Piaget (1929) argued that young children fail to distinguish between the same transition that Western philosophy did, albeit more quickly. Reality only recently became a problem. Philosophers once thought of the mind as a mirror that reflected the images of objects in the world, which meant that people could generally trust their perceptions. Tigers caused people to see tigers, people saw tigers where tigers stood, and the philosopher’s job was to figure out how the mind accomplished this act of reflection without the use of polished glass. But Kant (1781/1965) traded the innocence of realism for the complexity of idealism, which held that “the world as we know it is a constructed article, to which the mind contributes as much by its moulding as by its stimuli” (Durant, 1926, p. 272). The perception of tigers, idealists argued, was caused by knowledge of tigers, memories of tigers, belief in tigers, expectations of tigers, and sometimes (but not always) by tigers, too. Minds did not passively reflect, they actively constructed, and thus knowing the real world was a real problem.

WHERE DOES REALISM GO WHEN IT GOES AWAY?

Psychologists have suggested that ordinary people make roughly the same transition that Western philosophy did, albeit more quickly. Piaget (1929) argued that young children fail to distinguish between their experience of an object and the object’s actual properties, but that with maturity they come to realize that this experience can be influenced by factors extraneous to the object itself. He concluded that “the child is a realist in its thought” and that “its progress consists in ridding itself of this initial realism” (p. 166). Like philosophers, ordinary people start out as realists but get over it soon enough. The only problem with Piaget’s argument was that most adults can be counted on to think like children under the right circumstances, which suggests that if realism goes away, it does not get very far. The ink in Piaget’s book was barely dry when theorists such as Isaacs (1930) suggested that Piaget was wrong, that idealism did not replace realism, but rather, that idealism joined realism in a varied repertoire of cognitive strategies (see Ruble & Goodnow, 1998). Unlike philosophers, adults could be realists one moment and idealists the next.

Recent work on the interaction of automatic and effortful processes (see Wegner & Bargh, 1998) points toward another possibility. People may neither outgrow realism nor revert to it on some occasions; rather, cognitive activities may normally be characterized by an initial moment of realism that is quickly followed by an idealist correction. According to this correction model, when people attempt to understand the objective properties of a stimulus, they automatically assume that their subjective experience of the stimulus is a faithful indicator of its properties, and then—if they have the time, energy, and ability—they rapidly “undo” that assumption by considering the possibility that extraneous factors may have shaped their experience.

Piaget (1929) described realism as “a spontaneous and immediate tendency to confuse the sign and the thing signified” (p. 124), and according to the correction model, this tendency to equate objects and their representations remains spontaneous and immediate throughout life. It does not go away forever, and it does not go away on occasion. Rather, it is brief, unarticulated, and rapidly unraveled, but it is always the first step in the person’s representation of the world.

Research in a variety of domains provides support for this claim. For example, research in visual perception suggests that when a person’s head is tilted and he or she is placed under cognitive load, the person will see the object as tilted instead (Rock & Nijhawan, 1989; see also Epstein & Broota, 1986; Epstein & Lovitts, 1985). One way to interpret this finding is that people initially construct realist representations of objects (“The top of the chair appears on the side of my retina, so this chair is tilted”) and only then correct for the influence of temporary factors, such as their own spatial orientation (“But I’m lying down”). Presumably, cognitive load impairs this second step. Investigations of the processes by which people evaluate the veracity of propositions (Gilbert, 1991), form impressions of others (Gilbert, in press; Martin, Seta, & Crella, 1990), and decode intentions in conversations (Keysar, Barr, Balin, & Paek, 1998; Keysar, Barr, & Horton, 1998) have documented similar tendencies. In each case, people who were under cognitive load or time pressure embraced realist interpretations of a stimulus—as if they initially believed their mental representations of the stimulus to be faithful reflections of its properties, and then failed to execute a second mental operation that would correct for the fact that their representations were “contaminated” by extraneous factors (Wilson & Brekke, 1994).
Whereas correction models suggest that people draw realist interpretations of stimuli and then make idealist corrections, selection models suggest that people arrive at realist interpretations only when they are under load or time pressure, and that they otherwise arrive at idealist interpretations without ever experiencing a moment of realism (Gilbert, 1999). Although selection models can account for the effects of time pressure and cognitive load after the fact, they fail to account for two other kinds of evidence. First, when realist interpretations would naturally give rise to an affective reaction, even unloaded and unhurried individuals experience that reaction, despite the fact that their judgments are perfectly idealist (Gilbert, Giesler, & Morris, 1995). Second, when the realist and idealist interpretations of an utterance would naturally cause persons to look in different directions, those who interpret the utterance ideally initially look in the realist direction (Keysar, Balin, & Paek, 1998; Keysar, Barr, & Horton, 1998). This evidence suggests that people do experience an initial moment of realism, as the correction account suggests and the selection account denies.

THE PROBLEM OF THE PERMANENTLY TILTED

Embracing and then repudiating one’s mental representations would seem to be a marvelously inefficient way for people to go about the business of understanding reality. So why do so many psychological systems seem to work this way? In fact, the correction mechanism is a very efficient tool for those systems in which correction is only rarely required. Heads are usually held in an upright position, and thus it makes sense for people to assume that the top of an object will appear at the top of the retina and then to correct for the occasional romp on the monkey bars. Utterances are usually meant to convey accurate information, and thus it makes sense for listeners to assume that what they hear is true and then correct for the occasional presidential address. Conversation partners usually share basic cultural knowledge, and thus it makes sense for each to assume that the other knows what he or she knows and then correct for the occasional tourist. In short, it makes sense to assume the usual and correct for the unusual, and it is difficult to imagine a heuristic that might deal more efficiently with a world composed of enduring realities whose representation is occasionally distorted by temporary factors.

Of course, not all the factors that distort representations are unusual or temporary. Indeed, the categories or “moulding forms” that forever separated Kant’s observers from the dinge an sich were thought to be permanent fixtures of the observers’ minds—not spectacles temporarily balanced on the bridge of their noses, but contact lenses permanently affixed to their eyeballs. For idealists, it was not the momentary tilt of an observer’s head that made reality so elusive; rather, it was the enduring, dispositional properties of the observer’s mind that exerted ubiquitous influence on its representations and thus left the observer unable to tell how much of his or her subjective experience had been “given” by the object and how much had been “added” by his or her brain. It may be easy for people to consider how their perceptions of the world might have been influenced by temporary factors such as inadequate lighting or the tilt of their heads, but not so easy for them to consider how their perceptions might have been influenced by enduring factors, such as their acculturation, personalities, intelligence, or temperaments.

If correction is efficient only in those instances in which people correct for temporary influences on their representations, then by what mechanism might they take account of the influence that their own enduring, dispositional properties exert? The possibility explored in the present article is that the mind deals with enduring, dispositional influences on its mental representations just as it deals with temporary, circumstantial influences—namely, post hoc. We argue that, efficiently or not, when people are aware of their enduring tendencies to see the world in certain ways, they nonetheless take these tendencies into account only after their effects have been wrought. Thus, the momentary realism that characterizes temporarily tilted heads characterizes permanently tilted minds as well. In the studies that we report here, we investigated how people who are dispositionally prone to see the world in overly positive or negative ways deal with the influence that these dispositions exert on their perceptions of reality. In Study 1, we sought to demonstrate that such people initially assume that a stimulus is isomorphic with their “dispositionally tinted” experience of it, and only then correct for the possibility that their subjective experience was “colored” by their dispositions.

STUDY 1

In this study, dispositionally happy and unhappy participants saw brief flashes of color on a computer screen, were falsely told that they had been exposed to 12 words at subliminal speed, and were asked to guess the words while under time pressure (hurried condition) or no time pressure (unhurried condition). We expected that all participants would use their subjective experience to identify the words, but that unhurried participants would then correct for the influence of their own dispositions.

Method

Identifying participants’ dispositions

Twenty-seven female and 21 male undergraduates at the University of Texas at Austin completed the Positive Affect/Negative Affect Schedule, or PANAS (Watson, Clark, & Tellegen, 1988). The instructions emphasized that participants were to make dispositional ratings (“Tell us how you see yourself on average”) rather than ratings of their current mood. Watson et al. (1988) have shown that the PANAS may be used in this way to measure dispositional affective propensities. Pilot testing revealed relatively little variance in this population on the positive affect subscale, and thus only the negative affect subscale was used as a classification tool. Participants who scored above the mean of 18 on the negative affect subscale were classified as dispositionally unhappy (11 males and 10 females), and the remaining participants were classified as dispositionally happy (10 males and 17 females). Technically, of course, these participants differed only in their propensity to experience negative affect, but for narrative convenience, we refer to them as unhappy and happy participants.

Ensuring participants’ awareness of their dispositions

The experimenter claimed to have entered the subjects’ responses into a computer that could provide feedback about their personalities. Happy participants were given feedback stating that they typically displayed “very high patterns of positive affect and positive emotion” and could be classified as “a very happy individual.” Unhappy participants were told that they typically displayed “very high patterns of...
negative affect and negative emotion” and could be classified as “an occasionally unhappy individual prone to mild depression.” This feedback was always congruent with the participant’s self-report and was intended to ensure that participants were aware of their own dispositions.

Measuring participants’ inferences

Participants were shown a 1.5-min videotape in which alphabetic characters appeared for 100 to 2,000 ms at different locations on a video monitor. Participants were instructed to search for vowels. On 12 occasions, an unexplained burst of color appeared on the screen for about 30 ms. When this task was completed, participants were falsely told that a word had been presented at subliminal speed during each of the bursts of color. The experimenter claimed that “when certain words are flashed for a very brief time . . . they can have a small but real effect on a person’s moods—even though the person cannot consciously recall having seen the words.” Next, participants were told that on each of 12 trials they would be shown three words, and they should choose the one that they believed had been presented subliminally. Participants were shown 12 cards, each containing a positive word (“love”), a negative word (“pain”), and a neutral word (“fact”). Each word was numbered, and participants were instructed to call out the number of the word on each card that they thought they had seen earlier. The order in which the words were printed on the cards was counterbalanced across cards, and the words on each card were matched for length and frequency.

Impairing participants’ ability to correct their inferences

Participants in the hurried condition were told that they had to announce their choice within 2 s, whereas participants in the unhurried condition were told that they needed to wait at least 10 s before announcing their choice. Research suggests that time-pressure manipulations, like divided-attention manipulations, impair people’s ability to correct their judgments (e.g., Gilbert, Krull, & Malone, 1990; Gilbert, Tafarodi, & Malone, 1993; Kruglanski & Freund, 1983; Swann, Hixon, Stein-Seroussi, & Gilbert, 1990).

Results and Discussion

Four female and 4 male participants failed to follow instructions and their data were not analyzed, thus leaving 10 happy and 10 unhappy participants in each condition. We assumed that happy participants were naturally experiencing more positive affect than unhappy participants, that all participants would use that subjective experience to infer the objective properties of the words, but that unhurried participants would then consider their dispositions and correct their judgments. Thus, we expected hurried participants to be more likely than unhurried participants to choose words that were congruent with their dispositionally induced subjective experience. A word-choice index was created by awarding 1 point for each positive word chosen and subtracting 1 point for each negative word chosen. The index had a range of −12 to 12. As recommended by Rosenthal and Rosnow (1985; Rosnow & Rosenthal, 1995), the data were analyzed using planned contrasts, which revealed that when participants were hurried, happy participants were more likely to report seeing positive words than were unhappy participants (Ms = 6.1 and −1.2, respectively), t(1, 36) = 4.8, p < .001, but that when participants were unhurried, happy and unhappy participants were equally likely to report seeing positive words (Ms = 1.8 and 1.1, respectively), t < 1. Analysis of variance (ANOVA) also revealed the expected Condition × Disposition interaction, F(1, 36) = 9.45, p < .005.

The correction account of these data suggests that (a) all participants initially used their subjective experience to estimate the valence of the words, (b) all participants recognized that their dispositions might have influenced their subjective experience, (c) only unhurried participants were able to use this information to correct their estimates, and (d) as a result, unhurried participants purposefully selected a word whose valence was incongruent with their subjective experience. The data show that unhurried participants selected words that were incongruent with their subjective experience, but how do we know that they did so purposefully (rather than randomly) and that they did so because they recognized the role that their dispositions played in creating their subjective experience (rather than for some other reason)?

These questions were examined in Study 2. First, we examined the claim that participants purposefully chose incongruent words by installing a “purposeful choice detector” in the stimulus materials. We reasoned that if, for example, happy participants were purposefully avoiding congruent words (“love”) and purposefully choosing incongruent words (“fact” or “pain”), then the likelihood that they would choose an incongruent word would not be influenced by the presence of an additional congruent word (“hope”) in the set. However, if participants were choosing incongruent words randomly rather than purposefully, then the presence of an additional congruent word would decrease the probability that they would choose an incongruent word. Second, we examined the claim that unhurried participants chose incongruent words because they thought that their dispositions might be influencing them to do the opposite. This claim suggests that participants should choose an incongruent word only when they are acutely aware of their dispositions. In Study 2, we manipulated participants’ levels of awareness by exposing only some of them to the bolstering manipulation used in Study 1.

STUDY 2

In Study 2, only some participants were shown an additional congruent word, and only some were given bolstering feedback. None were hurried. First, we predicted that participants who received bolstering feedback would (like unhurried participants in Study 1) consider their dispositions and purposefully choose incongruent words, and that the presence of an additional congruent word would not influence the likelihood that they would do so. Second, we predicted that participants who did not receive bolstering feedback would not consider their dispositions and would not purposefully choose an incongruent word, and that the presence of an additional congruent word would therefore decrease the likelihood that they would choose an incongruent word.

Method

Dispositionally happy participants in the bolstered condition received feedback about their dispositions, and participants in the unbolstered condition did not. Participants then engaged in the bogus visual tracking task used in Study 1. Afterward, participants in the
three-word condition were presented with a negative, neutral, and positive word on each of six trials and were asked to determine which of these they had seen. Participants in the four-word condition were presented with a negative word, a neutral word, and two positive words, and were asked to make the same determination.

Identifying participants’ dispositions

Several hundred undergraduates at the University of Texas at Austin completed the PANAS (as modified to measure dispositions for Study 1) early in the semester, and only those who were classified as dispositionally happy (per the criteria used in Study 1) were invited to participate in the experiment. Participants were preselected in this way so that the PANAS would not need to be administered at the beginning of the experiment, which might have made unbolstered participants aware of their dispositions. Twenty-three males and 25 females participated in the study.

Bolstering participants’ awareness of their dispositions

Participants in the bolstered condition completed the PANAS at the beginning of the experimental session and were given the feedback given to happy participants in Study 1. Participants in the unbolstered condition did not complete the PANAS at the beginning of the experimental session and were given no feedback.

Manipulating the number of congruent alternatives

All participants performed the visual tracking task used in Study 1, all were told that they had been exposed to six words at subliminal speed, and all were asked to determine which of the words printed on each of six cards had been presented subliminally. Participants in the three-word condition were shown cards that contained one positive, one negative, and one neutral word. Participants in the four-word condition were shown cards that contained the same words used in the three-word condition plus one additional positive word. Other than the number of trials and words, the procedure was identical to the procedure used in the unhurried condition of Study 1. Finally, at the end of the experimental session, participants in the unbolstered condition completed the PANAS.

Results and Discussion

The data from 4 males and 4 females were omitted for the following reasons. One participant was suspicious of the procedures. Seven participants provided new PANAS scores that (a) diverged significantly from the PANAS scores collected earlier in the semester and (b) would have made them ineligible to participate in the experiment. Of these 7 participants, 2 were in each of three conditions and 1 was from the remaining condition. With the data from these 8 participants omitted, 10 participants remained in each condition.

We expected the presence of an additional congruent word to decrease the number of incongruent words chosen by unbolstered participants, but not by bolstered participants. Planned contrasts performed on the word-choice index (computed as in Study 1, with a range of −6 to 6) revealed that the presence of an additional congruent word affected only the unbolstered participants. Unbolstered participants chose more positive words in the four-word condition than in the three-word condition ($M_s = 3.5$ and 1.3, respectively), $t(1, 36) = 2.06, p < .05$, but bolstered participants chose an equal number of positive words in the two conditions ($M_s = 1.3$ and 1.1, respectively), $t < 1$. ANOVA revealed only a weak Number of Words × Bolstering interaction, $F(1, 32) = 2.76, p = .12$. These findings support our claim that dispositionally happy participants who were aware of their dispositions purposefully chose words that were not positive, and that they did so because they had considered the impact of their dispositions on their subjective experience.

GENERAL DISCUSSION

People seem to equate appearances and realities, and question this equivalence only subsequently and under special circumstances. The present studies demonstrate that even when reality’s appearance is distorted by enduring personal dispositions about which observers are well aware, observers may correct for that distortion only after it has been introduced. Unless and until such corrections are made, observers seem to take their subjective experience as a faithful reflection of objective reality. It is not surprising, perhaps, that people naturally trust in the faithfulness of representations that are distorted only on rare occasions—such as visual representations of the tilt and size of an object. It is somewhat more surprising to find that people initially trust in the faithfulness of representations that are distorted by an enduring, constant, and salient influence of which they are fully aware. These findings fit nicely with a correction account of judgmental realism, which suggests that realism is neither absent nor dormant, but is continually active and occasionally outmaneuvered.

The present studies have implications for many of the theoretical issues discussed earlier, as well as more practical consequences. Effective communication, harmonious interpersonal relations, and behavioral prediction all require that people understand how others see the world, and to do that they must consider the influence that their dispositions exert on their perceptions (“I love this salsa, but I have an asbestos mouth so others may find it too spicy”). Several investigators (Dunning, Griffin, Milojkovic, & Ross, 1990; Griffin, Dunning, & Ross, 1990; Griffin & Ross, 1991; Robinson, Keltner, Ward, & Ross, 1995) have noted that realism is often at the root of social conflict because people who believe they have seen the world “as it really is” are often suspicious of those whose backgrounds or agendas cause them to see it differently. If all inferences are momentarily realistic, then we would expect such intransigence to be the rule in social life, and we would expect the appreciation of differing perspectives to be the exception. Indeed, participants in our studies corrected their inferences only when knowledge of their dispositions was reinforced by the kind of explicit feedback that is relatively rare in everyday situations. Furthermore, as Wilson and Brekke (1994) have argued, correction is only as good as the naive theories on which it depends, and to the extent that people have incorrect beliefs about their own dispositions (“I’m a very open-minded person”), their attempts at correction may make matters worse rather than better (“So if I can’t stand hip-hop, it must be totally worthless”). In short, if realism is a necessary prelude to idealism, then the opportunities for social misunderstanding seem particularly rich.

These studies are the first to show that people correct—and sometimes fail to correct—for the influence of their own dispositions on their judgments of stimulus properties. The findings do not “prove” the correction account, of course, as any phenomenon can be explained by more than one model. But because so much evidence
consistent with this account has now been amassed, and because no inconsistent evidence has arisen, it seems reasonable to embrace it as an explanation of judgmental realism. The burden, we believe, is now on alternative accounts to explain all the data as efficiently.

Acknowledgments—We thank Boaz Keysar and several anonymous reviewers for comments on this article. We gratefully acknowledge the support of Grant No. RO1-MH56075 from the National Institute of Mental Health to Daniel T. Gilbert and Timothy D. Wilson.

REFERENCES


(Received 8/27/99; Revision accepted 3/21/00)