a variety of researchers.1

For example, in research conducted by Golding et al. (1990), participants were given information about a target person. Specifically, they were provided with lists of honest, neutral, and unkind behaviors supposedly performed by the target person (e.g., one honest behavior was “Told a cashier she had given him too much change”). Later, some of the information was identified as inappropriate for use in judgments of the target person because the experimenter had presented information that was supposed to be confidential or because the information was actually about a person other than the target. Golding et al. (1990) found consistent corrections in final judgments away from the implications of false information but found no such corrections away from confidential information. In fact, participants actually made target judgments that were consistent with the confidential information (for related paradigms and results, see W. C. Thompson et al., 1981; Wyer & Budesheim, 1987).

Corrections have also been found in the context of prior priming of trait dimensions relevant to ambiguous behaviors of a target person. For example, Strack, Schwarz, Bless, Kubler, and Wanke (1993) had participants engage in a priming task that activated either positive (i.e., friendly/helpful) or negative

---

1 Of course, corrections in assessments of targets do not necessarily render the assessments more accurate. In fact, according to our flexible correction model (see also Petty & Wegener, 1993; Wegener, 1994), people might overcorrect (i.e., adjust judgments of the target farther than the biasing agent or agents had influenced assessments without the correction), correct in the wrong direction (i.e., exacerbating the actual bias), or even correct for a perceived bias that does not exist.
(i.e., dishonorable) trait dimensions before they were asked to form an impression of an ambiguous target person later in the experimental session (with a distraction task between the two activities). In addition, Strack et al. (1993) either reminded participants of the priming task before the impression formation task or did not. Strack et al. found corrections away from the implications of the primed trait when participants were reminded of the task (i.e., they contrasted target impressions away from the primes) but found no corrections away from the primes when participants were not reminded of the priming task. In fact, in the latter case, participants assimilated target impressions to the primes (for similar results using priming and impression formation tasks or using prior survey questions as primes, see Lombardi, Higgins, & Bargh, 1987; Newman & Uleman, 1990; Schwarz, Strack, & Mai, 1991).

A number of models have been developed to account for these various correction effects. For example, Martin’s (1986) set-reset model and Schwarz and Bless’s (1992) inclusion-exclusion model have dealt primarily with priming contexts similar to those used by Strack et al. (1993; see also Strack, 1992a, 1992b). Schwarz’s (1990) feelings as information approach has dealt with the implications of mood states for judgments, as has Berkowitz’s cognitive-neoassociationistic analysis of emotional effects on judgment and behavior (see Berkowitz, in press). One common feature of all of these positions is that contexts (whether they be external primes, internal reactions to primes, internal mood states, or information about the target) are postulated to have assimilative effects when the cognitive effort of social judges in that context is relatively low (i.e., when more effortful correction processes are unlikely to occur). That is, prevailing models have assumed that the default effect of contexts is generally to produce assimilation, rendering judgments closer to the contextual stimuli than would normally be the case. Although the models differ somewhat in their accounts of how assimilation occurs and in precisely how effortful correction processes subvert the default assimilation (and some models predict high-effort contrast away from the context, whereas others simply predict less assimilation or no effect of the context), the prevailing conceptual positions share the feature that increases in cognitive effort by judges aware of some contextual influence lead to less assimilative impact of the context. That is, corrections lead to target judgments less like the context than in no-correction settings.2

For example, deriving predictions from the set-reset model, Martin, Seta, and Crelia (1990) hypothesized that ratings of a target person described in ambiguous terms would be assimilated to the implications of a previous blatant prime in situations in which previous research had shown that cognitive effort would be relatively low. On the other hand, judgments were predicted to be contrasted from the implications of the blatant primes in situations in which past research had shown that cognitive effort was relatively high. In support of this view, Martin et al. (1990) found that when people were low in need for cognition (Cacioppo & Petty, 1982), were distracted during the judgment task (Petty, Wells, & Brock, 1976), or shared responsibility for the judgment (Petty, Harkins, & Williams, 1980), assimilation to the context occurred. However, judgments were contrasted from the implications of the blatant prime when people were high in need for cognition, not distracted, or individually responsible for the judgment. According to the set-reset model, assimilation occurs because of the overlap between reactions to the context and reactions to the judgment target (Martin et al., 1990). If that overlap is not recognized (and removed), the target is viewed as more like the context than would normally be the case (i.e., “setting” takes place). However, if the person becomes aware that his or her reactions toward the target are being influenced by the overlap in reactions toward the context and target, the person might effortlessly subtract or partial out reactions toward the context from reactions toward the target (i.e., “resetting” takes place). This effortless correction makes assessments of the target less like the context, and “overcorrection” (by subtracting even true reactions toward the target) leads to contrast of the target judgments away from the context (see Martin, 1986; Martin et al., 1990).

An alternative way to organize the various correction phenomena is to propose a more flexible set of correction processes driven by respondents’ naive theories of how any given contextual factor has influenced their perceptions of the target (e.g., Petty & Wegener, 1993; Wegener, 1994; Wilson & Brekke, 1994). That is, whereas past models of correction have focused on aspects of the context such as the overlap between mental representations of the context and target (Martin, 1986;...

2 Some researchers have also attempted to account for such findings without invoking the concept of corrective processes per se. For example, Lombardi, Higgins, and Bargh (1987) found assimilation to primes when participants could not recall any of the priming episode but contrast to primes when they could recall at least one of the primes. In a second study, Lombardi et al. found that recall of any of the primes was still associated with assimilation if participants believed that the priming task had been interrupted (see also Martin, 1986). According to Lombardi et al. (1987), these results suggest that participants conscious of the priming events were capable of adopting differing processing strategies, depending on whether the priming task had been interrupted or not. Higgins (1989) explained this view by noting that primes that provide a clear instantiation of a given trait might form a relatively extreme standard of comparison against which ambiguous target behavior is compared and thus contrasted when judges can recall the prime(s). Thus, according to this view, contrast is due to primes acting as a standard of comparison rather than to attempts to remove the influence of the primes from judgments of the target (cf. Martin, 1986; Schwarz & Bless, 1992). However, it is not clear why the same prime(s) would be viewed as less extreme or why participants would choose to use the prime(s) differently in judging the target when they believe that the priming event is incomplete (see Martin & Achee, 1992, for further discussion of this perspective). Within more correction-based models such as Martin’s set-reset theory, assimilation when the priming task is incomplete is taken as support for the notion that thought perseverence elicited by an incomplete priming task makes it more difficult to remove the influence of primes on target ratings (e.g., see Martin, 1986; Martin & Achee, 1992).

Another way that researchers have explained the attenuation of priming effects is by hypothesizing that features of a context might prime a dimension of judgment, and judges might especially scrutinize information about the target along that dimension of judgment (e.g., Somin, Carlson, & Isham, 1993). In addition, explicitly providing accuracy (Ford & Kruglanski, in press; E. P. Thompson, Roman, Moskowitz, Chaiken, & Bargh, 1994) or other goals (e.g., Sedikides, 1990) might motivate people to process or reprocess information in ways leading to various interpretations of information about the target that might or might not match implications of the priming stimuli.
Schwarz & Bless, 1992) or changes in meaning of the scale
anchors when the context is used to define those anchors (Schwarz
& Bless, 1992), an alternative approach to organizing past work
focuses on respondents' theories of how the context might have
influenced (i.e., biased) their perceptions of the target and on
the necessary steps to correct the perceived biasing influence(s).3

Work on theory-based correction follows directly from the
notion that avoiding the influence of a stimulus would require
awareness of the influence of the stimulus rather than awareness
of the stimulus itself (e.g., see Bargh, 1992; Higgins & Bargh,
1992; Jacoby & Kelley, 1990). For some time, researchers have
noted that people are likely to possess or generate theories about
how various contextual factors might influence their perceptions
of target objects (e.g., Nisbett & Wilson, 1977; Wilson,
Laser, & Stone, 1982). For example, Nisbett and Wilson (1977)
documented that people might believe that a factor influenced
their perceptions (e.g., noise in an adjacent room) even if the
factor had little demonstrable effect. As in this example, much
of the attention given these theories has focused on the overall
accuracy or inaccuracy of the theories (e.g., Nisbett & Wilson,
1977; Wilson et al., 1982) rather than on how those naïve theo-
ries of bias are used in attempts at correction. Some researchers
have speculated, however, that respondents might use their naïve
perceptions of how a biasing factor has influenced them to
adjust target ratings to compensate for the undue influence of
the biasing factor (e.g., see Strack, 1992a; W. C. Thompson et
Unfortunately, these researchers have not provided explicit em-
pirical evidence for theory-based correction processes.

In addition, researchers have used the theory-based correc-
tion notion in conjunction with assumptions about theories
judges might hold to account for empirical outcomes not ex-
plained by some other judgment model. For example, theory-
based correction processes have been used to attempt to explain
why adjustments in target ratings have generally been larger for
negative or convicting information than for positive or vindic-
ating information (e.g., see W. C. Thompson et al., 1981; Wyer
& Budesheim, 1987). That is, because people might begin with
relatively positive expectations of others not previously encoun-
tered, positive information might be seen as less biasing than
negative information and, thus, less likely to induce correction
processes. Unfortunately, no empirical tests exist of whether
these postulated corrections are driven by naïve theories of bias.
In fact, empirical tests of flexible correction notions and ex-
tended discussions of factors involved in theory-based correc-
tion have only begun to appear in the literature.

Recently, we introduced a model of theory-based flexible
corrections and provided relevant empirical evidence (see
Petty & Wegener, 1993; Wegener, 1994). According to this
flexible correction model, in some situations social judges
have an initial reaction to a judgment target that might or
might not actually have been influenced by contextual fac-
tors. Nevertheless, at times the social perceiver is motivated
and able to identify potential sources of bias and comes to
believe that some contextual influence (i.e., bias) is operating
(e.g., because of external prompts or reminders in the judg-
ment situation or because of personal factors that prompt
searches for possible biases). When the person is also moti-

3 Some researchers have noted the possibility of theory-based correc-
tion but have included assumptions seemingly inconsistent with the
flexible nature of theory-based correction. For example, Strack (1992a)
noted that "a correction ... may be based solely on a theory-guided
assumption that a potentially distorting influence may be operating" (p.
258) but also stated that "representativeness is the default option" and
"representative information or experiences will influence the judgment
in the direction of their semantic or affective implications and lead to
an assimilation effect" (p. 261). Strack (1992a) also noted that if in-
formation is deemed nonrepresentative, judges can attempt to ignore
the activated content (producing no bias in judgment), can engage in
theory-based correction, or can use the activated information in for-
ming a standard of comparison (each of which leads to less assimilation
and can yield a contrast effect). Furthermore, Strack (1992b) noted that in-
fuences of previous questions in a questionnaire can be understood in
the following way. When respondents are not aware of the previous ques-
tion, the result is an assimilation effect (due to activation of information
by the previous question). When respondents are aware of the previous
question and Grecian maxims do not apply, the effect of the question is
contrast (due to correcting for the presence of the previous question;
see Strack, 1992b, pp. 33-34). Thus, although Strack's (1992a) general
statements about theory-based correction are largely consistent with
our view, his assumptions regarding the assimilation-contrast literature
are not. This is because Strack discussed assimilation and contrast by
making the same effort and correction assumptions as the Martin, Seta,
and Creila (1990) set-reset model and the Schwarz and Bless (1992)
inclusion-exclusion model (i.e., that assimilation is the default low-
effort outcome and contrast requires additional corrective effort).

4 This is not to suggest that flexible theory-based corrections occur
only after an initial reaction. More generally, we assume that corrective
processes ensue when people become aware of a potential bias. This
awareness can occur before, during, or after operation of a judgment
process. Accordingly, corrections for bias need not occur only after
judgment; people might also attempt to avoid the perceived bias by
changing the way in which information about the judgment target is
gathered or scrutinized.

5 Some of these notions are similar to the views of Jacoby and his
colleagues on conscious attempts to subvert influences on memory and
other subjective experiences (e.g., Jacoby & Kelley, 1987, 1990). For
example, Jacoby and Kelley (1987) reported a study in which partici-
pants were asked to judge how difficult other people would find the sol-
sution of various anagrams. In one condition, the solution of the ana-
gram was presented along with the anagram so that participants could
not rely on their subjective experience of difficulty by solving the anagram
themselves before judging difficulty for others. In this condition,
Jacoby and Kelley noted, participants could replace subjective experi-
ence only by a theory of what makes anagrams difficult. Jacoby and
Unlike models that assume that assimilative biases are default outcomes and that contrastive effects of those contexts are due to effortful correction processes, the flexible correction model assumes that either assimilation or contrast can be the perceived effect of a context and that either perceived bias can be corrected. As noted earlier, according to the flexible correction model, theory-based adjustments in judgments of targets would take place even if the perceived biases did not actually occur, as long as the judge believed that an unwanted bias was operating. An important implication of the flexible correction view is that either assimilation or contrast outcomes could be due to effortful correction processes or to low-effort lack of correction processes. That is, either could be the default outcome or the (over)corrected outcome. If the perceiver is unmotivated or unable to correct assessments of the target in light of contextual biases, his or her final judgment is likely to reflect his or her mapping of the initial reaction to the target (i.e., the default reaction) onto the response scale available for that rating (see also Wegener, 1994).

In an initial test of the flexible correction model, Petty and Wegener (1993) found that people can hold theories of contrast as the “natural” (i.e., uncorrected) influence of some context(s). Furthermore, using a context for which participants held a contrastive theory, Petty and Wegener (1993) found that both blatant and subtle prompts to consider possible effects of the context on target ratings led to corrections of target ratings in a direction toward (rather than away from) the context. This correction is contrary to theories of assimilation and contrast that assume that effortful corrections inevitably shift target ratings away from the person’s reactions to the context.

In addition to potentially organizing correction findings across different paradigms and domains, the notion of flexible theory-based corrections also generates hypotheses that are not readily derivable from models not based on flexible use of theories of bias in correction. For example, as noted earlier, the flexible correction framework posits that either assimilation or contrast can be the default result of contextual influence and that either assimilation or contrast outcomes can be the result of effortful overcorrections. The flexibility of theory-based corrections has yet to be fully developed and demonstrated, however. For example, no research has reported opposite theory-based corrections. That is, it is not clear that theory-based corrections can result in judgments displaced away from the context for some tasks or individuals but that opposite theories dispose corrections to result in displacements toward the context for other tasks or individuals. This is important because, within the flexible correction perspective, corrections guided by judges’ naive theories of bias move in opposite directions to the extent that people’s theories of bias denote opposite biases. This would occur even if different people have different theories about the same context (e.g., if one person thinks that blatant primes make target judgments more like the prime, but another person thinks that blatant primes make judgments less like the prime). Opposite correction would also occur if people have different theories about how the same context influences judgments for different targets (e.g., if a person believes that an extreme blatant prime makes judgments about one target less like the prime but believes that the same prime makes judgments of another target more evaluatively consistent with the primed evaluation). Although some models of assimilation and contrast predict that different effects on target judgment might occur for the same context (e.g., for some targets, the context might be included in the representation of the target, which leads to assimilation, but the same context might be excluded from the representation of another target, which leads to contrast; Schwarz & Bless, 1992), only a theory of correction based on judges’ naive theories of bias would predict different corrections (i.e., away from and closer to the context) when the same context is believed by judges to have different biasing effects.

To examine the viability of the critical opposite-correction outcome derived from the flexible correction model, we sought, in an initial study, to identify contexts and judgments for which people possess opposite theories of how the same context influences judgments about different targets. Then, in Studies 2 and 3, we tested whether the opposite theories identified produce opposite corrections in judges’ ratings of the targets when judges are prompted to correct their ratings in light of the possible contextual influence. Finally, in Study 4, we tested the important assumption that the theories of bias generated by each individual predict that person’s unique adjustments in target ratings under conditions that encourage corrections.

Study 1

Do people possess theories that the same context can have differing influences (i.e., assimilation vs. contrast) on ratings of different targets, or do people tend to believe that a given context will have basically the same influence on all judgments? People might believe that contexts generally have one effect on judgments (i.e., that there are assimilation contexts and contrast contexts). On the other hand, people might believe that a given context has one default (i.e., uncorrected) effect on judgments of target x but has another default effect on judgments of target y: If one can identify specific combinations of contexts and targets for which people believe that uncorrected contrast versus assimilation effects are likely, such combinations could be used to determine whether people flexibly correct for these perceived influences. That is, one could determine whether people make corrections aimed at removing perceived biases according to the differing naive theories about the biasing effect the context has on the judgment. To identify possible contexts for which assimilation and contrast theories exist for different targets, we described various contexts and judgment targets to
participants and asked them what the natural (i.e., uncorrected) effect of these contexts would be on people’s target ratings.

Method

Participants and Procedure

Two sets of individuals (each consisting of 31 undergraduate psychology students) participated in partial fulfillment of a class requirement. The first set of participants received a questionnaire describing a context and judgment to be made. Then the context was repeated with a second judgment. The second set of participants received a questionnaire describing two contexts, each paired with two different judgments. For both sets, instructions explained that contexts can sometimes create biases in people’s judgments and that we wanted to find out what people thought about how some specific contexts might affect their ratings. Participants were asked to decide what effect they thought that each context would tend to have on people’s initial perceptions (i.e., whether it would make people see the target as more favorable or less favorable).

For the first set of participants, the context was imagining the weather in desirable vacation locations. The two judgments were created such that past research led us to expect that theories of contrast might be likely for one judgment (i.e., participants might be likely to identify contrast as the probable default or uncorrected outcome), whereas theories of assimilation might be likely for the other judgment (i.e., participants might be likely to identify assimilation as the probable default outcome). The contrastive target, adapted from Petty and Wegener (1993), was making ratings of how desirable the weather would be in midwestern cities such as Indianapolis. Thus, for this judgment, we expected that participants might believe that thinking about the weather in desirable vacation spots would make the weather in midwestern cities seem worse than what would normally be the case (i.e., contrast would be the perceived default outcome). Unlike Petty and Wegener (1993), however, we also designed a target within the same context for which participants’ perceptions of default assimilation might be likely. That is, respondents were asked how thinking about the weather in the vacation locations would influence perceptions of how satisfied people are with their jobs in places like Hawaii (cf. Schwarz & Clore, 1983). For this judgment, we expected that participants might believe that thinking about the weather in desirable vacation spots would make people’s satisfaction with jobs in those locations seem higher than would normally be the case (i.e., assimilation might be the perceived default outcome).

For the second set of participants, one of the contexts was very similar to the context in the first set (i.e., imagining being in desirable vacation locations). The other context was imagining being very attractive actresses and models. As in Set 1, two judgments were created for each context. The two judgments were created such that past research led us to expect that theories of contrast might be likely for one target and theories of assimilation might be likely for the other target. The contrastive target for the vacation context, taken directly from Petty and Wegener (1993), was making ratings of how desirable midwestern cities such as Indianapolis would seem (cf. Dermer, Cohen, Jacobsen, & Anderson, 1979). For the attractiveness context, the contrastive target was rating the attractiveness of an average-looking woman (cf. Kenrick & Guiterrres, 1980). As in Set 1, we also designed a target for which participants’ perceptions of default assimilation were more likely. That is, for the vacation context, respondents were asked how thinking about the vacation qualities of the locations would influence perceptions of how satisfied people are with their jobs in places like Hawaii. For the attractive model/actress context, respondents were asked how seeing the attractive women would influence one’s perceptions of a product the women endorsed (cf. Petty, Cacioppo, & Schumann, 1983).

Theory Measures

Ratings of the probable effects of the contexts were tailored to the individual context and target. For the weather context (Set 1), participants’ perceptions of the hypothesized contrastive judgmental effects of rating the desirability of the weather in vacation spots like Jamaica, the Bahamas, or Hawaii were measured by asking how considering the weather in those places would influence one’s perceptions of how desirable the weather would be in midwestern cities like Indianapolis. Responses were made on a 9-point scale ranging from make weather in Indianapolis seem less desirable than if no vacation spots were considered (−4) to make weather in Indianapolis seem more desirable than if no vacation spots were considered (4). The hypothesized assimilative effects of thinking about the vacation weather were measured by asking how thinking about the weather in those vacation locations would influence one’s perceptions of how satisfied people are with their jobs in places like Hawaii. Responses were made on a 9-point scale ranging from make satisfaction with jobs seem lower than if weather weren’t considered (−4) to make satisfaction with jobs seem higher than if weather weren’t considered (4). The order of the two judgments in Set 1 was counterbalanced.

The wording of the vacation context in Set 2 varied only slightly from the vacation weather context in Set 1. Participants’ perceptions of the hypothesized contrastive judgmental effects of rating the desirability of vacation spots like Paris, the Bahamas, or Hawaii were measured by asking how considering those places would influence one’s perceptions of midwestern cities like Indianapolis. Responses were made on a 9-point scale ranging from make Indianapolis seem less desirable than if no vacation spots were considered (−4) to make Indianapolis seem more desirable than if no vacation spots were considered (4). The hypothesized assimilative effects of thinking about the vacation contexts were measured by asking how thinking about the vacation qualities of those locations would influence one’s perceptions of how satisfied people are with their jobs in places like Hawaii. Responses were made on a 9-point scale ranging from make satisfaction with jobs seem lower than if vacation qualities weren’t considered (−4) to make satisfaction with jobs seem higher than if vacation qualities weren’t considered (4).

Perceptions of the hypothesized contrastive effects of seeing a number of actresses and models generally rated as very attractive were assessed by asking how seeing the attractive women would influence one’s perceptions of an average-looking woman. Responses were made on a 9-point scale ranging from average woman would seem less attractive than if the attractive women weren’t there (−4) to average woman would seem more attractive than if the attractive women weren’t there (4). The possible assimilative effects of seeing a number of actresses and models who are very attractive and then learning that these people endorse a product were assessed by asking how seeing the attractive women would influence one’s perceptions of how desirable the product would be. Responses were made on a 9-point scale ranging from make the product seem less desirable than if the women were not attractive (−4) to make the product seem more desirable than if the women were not attractive (4). In Set 2, the order of presentation of contexts and targets was contrastive targets for the vacation and attractiveness contexts and then assimilative targets for the vacation and attractiveness targets for all participants.

Results and Discussion

Ratings of the perceived judgmental effects of each context were analyzed by testing the difference between the average rat-
ing for each context and the zero point of each scale (i.e., no perceived influence of the context) with the Student’s *t* statistic (with 30 degrees of freedom). Across both sets, the context–target combinations that were expected to lead to contrastive theories of bias did just that. That is, participants believed that rating the weather in vacation spots like Jamaica, the Bahamas, or Hawaii would make the weather in midwestern cities like Indianapolis seem less desirable (*M* = −1.55; *t* = 2.45, *p* < .002). Also, they believed that being in locations like Paris, the Bahamas, or Hawaii would make midwestern cities like Indianapolis seem less desirable (*M* = −.94; *t* = 2.17, *p* < .038) and that seeing a group of very attractive actresses/models would make an average-looking person seem less attractive (*M* = −1.03; *t* = 2.96, *p* < .006).

The same contexts led to assimilative theories of bias when they were paired with different target judgments. That is, participants believed that thinking about the weather in vacation locations would make people’s satisfaction with jobs in locations like Hawaii seem higher (*M* = 1.13; *t* = 2.68, *p* < .012). They also believed that thinking about the vacation qualities of locations like Paris, the Bahamas, or Hawaii would make people’s satisfaction with jobs in locations like Hawaii seem higher than if vacation qualities were not considered (*M* = 1.13; *t* = 3.20, *p* < .003) and that seeing a group of very attractive actresses/models would make a product endorsed by the actresses/models seem more desirable (*M* = .71; *t* = 2.06, *p* < .048).

Thus, it seems possible to construct situations such that people perceive the same context as having a positive effect on one judgment and a negative effect on another. That is, people might perceive contrast and assimilation effects to be the natural consequence of the same context on different respective judgment targets. If the flexible correction model is correct, this implies that different corrections can occur in response to different combinations of the same context with different target items. Flexible corrections that demonstrate this bidirectional flexibility are not derivable from models of correction processes that are not based on the theories of bias possessed by judges. Of course, simply showing that people believe that some context–target combinations lead to contrast effects and others lead to assimilation does not ensure that they flexibly correct for such varied contextual effects.

The flexible correction model predicts that people induced to correct for a context they believe would create a contrast effect will correct for that context by adjusting their ratings of the target back toward the context. On the other hand, people induced to correct for a context they believe would create an assimilation effect will correct for that context by adjusting their ratings of the target away from the context. This should occur even though the context is the same in both cases. To examine the hypothesis that people flexibly correct toward or away from the same context depending on the theory of bias associated with the context–target combination, we explicitly asked participants in some conditions of Studies 2 and 3 to attempt to keep contextual stimuli from influencing their judgments of targets.

**Study 2**

In Study 2 we used the same contexts and targets identified in Study 1 as associated with generally shared theories of bias. By using these context–target combinations, we were able to assess the bidirectional flexibility of corrections. Study 2 included three tests of flexible correction in which participants judged one of the three contexts followed by targets that were believed by participants to be either assimilated or contrasted to the context (based on Study 1). These context and target judgments took place either with no prompt for participants to consider the context’s potential effects on target judgments or with a prompt for participants to try not to let the context influence their ratings of targets (without alerting them to what that influence might have been).

**Method**

**Participants and Design**

Study 2 included three independent sets of participants. Each set of individuals was composed of undergraduate psychology students participating in the experiment in partial fulfillment of a class requirement, and each set received one of the contexts from Study 1. Forty-nine participants received the vacation weather context, 63 received the vacation desirability context, and 54 received the attractiveness context. Within each of these groups, participants were randomly assigned to a 2 (theory of default influence: contrast vs. assimilation) × 2 (correction: instruction vs. none) between-subjects design.

**Procedure**

Students in Set 1 participated in groups of 3 to 8; those in Sets 2 and 3 participated in classroom sessions of approximately 30 people each. For all participants, the first page of a packet of experimental materials contained all of the items for the experiment (the packets continued by presenting a number of additional tasks for other unrelated experiments). At the top of the page, a set of instructions informed participants that the experiment was “interested in people’s perceptions of a number of things.” Participants were asked to “read each of the judgment items carefully and to respond on the scale following each statement by circling the number that best represents your perception of the item.” The page included three context items and two target items. After completing the entire experimental packet, participants were thanked, debriefed, and dismissed.

**Independent Variables**

**Theory of default influence.** Participants began by rating three statements that formed the context for subsequent judgments. Within each of the three sets of participants, the same context was used for everyone in the set. For Set 1, the context was the desirability of the weather in each of three very desirable vacation locations (i.e., Jamaica, Hawaii, and the Bahamas). People were asked to rate the desirability of the weather in Jamaica, the Bahamas, and Hawaii on 9-point scales ranging from not at all desirable (1) to very desirable (9). After the three context items, participants encountered two target items. These items matched the category of targets believed by participants in Study 1 to be either naturally assimilated or contrasted to the given context. That is, for Set 1, the targets corresponding to a contrastive theory of bias were ratings of the desirability of the weather in midwestern cities (i.e., Indianapolis and Kansas City). The target items that corresponded to contrastive theories were rated on the same 9-point scales just described. The targets corresponding to an assimilative theory of bias were ratings of how satisfied people in Hawaii and the Bahamas are with their jobs.
Ratings of projected satisfaction were made on 9-point scales ranging from *not at all satisfied* (1) to *very satisfied* (9). For Set 2, the context was staying 2 weeks in very desirable vacation spots (i.e., Hawaii, Paris, and the Bahamas). The vacation locations were rated on 9-point scales ranging from *like as much as staying in Hawaii or Paris (like very much)* (1) to *dislike as much as staying in Cambodia or Iran (dislike very much)* (9). This "specific" scale was also used for ratings of the target items believed to be naturally contrasted to the context so that changes in meaning of the response scale could not easily account for the adjustments in target ratings across correction conditions (see Petty & Wegener, 1993). The targets corresponding to a contrastive theory of bias were ratings of how much people would like to stay in neutral midwestern cities (i.e., Indianapolis and Kansas City). The targets corresponding to an assimilative theory of bias were ratings of how satisfied people in Hawaii and the Bahamas are with their jobs. Ratings of projected satisfaction were made on 9-point scales ranging from *not at all satisfied* (1) to *very satisfied* (9).

For Set 3, the context was rating the attractiveness of very attractive actresses and models (i.e., Christie Brinkley, Michelle Pfeiffer, and Kim Basinger), each as identified by previous pretest participants. The attractiveness items were rated on 9-point scales ranging from *as unattractive as Phyllis Diller or Roseanne Arnold (not at all attractive)* (1) to *as attractive as Michelle Pfeiffer or Kim Basinger (very attractive)* (9). For the attractiveness context, the targets believed to be naturally contrasted to the context were two people rated by pretest participants as rather average in attractiveness (i.e., Hillary Clinton and Tipper Gore). The targets believed to be naturally assimilated to the context were ratings of how desirable two products endorsed by Michelle Pfeiffer and Kim Basinger, respectively, were likely to be. Ratings of perceived product desirability were made on 9-point scales ranging from *not at all desirable* (1) to *very desirable* (9).

**Correction manipulation.** Participants in all three sets either rated the three context items and then the two targets (no instruction) or rated the three context items and then were asked to "please try to make sure that your perceptions of the __ do not influence your ratings of __." The first blank was replaced by "weather in the vacation spots above" for Set 1, by "vacation qualities of the locations above" for Set 2, and by "attractiveness of these people" for Set 3. The second blank was also tailored to the context and targets involved. For Set 1, the second blank was replaced by "the weather in the following places" (in contrastive theory conditions) or "how satisfied people are with their jobs in these locations" (in assimilative theory conditions). For Set 2, the second blank was replaced by "the following places" (in contrastive theory conditions) or "how satisfied people are with their jobs in these locations" (in assimilative theory conditions). For Set 3, the second blank became "the attractiveness of the following people" (in contrastive theory conditions) or "how desirable a product might be if it is endorsed by these people" (in assimilative theory conditions). The target items at the bottom of the first page of the experimental packet actually formed the end of the experimental materials for the study.

**Hypotheses.**

On the basis of the preliminary theory identification study, participants were expected to believe that rating contexts that were extremely positive objects on the same dimension of judgment as the targets (e.g., rating weather in Hawaii before rating weather in Indianapolis) would bias their default reactions to be less positive than would normally be the case (i.e., would lead to contrast). On the other hand, participants were expected to believe that rating these same context items would bias their default reactions of the targets on different but related dimensions of judgment to be more positive than would normally be the case (i.e., would lead to assimilation). Thus, for each set of participants, a Theory × Correction interaction was expected such that correction prompts would lead to less positive target ratings when participants believed that context ratings would lead to assimilation and more positive ratings when they believed that context ratings would lead to contrast in uncorrected settings.

**Results.**

For each set, ratings of the two targets were summed to form the primary dependent measure. These ratings were submitted to a 2 (theory of default influence: contrast vs. assimilation) × 2 (correction: instruction vs. no instruction) between-subjects analysis of variance (ANOVA). As explicated in the following sections, the expected Theory × Correction interaction was obtained within each set of participants (each of whom received a different set of contexts and targets).

**Set 1: Vacation Weather Context.**

Results showed only a significant Theory × Correction interaction, *F*(1, 45) = 4.35, *p* < .043. Participants' corrected judgments were closer to the context (*M* = 10.23) than when no correction instruction was given (*M* = 8.58) for targets believed to be naturally contrasted; however, participants' corrected judgments were farther away from the context (*M* = 10.17) than when no correction instruction was given (*M* = 11.42) for 7

---

7 One possible cause of shifts in ratings does not depend on the operation of theories of bias. This cause of rating adjustments is shifts in perspective or response language associated with the response scale (see Ostrom & Upshaw, 1968; Schwarz & Bless, 1992). That is, if participants use the positive contextual items to define the positive content endpoint of the response scale (i.e., if they use Hawaii-like weather to define very desirable weather), the meaning of this endpoint might be more positive than the person's usual idea of desirable weather. If this were so, judges' ratings of the desirability of weather in Indianapolis could be pushed lower on the scale than would normally be the case, only to return to their normal position when participants were asked to not let their perceptions of the context items influence their ratings of the targets (see Petty & Wegener, 1993, for a more thorough discussion of this alternative to theory-based correction).

There are a number of reasons to believe that such shifts in response language are not responsible for all corrections in target ratings along contextual dimensions of judgment. Chief among those reasons is that corrections for contrast along the same dimension of judgment as that used in context ratings have been shown when redefinition of scale end-point meaning is highly unlikely. That is, Petty and Wegener (1993, Experiment 3) used a response scale with "specific" endpoints (i.e., endpoints that included some of the contextual items to define that scale) whose meaning was quite clear and constant across correction conditions (see also Harvey & Campbell, 1963; Krantz & Campbell, 1961; Manis, Biernat, & Nelson, 1991). A prompt to consider the possible influence of the context on target ratings led to corrections of a magnitude equal to that observed using "abstract" scale endpoints in which redefinition of the scale was more likely. Thus, at least in one study, corrections along the contextual dimension of judgment have not appeared to be due to scale redefinition. In the current study, we used specific anchors like those of Petty and Wegener (1993, Experiment 3) to anchor response scales along the dimension of judgment used for context ratings in Sets 2 and 3. Specific endpoints were not used for ratings along dimensions that differed from the contextual dimension of judgment because such shifts in judgment dimension would not be expected to lead to expanded perspective along the new dimension of target rating (see Martin, 1986; Upshaw & Ostrom, 1984).
targets believed to be naturally assimilated. Neither of these simple effects reached conventional levels of significance when the pooled error term from the $2 \times 2$ ANOVA was used (see Table 1, however, for separate significance levels involving only the participants within each theory condition). Ratings of the context were not affected by any of the manipulated variables (all $F$s < 1).

**Set 2: Vacation Desirability Context**

Once again, results showed only the predicted Theory $\times$ Correction interaction, $F(1, 59) = 4.12, p < .047$. Participants' corrected judgments were closer to the context ($M = 9.29$) than when no correction instruction was given ($M = 7.53$) for targets believed to be naturally contrasted; however, their corrected judgments were farther away from the context ($M = 8.53$) than when no correction instruction was given ($M = 10.36$) for targets believed to be naturally assimilated. Neither of these simple effects reached conventional levels of significance. Ratings of the context were not affected by any of the manipulated variables (all $F$s < 1).

**Set 3: Attractiveness Context**

In this set, results showed that targets believed by participants in the preliminary study to be naturally assimilated to the positive context were rated more positively ($M = 11.25$) than targets believed to be naturally contrasted to the positive context ($M = 8.57$), $F(1, 50) = 8.38, p < .006$. More important, there was also a significant Theory $\times$ Correction interaction, $F(1, 50) = 6.89, p < .012$. Participants' corrected judgments were closer to the context ($M = 10.13$) than when no correction instruction was given ($M = 7.0$) for targets believed to be naturally contrasted; however, their corrected judgments were farther away from the context ($M = 10.38$) than when no correction instruction was given ($M = 12.13$) for targets believed to be naturally assimilated. Only the correction under the contrastive context was significant ($p < .03$). Ratings of the context were not affected by any of the manipulated variables (all $F$s < 1).

**Summary of Corrections in Study 2**

Across three sets of participants, corrections for the perceived effects of a judgmental context were consistently in a direction opposite to the rated direction of uncorrected influence identified by participants in our preliminary study. Table 1 displays the directions, $Z$ scores, directional probabilities (using only the participants in the conditions associated with the test), and sample size for each of the corrections evident in each set of participants from Study 2. When we combined the corrections for targets identified as naturally assimilated to the context presented (i.e., Tests 1, 3, and 5) using the Mullen (1989) basic meta-analysis program, the $Z$ score associated with the overall correction for perceived assimilation was $-2.20$ ($p < .014$). As the sign of the $Z$ score denotes, the direction of correction was away from the positive context, as predicted. The overall effect sizes for this correction were as follows: $r = -.244$ and $d = -.502$ (see Rosenthal, 1991). When we combined the corrections for targets identified as naturally contrasted to the context presented (i.e., Tests 2, 4, and 6), the $Z$ score associated with the overall correction for perceived contrast was $3.14$ ($p < .0008$). As the sign of the $Z$ score denotes, the direction of correction was toward the positive context, as predicted. Overall effect sizes for this correction were as follows: $r = .367$ and $d = .750$. A comparison of the difference in corrections for judgments believed to be assimilated versus contrasted (based on significance levels of the tests) yielded a $Z$ score of $3.85$ ($p < .0001$). A comparison of the difference in corrections based on the effect sizes of the corrections yielded a $Z$ score of $3.78$ ($p < .0001$; see Rosenthal, 1991). Although participants were not randomly assigned across the contexts, tests of homogeneity within theory conditions showed that, across the three sets, they did not act differently from one another. Thus, the three groups of participants in Study 2 showed consistent support for our hypothesized pattern of results.

**Discussion**

Study 2 provides initial support for the existence of opposite corrections for the same context that correspond to measured theories of how the context would influence target perceptions. That is, when participants were confronted with a context and target for which contrast was the perceived uncorrected impact of the context, prompts to consider the possible impact of the context led to adjustments that moved target ratings closer to the context than when no prompt was present. On the other hand, when they were confronted with a context and target for

<table>
<thead>
<tr>
<th>Test (set, condition)</th>
<th>Correction direction</th>
<th>$Z$ for correction</th>
<th>Directional probability</th>
<th>$n$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 (Set 1, assimilation theory)</td>
<td>Away from context</td>
<td>$-1.04$</td>
<td>.1497</td>
<td>24</td>
</tr>
<tr>
<td>2 (Set 1, contrast theory)</td>
<td>Toward context</td>
<td>$2.04$</td>
<td>.0206</td>
<td>25</td>
</tr>
<tr>
<td>3 (Set 2, assimilation theory)</td>
<td>Away from context</td>
<td>$-1.41$</td>
<td>.0811</td>
<td>29</td>
</tr>
<tr>
<td>4 (Set 2, contrast theory)</td>
<td>Toward context</td>
<td>$1.41$</td>
<td>.0791</td>
<td>34</td>
</tr>
<tr>
<td>5 (Set 3, assimilation theory)</td>
<td>Away from context</td>
<td>$-1.33$</td>
<td>.0925</td>
<td>28</td>
</tr>
<tr>
<td>6 (Set 3, contrast theory)</td>
<td>Toward context</td>
<td>$2.19$</td>
<td>.0143</td>
<td>26</td>
</tr>
<tr>
<td>Combined assimilation (Tests 1, 3, and 5)</td>
<td>Away from context</td>
<td>$-2.20$</td>
<td>.0140</td>
<td>81</td>
</tr>
<tr>
<td>Combined contrast (Tests 2, 4, and 6)</td>
<td>Toward context</td>
<td>$3.14$</td>
<td>.0008</td>
<td>85</td>
</tr>
</tbody>
</table>
which assimilation was the perceived uncorrected impact of the context, prompts to consider the possible impact of the context led to adjustments that moved target ratings farther away from the context than when no prompt was present. These adjustments led to the significant Theory × Correction interactions observed.

Study 3

In Study 2, the corrections were all in the direction predicted by the flexible correction model, and the corrections toward and away from contrastive and assimilative contexts were significant when combined meta-analytically. Yet, one could gain further confidence in the theory-based correction framework if stronger corrections toward and away from the context were found in one set of participants randomly assigned to conditions. Given the small number of participants in each cell of the previous experiment, theories of assimilation and contrast might not have been strong enough to produce significant corrections. Thus to enhance power over our three small sets of participants from Study 2, we conducted another experiment with a larger number of participants. In addition, in Study 3, we took a measure of participants' initial ratings of the targets (before the context induction) so that we would have a baseline of their perceptions against which to compare the effects of context and correction.

Method

Participants and Design

One hundred eighty-eight undergraduate psychology students participated in the experiment in partial fulfillment of a class requirement. Participants received experimental packets that were randomly assigned to a 2 (theory of default influence: contrast vs. assimilation) × 2 (correction: instruction vs. no instruction) × 2 (context: vacation locations vs. attractive actresses/models) mixed design; theory and correction were between-subjects variables and context was a within-subject variable.

Procedure

Students participated in one of four large classroom sessions consisting of approximately 40–50 people per session. Procedures were similar to those used in Study 2, with two exceptions. First, at the beginning of the experimental packet, the same target items as presented in the experimental context were presented with no context items so that participants' perceptions of the targets could be assessed outside the experimental judgment context (and could thus be used as a baseline against which to compare the effects of the experimental context and participants' corrections for the context).

Thus, participants first completed the target items that they would later complete within a related context. Next, they completed filler ratings of unrelated targets. The filler ratings took approximately 5 min. Finally, they rated the context and target items from the vacation context and then from the attractiveness context (the within-subject manipulation of context). As in Study 2, participants were either requested to not let the context influence their target ratings or were given no such instruction just before responding to the target items.

The instructions, context and target items, and measures were identical to those used in the vacation desirability and attractiveness sets from Study 2, with the exceptions noted earlier (i.e., context was manipulated within subjects and participants responded to initial target ratings outside the judgmental context, followed by filler items). All of the materials for this study were contained in the first portion of a larger experimental packet. After participants had completed the entire packet, they were thanked, debriefed, and dismissed.

Results

Ratings of the two targets that followed the context items were summed, as were the ratings of the same targets completed at the beginning of the packet before any judgmental context had been introduced. The ratings of the targets presented outside the judgment context were subtracted from the ratings of the targets presented inside the judgment context to form the primary dependent measure. Thus, the primary dependent measure indexed the magnitude of the shift in ratings induced by the context and by correction processes. This measure (constructed within each of the contexts) was submitted to the 2 × 2 × 2 ANOVA described earlier.4

Results for the shift in target ratings measure showed two effects. First, a main effect of theory, F(1, 184) = 4.74, p < .04, indicated that targets believed to be contrasted created more positive shift scores than targets believed to be assimilated. More important, the expected Theory × Correction interaction, F(1, 184) = 10.48, p < .001, also emerged (see Figure 1).

4 The same ANOVA was also run on the baseline measures (i.e., the prereflection measures of target perceptions). Not unexpectedly, there was a main effect of theory, F(1, 184) = 108.01, p < .0001, such that targets to be assimilated to the context were originally viewed more positively (M = 12.24) than targets to be contrasted to the context (M = 8.70). In addition, there was a Context × Theory interaction, F(1, 184) = 15.29, p < .0001, such that the baseline ratings of targets within the attractiveness context differed more from one another (Ms = 12.82 and 8.18 for the targets to be assimilated and contrasted, respectively) than the baseline ratings of targets within the vacation context (Ms = 11.68 and 9.23 for the targets to be assimilated and contrasted, respectively). Thus, to the extent that the primary dependent measure (which controls for individuals' baselines) shows similar results across the within-subject context variable, the correction results can be interpreted as robust with respect to differential extremity of the baseline measures.
For participants who received targets believed to be assimilated, the difference between ratings of context-embedded targets and context-independent targets was less positive when a correction instruction was given ($M = -0.84$) than when no correction instruction was given ($M = 0.37; p < .039$, one-tailed). However, for participants who received targets believed to be contrasted, the difference between ratings of context-imbedded targets and context-independent targets was more positive when a correction instruction was given ($M = 1.71$) than when no correction instruction was given ($M = -0.13; p < .003$, one-tailed). This interaction was the same for both the vacation and attractiveness contexts ($F < 1$ for the three-way interaction).

**Discussion**

Study 3 provides consistent evidence supporting the flexible correction notion that people correct in opposite ways for the same context so long as they hold opposite theories about how the context affected their judgments. We also found, for the first time, that these corrections can result in overcorrection of targets believed to have been naturally contrasted, leading to corrected assimilations of the targets. That is, shifts in target ratings when correction instructions were present led to shifts in target ratings that became more positive than ratings of the targets outside any context (i.e., the shift score under correction conditions for targets believed to have been contrasted was significantly higher than zero, or accurate correction: $p < .001$). Perhaps such overcorrection occurred in this case because the contexts alone (i.e., no correction conditions) did not markedly shift target perceptions away from the precontext ratings participants had provided earlier in the session. This might have been due to rating the same targets twice within the experimental setting, although the presence of two ratings did not keep participants from shifting far enough to be captured under correction instructions. This might also have been due to the use of relatively few context ratings.

The results of this study also nicely illustrate that corrections are made for the expected or perceived bias rather than for the actual bias present in the situation. That is, although participants significantly corrected in directions consistent with the shared theories of bias identified in the preliminary study, they did so in a situation in which the actual bias created by the contexts was negligible. Especially because of the overcorrection we observed and because of our use of specific endpoints in target ratings that followed the same dimension as context ratings, the current correction results are not easily accounted for by changes in response language or by models that rely on "ignoring" the contexts in our correction conditions (e.g., Schwarz & Bless, 1992; see Petty & Wegener, 1993, for a discussion).

In sum, the current results constitute strong evidence of the flexibility of the theory-based correction processes that can operate in social judgment settings. The current experiments documented opposite corrections that followed opposing normative theories of how context influences perceptions of targets. In addition, we have shown that people can flexibly hold opposite theories of bias for the same context and that these opposite theories then correspond to the corrections for this context. In Studies 2 and 3, these flexible corrections were consistent across four groups of participants, and across three different contexts.

A critic might argue, however, that we have yet to demonstrate the crucial idiographic role of theories of bias in these flexible correction processes. That is, although it is a reasonable assumption, we have no evidence that the participants in Studies 2 and 3 held the same theories of bias as did those in our theory identification study (Study 1); therefore, we have not shown that an individual’s theory of how a context has affected his or her perceptions of the target drives corrections for that context. Also, consideration of an individual’s theory of the biasing effect of a context allows for prediction of the magnitude as well as the direction of a correction. For example, a person who thinks that attractive endorsers of a product have a large biasing impact on judgments of the product should show a larger correction for the attractive endorsers than should a person who thinks that attractive endorsers have a small biasing impact. In Study 4, we measured participants' unique theories of bias and later exposed them to target judgments within conditions encouraging correction of target ratings for perceived context-induced bias.

**Study 4**

In our final study, we aimed to examine whether participants corrected for the magnitude and direction of their own theories.

---

9 One of the reviewers raised a possible alternative explanation for the results of Study 3. That is, because the baseline values of the two sets of targets (i.e., the targets participants believed would be assimilated vs. contrasted to the context) differed from one another, this reviewer proposed that aiming the perceptions of targets twice might have led to polarization of the perceptions (Lesser, 1978) in the no correction instruction condition and that this polarization (rather than perceived effects of the context) was moderated in the correction instruction conditions. There are a number of reasons not to favor this interpretation. First, this explanation does not account for obtaining the same pattern of data in Study 2 even though only one measure of the targets took place. In addition, the correction-for-polarization alternative does not fully account for the pattern of data in the current study. To test this alternative empirically, we classified participants within the perceived contrast conditions according to whether their average baseline responses (combined across the midwestern city and average attractiveness targets) were above or below the midpoint of the rating scale (because none of the combined baseline measures for targets in the perceived assimilation conditions fell below the midpoint, the following analysis was impossible for that condition). According to the correction-for-polarization alternative, the effects of the correction instruction should differ for these two groups, with corrections for polarization being consistent with our hypotheses in the negative baseline group but corrections for polarization being in the opposite direction of our hypotheses in the positive baseline group. According to our theory-based correction view, however, both groups should correct by moving target ratings toward the context (because both groups are correcting for perceived contrast). In fact, when shift scores for these participants were submitted to a 2 (correction: instruction vs. no instruction) × 2 (baseline positivity: positive vs. negative) ANOVA, the results strongly supported our theory-based view. There was a strong main effect of correction, $F(1, 95) = 8.44, p < .005$, and no Correction × Baseline interaction ($p > .17$). That is, for targets with either positive or negative baseline ratings, correction instructions led to more positive target shifts than lack of correction instructions. In fact, the nonsignificant interaction was actually in the direction of correction instructions leading to greater positive shifts for targets with positive baseline ratings than for targets with negative baseline ratings (which was opposite to the correction-for-polarization alternative).
of how context influences target perceptions. If individuals' theories of bias drive corrections, then regressing participants' corrected target judgments on their theories of bias should yield a significant negative relationship. That is, the more negative (or less positive) a person's theory of how the context might bias his or her perceptions of the target, the more positive (or less negative) the person's shift in target ratings should be when attempting to correct for the perceived bias. Thus, negative theories of bias should lead to more positive corrections of target ratings than should positive theories of bias (as found in Studies 2 and 3). Furthermore, within valence of theory (e.g., within negative theories), more extreme theories should lead to larger corrections (i.e., adjustment for the perceived magnitude of bias).

Method

Participants

Forty-four undergraduate psychology students participated in the study in partial fulfillment of a class requirement. Individuals were recruited for the study from a pool of participants who had completed pretesting materials 4 to 7 weeks earlier. During this pretesting, participants completed all of the target items used in the correction portion of the study with no context preceding the target ratings (as in Study 3).

Procedure

Study 4 involved a multiple-experiment format in which separate experimenters (one woman and one man) administered each portion of the session under the guise of the session containing three separate experiments. For the "first experiment," participants received the same measurement of theories of bias as used in the second set of Study 1 (i.e., the vacation desirability and attractiveness contexts). For example, participants were asked to estimate how thinking about desirable vacation locations (e.g., Paris, the Bahamas, or Hawaii) would influence their perceptions of midwestern cities like Indianapolis. The one-sheet assessment of participants' theories of bias from Study 1 was followed by an additional sheet of assessments of perceived biases unrelated to the theories of bias on the first page. For example, one of these questions asked participants how they thought static (i.e., white noise) on an audiotaped message would affect their view of the message. After participants had completed this packet, the female experimenter collected the packets and left the laboratory.

Participants next engaged in a filler task introduced by a new experimenter who entered the room as the first experimenter departed. The filler task was reading and scrutinizing a counterattitudinal persuasive message on the topic of instituting senior comprehensive exams at the participants' university. After reading the message, participants reported their attitude on the topic; they also reported their perceptions of the desirability of each of the consequences of instituting senior comprehensive exams noted in the arguments, as well as the likelihood that each consequence would occur if the exams were implemented. The persuasion task took approximately 10 min to complete. To enhance participants' perceptions that the new experiment was separate from the first, we used, in the materials for the filler task, a print font quite different from that used in the theory measurement task.

Finally, when participants had completed the filler task, the second experimenter gave them one last sheet for a "different experiment." This sheet contained ratings of one of the context-target combinations used in Studies 2 and 3. That is, participants received the vacation desirability context with ratings of the desirability of midwestern cities. All of the ratings used the same response scales as in Studies 2 and 3 and were completed with the same correction instructions used in those experiments. The print font used for the correction packet was the same as that for the filler task (which was different from that used for the measure of theories). When participants had completed the correction sheet, the second experimenter announced that one of the factors that can influence people's ratings on these kinds of tasks is what each person believes is being studied. The experimenter continued by asking participants to turn over their sheets and to write on the back what they believed was being studied in each of the two tasks he had given them. After participants had completed these activities, they were thanked, debriefed, and dismissed.

Before analysis, participants' free responses were coded for suspicion of connections between the theory measurement task and the correction judgment task that used one of the contexts described in the theory measurement task. None of the 44 participants noted any relationship between the earlier theory measurement task and the correction judgment task. That is, although a number of people generated possible connections between the persuasive message (i.e., the "second" experiment) and the correction judgment task (e.g., looking at persuasion through presentation of external information versus persuasion through a person's own perceptions), none mentioned any role of the materials or questions presented by the first experimenter in the correction judgment task. In fact, no participants even noted that correction or perceptions of bias had anything to do with the judgment task. Many simply speculated that the correction judgment task involved looking at how average objects are perceived when viewed together with desirable objects.

Results and Discussion

As in Study 3, the difference between ratings of the pair of targets following the context and the same ratings completed outside the context (i.e., in the prescreening session several weeks earlier) was computed as an index of each participants' shift in target ratings. Regression analyses were performed to test whether idiographic theories of bias reported by participants predicted the corrections in which those participants engaged when conditions encouraged corrections. For this analysis, the shift index was regressed on participants' theories of bias for the context. The regression coefficient for participants' theories indicated the prediction of their shifts in target ratings from their theories of bias (with the shift index controlling for participants' perceptions of the targets outside the experimental context). Our hypothesis was that the more a participant believed that the context would exert a negative impact on his or her judgments, the more the shift from prescreening to postcontext judgment should be positive when attempting to correct for the bias.

Results showed that theories of bias significantly predicted participants' shifts in target ratings ($B = -.59, T = -3.57, p < .001$). As predicted, the regression coefficient was negative, indicating that as theories of bias became more negative, shifts in target ratings became more positive. Within this overall analysis, there was evidence of correction for both direction and magnitude of the perceived bias. That is, if the theory predictor is replaced by a dichotomous variable denoting direction of perceived bias (positive or negative), this variable accounts for significant variance in the positivity of participants' shifts in target ratings ($B = -.75, T = -2.77, p < .01$). That is, negative theories of bias lead to more positive shifts in target ratings than do positive theories of bias.
In addition, results showed evidence of participants correcting to a greater extent as theories of bias become more extreme (controlling for the direction of bias). This was tested by "folding the regression line" at the zero point along the theory variable. That is, for people who reported positive theories of bias (which differed from the normative theory of bias reported by participants in the preliminary theory identification study; \( n = 10 \)), the value of the reported theory was reverse scored. In addition, the difference between the shift score for each person and the shift score predicted at the intercept of the regression line identified in the analysis reported earlier (i.e., the value of the shift score predicted by the regression line at a theory value of zero; intercept = 0.807) was added to the intercept value. Thus, for example, if a person reporting a theory of 1 had a shift value of -1 (i.e., 1.807 below the intercept), then the shift value for the magnitude analysis would be 0.807 + 1.807, or 2.614. Therefore, shift values for participants reporting positive theories of bias were recoded to take on the same difference between the intercept and their original value but were now on the opposite side of the intercept value. The magnitude analysis of these data included the recoded theory variable (now representing theory extremity and ranging from -4 to 0), a direction variable denoting whether the theory was originally positive or negative (in two separate analyses, this variable included zero theory values in either the positive or negative category), and the interaction between theory and direction. Thus, if participants corrected for the perceived magnitude of the bias, this would be evidenced by a negative regression weight for the theory variable (denoting that as theories become more extreme — coded as negative — shifts in ratings become more extremely positive). Analyses including the zero theories in the positive and in the negative direction category supported the correction of target ratings for the perceived magnitude of the bias. That is, the regression weight for the theory variable was significant and negative in both analyses (\( B = \cdot 78 \) and \( \cdot 80 \), one-tailed \( p = \cdot 02 \) and \( \cdot 015 \), respectively). Also, in both analyses, there was no difference in the regression weight for positive as opposed to negative theories (Direction × Theory Extremity interaction \( F = \cdot 00 \) and \( \cdot 01 \), \( p = \cdot 98 \) and \( \cdot 92 \), respectively). Deleting the zero theories also did not significantly change the results.

Study 4 provides the first evidence linking individuals' idiosyncratic theories of bias to their corrections for the contexts associated with those perceived biases. These results are consistent with the findings of Studies 2 and 3 in that direction of the perceived bias accounts for significant variance in corrections (with more negative theories leading to more positive shifts in target ratings). Thus, if one assumes that participants in Studies 2 and 3 possessed theories of bias that were of a similar distribution to those reported by participants in Study 1, then Studies 2 and 3 reflect the same idiographic corrections based on theories of bias as shown in Study 4. In addition, however, Study 4 provides the first evidence of idiographic corrections varying in magnitude according to the extent to which people think that a context will bias perceptions of the target.

General Discussion

The current set of studies provides evidence for a variety of flexibilities in corrections of social judgments. In Study 1, we showed that people flexibly can possess opposite theories of bias, even for the same contextual factor influencing different judgmental targets. Studies 2 and 3 provided the first evidence of corrections for the same context proceeding in opposite directions according to the theories people hold (as identified in Study 1). These theory-based opposite corrections were demonstrated across a variety of judgment contexts and targets. In addition, Study 3 (using target measures outside the judgment context) provided the first evidence for significant assimilation effects resulting from theory-based corrections (i.e., overcorrections) for perceived contrast. Finally, Study 4 provided evidence of the idiographic nature of theory-based correction processes. That is, in Study 4, the direction and magnitude of participants' own reported theories of bias predicted their shifts in target ratings.

Thus, the current corrections seem quite different from the kinds of corrections studied within frameworks that do not focus on theory-based corrections. For example, a reset-based correction (in which contextually activated reactions are suppressed or partialed out of reactions to the target; e.g., Martin, 1986) would seem to have difficulty accounting for a correction that results in target ratings becoming more similar to reactions to the context. That is, the term set refers to "the use of a contextually activated response during formation of the target impression," and "when the contextual response is brought to bear [in forming the target impression], the evaluation of the target is assimilated toward the implications of the contextual stimuli" (Martin, 1986, p. 495). In addition, the term reset "refers to the suppressed use of the contextually activated response and the generation of a context-distinct response for the target" (Martin, 1986, p. 495). In resetting, "when the contextual response is not brought to bear [in forming the target impression], and the individual generates a context-distinct response, the evaluation of the target is contrasted with the implications of the contextual stimuli" (Martin, 1986, p. 495). Thus, because "subjects cannot use their reaction to the priming stimuli if their objective is to give their reaction to the target" (Martin & Achee, 1992, p. 210), reset-based corrections result in reactions to the target being less like reactions to the context than when setting occurs.

---

10 The same effects are found if magnitude of shift is simply indexed by the absolute value of the shift measure.

11 Testing the simple effects within the nonsignificant interaction of theory extremity and theory direction, one would find that the regression weight was similar for negative theories (i.e., theories ranging from -4 to 0), \( B = \cdot 80 \), and positive theories (i.e., theories ranging from 0 to 4), \( B = \cdot 43 \). Perhaps not unexpectedly, only the weight for negative theories (\( n = 34 \)) was significant (\( p < \cdot 023 \), one-tailed). Given the small sample size (\( n = 13 \)), the weight for positive theories was not significant (\( p < \cdot 18 \), one-tailed).

12 A reviewer noted that although our results are incompatible with a set-reset contrast effect (see Martin, Seta, & Crelia, 1990, Figure 1), they might be compatible with the schematic representation of the set-reset model (see Martin et al., 1990, Figure 2). In this representation, no particular default outcome is specified. That is, the default portion of the model is not labeled as assimilate target judgments. In fact, in such presentations of the set-reset model, the term set does not even appear, although the possibility of other default outcomes has not been given particular attention. If other default processes routinely operate, however, it is unclear why the name of the model reflects only the set.
One important asset of the flexible correction perspective is that it has the potential to go beyond the social judgment paradigm used in the present research. That is, theories of bias probably exist for many factors that do not fit easily into the judgment paradigm used in much past research on correction processes. Furthermore, when any perceived biasing factor is present in a context in which assessments of the true qualities of a target are sought, the correction process driven by theories of bias might proceed in a similar manner even though no rating of the biasing context takes place. For example, one factor that at least some people believe can bias perceptions of a target is noise in the environment when the target is encountered (e.g., see Nisbett & Wilson, 1977). To the extent that this notion becomes salient (i.e., is accessible), people should correct their ratings of the target in a direction opposite to the perceived bias associated with the environmental noise. This should occur even though people’s perceptions of the effect of such a factor might be quite inaccurate (see Nisbett & Wilson, 1977).

The theory-based flexible correction framework also has potentially important implications for integrating other traditional areas of social psychology that include corrective processes. For example, some recent models of attribution explicitly posit corrective processes as being the route by which relatively automatic associations of traits with targets are overridden by consideration of situational constraints on behavior. For example, Gilbert, Pelham, and Krull (1988) noted that “attributions are a product of dispositional inferences that are followed by situational adjustments” (p. 738; see also Gilbert & Osborne, 1989; Osborne & Gilbert, 1992, Quattrone, 1982). As noted by Petty and Wegener (1993), however, assuming that situational attributions invariably come about as a result of corrections of dispositional attributions is analogous to assuming that contrast is the result of corrections of default assimilative effects of contexts. Applying more flexible corrections in attribution contexts leads to the differing hypothesis that either situational or dispositional attributions can be the default outcomes and, thus, either situational or dispositional attributions can be the result of effortful corrective processes.

Consider the implications of these views for data recently reported by Morris (1993). Morris found that Chinese attributors weighted dispositional causes as less important and situational causes as more important than American attributors; he also found that a situational bias extended to attributions for murder in Chinese newspapers, whereas a dispositional bias extended to attributions for murder in American newspapers (for similar results using Hindu respondents, see Miller, 1984). Within the view of attribution espoused by Gilbert et al. (1988), these findings could be due to effortful correction generally taking place for Chinese but not American attributors. Within the flexible correction perspective, this could be the case; however, an alternative view also comes out of the flexible correction perspective. That is, according to our flexible correction model, either dispositional or situational attribution could be the default outcome, and either dispositional or situational attribution could be the result of effortful correction processes. Thus, in addition to the view of Chinese attributors as chronically active correctors—generated from the Gilbert et al. (1988) model—the flexible correction view also suggests that something common to Chinese attributors (e.g., the Chinese culture) might have made situational attributions the default (i.e., no-correction) outcome that can subsequently be corrected for dispositional influences. This notion should be addressed in future research.

In addition, it is likely that situational rather than personal factors could make situational attributions the default. That is, even with American attributors, some settings might encourage situational attributions for behavior, with information about the person having an impact only when attributors are able to effortfully consider the implications of the person’s character. Potential initial support for this notion was reported by Krull (1993). In Krull’s study, all participants watched anxious nonverbal behaviors performed by a target person. The results of the study replicated those of past work on corrections for dispositional attributions when participants were asked to focus on the extent to which the target person was anxious. That is, distracted participants rated the person as more anxious than nondistracted participants and rated the situation as less anxious. When participants were asked to focus on the extent to which the situation surrounding the target person was anxious, however, distracted participants rated the situation as more anxious than nondistracted participants and rated the person as less anxious. Thus, in the latter conditions, higher levels of cognitive ability were associated with less dispositional attributions of behavior. Thus, although participants in the Krull (1993) study did not receive any particular dispositional information on which to base their corrections, corrections of perceived default situational attributions might have been responsible for the observed effects.

Research on theory-based flexible corrections can take a variety of directions. For example, one factor that might prove useful in extending the flexible correction framework to corrections outside the assimilation–contrast literature is perceptions of the legitimacy or illegitimacy of bias. That is, even though a person might be fully aware of the perceived effect of a biasing factor, that bias might seem relatively legitimate in some settings or to some people but seem quite illegitimate in other settings or to other people. For example, people who score high versus low on measures of prejudice or racism might be equally
Theory-Based Correction

aware of the bias contained in the stereotype of a social group (e.g., Devine, 1989), but high scorers might view the bias in the stereotype as more legitimate than low scorers. Thus, high scoring judges might be unmotivated to correct their judgments away from the implications of stereotypic information even though they know that the information will make particular perceptions of the target more likely than if the information were not presented (i.e., that the information is biasing).

Similar processes might be at work in jury settings, or in research such as that conducted by Golding et al. (1990), in which situations surrounding judgments might influence people's perceptions of the legitimacy of the influence of certain kinds of information. For example, in Golding et al.'s (1990) study, participants might have believed that implications of confidential information for judgment are legitimate because they are valid and relevant to making accurate judgments of the target person.

In the case of false information, however, any implications of that information are not legitimately attributed to the target person, and, thus, judges are motivated to correct their assessments of the target away from the implications of the false information. Thus, one finds that judgments of targets after judges receive confidential information are close to those of judges who are not led to believe that the information is confidential, and judgments of targets by judges who receive false information are close to those of judges who receive no biasing information (Golding et al., 1990; Wyer & Budesheim, 1987). This reasoning also accounts for cases in which mock jurors continue to use evidence (such as past convictions for a similar charge) that appears relevant to judgments of probable guilt even though they are instructed to disregard the inadmissible evidence (e.g., W. C. Thompson et al., 1981).

Another of the directions for future research on theory-based corrections is to delineate the various processes by which such corrections might occur. For example, corrections directed by theories of bias might at times lead a person to seek out information about the target (stored either in his or her memory or in the judgment environment) that supports a judgment opposite to the person's initial "biased" reaction. In such cases, the magnitude of the search for additional information might be tied to the perceived magnitude of the bias at work. In other circumstances, however, theory-based corrections might proceed through reinterpretation of the information present at the time of the person's initial reaction (i.e., reinterpretation in light of the biased nature of the initial outcome). Also, to the extent that a potential bias is noted before target information is encountered, the theory of bias might guide initial scrutiny and information gathering regarding the target.

Adoption of a theory-based correction perspective also makes issues of theory formation and change particularly important. For example, stereotype-based biases might be one area in which individuals' perceptions of bias are quite stable over time. In the current research, measured theories of how a given judgment context might affect target perceptions were likely to be less stable and potentially short lived, especially as the participants gained further experience with that type of judgment setting. Therefore, further work in the area of bias correction is likely to profit from research investigating the factors that influence the generation and subsequent perseverance of theories of bias (e.g., Anderson & Sechler, 1986). In addition, the current work represents situations in which a single salient factor might or might not be associated with judgmental bias. In many settings, however, social judges are likely to face multiple possible factors that can influence their perceptions of targets. Therefore, future work might also investigate correction processes associated with identifying and adjusting for multiple factors in the judgmental context.

Our flexible correction model has the potential for organizing findings across paradigms and generating new hypotheses in many areas. Perhaps research and theory based on flexible correction notions will help to build a comprehensive and unifying framework within which correction processes across many areas of social psychology can be explained.

References


Higgins, E. T. (1989). Knowledge accessibility and activation: Subject-


Received December 7, 1993
Revision received May 18, 1994
Accepted May 18, 1994

---

Low Publication Prices for APA Members and Affiliates

**Keeping You Up-to-Date:** All APA members (Fellows; Members; Associates, and Student Affiliates) receive—as part of their annual dues—subscriptions to the American Psychologist and APA Monitor.

High School Teacher and International Affiliates receive subscriptions to the APA Monitor, and they can subscribe to the American Psychologist at a significantly reduced rate.

In addition, all members and affiliates are eligible for savings of up to 60% (plus a journal credit) on all other APA journals, as well as significant discounts on subscriptions from cooperating societies and publishers (e.g., the American Association for Counseling and Development, Academic Press, and Human Sciences Press).

**Essential Resources:** APA members and affiliates receive special rates for purchases of APA books, including the Publication Manual of the APA, the Master Lectures, and Journals in Psychology: A Resource Listing for Authors.

**Other Benefits of Membership:** Membership in APA also provides eligibility for low-cost insurance plans covering life, income protection, office overhead, accident protection, health care, hospital indemnity, professional liability, research/academic professional liability, student/school liability, and student health.

**For more information**, write to American Psychological Association, Membership Services, 750 First Street, NE, Washington, DC 20002-4242, USA