

When Does Introspection Bear Fruit? Self-Reflection, Self-Insight, and Interpersonal Choices

J. Gregory Hixon and William B. Swann, Jr.

Whereas earlier research suggests that the fruits of introspection may promote error and misperception, this research suggests that thinking about the self may sometimes foster self-insight. Participants who had opportunity to reflect on themselves were particularly inclined to display self-insight by (a) rating feedback that confirmed their self-views as self-descriptive (Experiments 1 and 3), (b) rating themselves in ways that matched their friends' appraisals of them (Experiment 2), and (c) choosing a self-verifying interaction partner rather than an overly favorable one (Experiment 4). These effects were moderated by the nature of the introspective activity (Experiment 3) and by its duration (Experiment 4). Implications of these findings for the nature of self-knowledge and the worlds people construct around themselves are discussed.

Self-contemplation is a curse—That makes an old confusion worse.

—Roethke, *The Collected Poems of Theodore Roethke*

When people peer inward, what do they see? Intellectual fog, Roethke might reply, and many would agree with him. In fact, many members of our society harbor a deep distrust of the power of introspection. Witness, for example, the tendency of elementary school teachers to advise their pupils to answer achievement test questions with “the first answer that pops into your head.” Similarly, coaches exhort athletes to “play through” their slumps rather than systematically analyze the source of their difficulties. The message underlying such injunctions is clear: Reflection may undermine, rather than promote, self-insight.

The research literature has fueled such skepticism regarding the consequences of reflection. Schooler and his colleagues (e.g., Schooler & Engstler-Schooler, 1990; Schooler, Foster, & Loftus, 1988), for example, have shown that simply striving to remember something may cause people to misremember it the next time around. Specifically, these researchers exposed people to a series of images and then had them complete a retrieval test. They found that the act of completing the retrieval test contaminated people's memories and diminished performance on a later memory test.

J. Gregory Hixon, Department of Psychology, University of Connecticut; William B. Swann, Jr., Department of Psychology, University of Texas at Austin.

This research and the preparation of this article were supported by research funds (MH 37598) and a research scientist development award (MH 00498) from the National Institute of Mental Health to William B. Swann, Jr., and a National Science Foundation predoctoral fellowship to J. Gregory Hixon. We thank Stephanie Yarmo for conducting Experiment 3. We are grateful to Dan Gilbert, Tim Wilson, and several anonymous reviewers for their helpful comments on earlier versions of this article.

Correspondence concerning this article should be addressed to J. Gregory Hixon, Department of Psychology, Box U-20, Room 107, University of Connecticut, Storrs, Connecticut 06269-1020.

In a similar vein, studies of nonconscious information processing and implicit learning have suggested that reflection sometimes undermines performance. Researchers have found, for example, that encouraging participants to reflect on and implement information-processing strategies that are ordinarily nonconscious interferes with learning (e.g., Brooks, 1978; Howard & Ballas, 1980; Reber, 1976; Reber, Kassir, Lewis, & Cantor, 1980). Reber (1989) speculated that such performance deficits grow out of a tendency for reflective participants to grope for rules that are simply inaccessible: “Looking for rules will not work if you cannot find them” (p. 223).

Just as the foregoing evidence indicates that contemplation interferes with certain cognitive processes, other recent research has shown that self-reflection actually undermines self-insight. To wit, Nisbett and Wilson (1977) showed that when people were asked to explain why they behaved as they did, they relied on shared theories about the causes of their behaviors rather than on the actual causes of their behaviors. To be sure, Nisbett and Wilson's work inspired a host of thoughtful critiques (e.g., Ericsson & Simon, 1980; Kraut & Lewis, 1982; Smith & Miller, 1978). The critics, however, were able to qualify but not dismiss the claim that the fruits of introspection were misleading. Furthermore, subsequent investigators extended and enriched Nisbett and Wilson's argument by demonstrating that thinking can undermine the relation between people's attitudes and behaviors (e.g., Millar & Tesser, 1986, 1989; Wilson & Dunn, 1986; Wilson, Dunn, Bybee, Hyman, & Rotondo, 1984; Wilson, Dunn, Kraft, & Lisle, 1989) and can reduce the quality of their decisions (e.g., Wilson & Schooler, 1991).

Both casual observation and controlled experimentation therefore prompt people to treat the products of reflection with a measure of incredulity. Yet, even the harshest critics of the power of introspection add a caveat to their analyses. For example, noting several instances in which thinking actually facilitated learning performance, Reber (1989) conceded that “Looking for rules will work if you can find them” (p. 223). Similarly, despite a general pessimism regarding the products of introspection, both Bem (1972) and Nisbett and Wilson (1977) allowed that only weak and ambiguous internal cues offer unreli-

able insights into the self. In support of this, Wilson and his colleagues have shown that thinking undermines performance only when information about the attitude object is either unavailable (Wilson, Kraft, & Dunn, 1989; Wilson et al., 1984) or inaccessible (Wilson, Hodges, & Pollack, 1991).

One implication of the foregoing research is that when the object of introspection is the self, reflection should prove quite fruitful. Why? First, self-relevant information is highly available—people have more information about themselves than they have about other people or things (Kihlstrom et al., 1988). Second, self-relevant information is generally more accessible than other types of information (e.g., Klein & Loftus, 1988; Rogers, Kuiper, & Kirker, 1977). This analysis suggests that self-reflection will ordinarily foster self-insight and that factors that interfere with introspection and hinder people's ability to access their self-concepts will undermine self-insight.

On what basis do people respond when they cannot access their chronic self-conceptions? Research by Swann, Hixon, Stein-Seroussi, and Gilbert (1990) offers a lead. These researchers proposed that people must access their self-conceptions if they are to choose interaction partners who appraise them in a manner that confirms their self-views. In support of this hypothesis, they found that only those participants who had cognitive resources available chose interaction partners who saw them as they saw themselves; those who were deprived of cognitive resources chose favorable evaluators, even if such evaluators appraised them more favorably than they appraised themselves. Apparently, people who are unable to access their self-concepts respond merely on the basis of the positivity of stimuli and follow a simple rule of the form "If the feedback is positive, approach it; if the feedback is negative, avoid it."

In short, when it comes to self-views, the opportunity to engage in self-contemplation should promote rather than impair self-insight.¹ Thus, for example, people who have ample cognitive resources available should be particularly facile in assessing the self-descriptiveness of social feedback. In contrast, people who have few cognitive resources available should simply embrace favorable descriptions of themselves. We tested these predictions in the initial investigation.

Experiment 1

To determine whether the ability to engage in self-reflection influences self-assessments, we presented participants who had been pretested on a measure of self-perceived sociability with favorable and unfavorable feedback. Participants in the low reflective condition were deprived of mental resources by having them judge the accuracy of the feedback more rapidly than participants in the high reflective condition. We expected that participants in the low reflective condition would tend to embrace the favorable feedback as more self-descriptive, regardless of their self-views. In contrast, we expected that participants in the high reflective condition would tend to use their self-conceptions to gauge the self-descriptiveness of the feedback, with the result that participants with negative self-conceptions would tend to rate unfavorable feedback as self-descriptive, and participants with positive self-concepts would tend to rate favorable feedback as self-descriptive.

Method

Participants

Forty-six women enrolled in introductory psychology at the University of Texas at Austin took part in this experiment in partial fulfillment of course requirements. To ensure clear differences in our subject groups, we recruited only those who scored in the highest 20% (positive self-concept) or lowest 20% (negative self-concept) on the Texas Social Behavior Inventory (TSBI; Helmreich, Spence, & Stapp, 1974), a measure of social self-esteem.

Procedure

A male experimenter who was blind to the participant's self-concept led each participant to a private cubicle. He told each subject that he was interested in the ability of clinical trainees to discern a person's true characteristics on the basis of personality profiles. He then explained that two graduate students in clinical psychology had reviewed the participant's responses to some personality tests that she had completed at the beginning of the semester. The responses, the experimenter continued, had offered each graduate student a basis for completing an evaluation of each participant. He explained that each participant would be asked to review these two evaluations and then estimate the accuracy of each evaluation.

The experimenter then told the participant that, because of time constraints, a limited amount of time would be available for reviewing the evaluations. To ensure that participants did not spend their limited time deciphering the format of the evaluations, the experimenter familiarized each participant with a blank evaluation form. Participants in the low reflective condition learned that they would have 10 s to review the evaluations; participants in the high reflective condition learned that they would have 45 s. The experimenter then presented the participant with evaluations that had ostensibly been prepared by two independent raters. Each evaluation sheet contained ratings of the participant's sociability, likability, and interestingness on scales that ranged from 0 to 10. One set of ratings was relatively favorable ($M = 8.33$), and the other was relatively unfavorable ($M = 5.0$). We selected these levels of favorability because they approximated the self-views of participants who scored high and low on the TSBI.

The experimenter placed the evaluations on a table in front of the participant. For some participants, the favorable evaluation was placed on the left and the unfavorable on the right; for other participants this order was reversed (spatial position had no effect on the results). After either 10 s or 45 s, the experimenter asked the participant to assess the accuracy of the evaluations on scales that ranged from -5 (not at all accurate) to $+5$ (extremely accurate). When the participant was finished, the experimenter thanked and debriefed her.

Results and Discussion

We expected that all low reflective participants, and high reflective participants with positive self-views, would endorse the accuracy of the favorable evaluation. In contrast, we expected that participants with negative self-views in the high reflective condition would show relatively more endorsement of the unfavorable evaluation. As Table 1 shows, the results confirmed our predictions. Examination of the difference scores,

¹ We use phrases such as "promotes self-insight" or "fosters self-knowledge" in the sense of obtaining a clearer picture of preexisting information about the self and not in the sense of learning something new about the self.

Table 1
Impact of Self-Reflection and Self-Conception on Rated Accuracy of Favorable and Unfavorable Evaluations (Experiment 1)

Self-concept	Self-reflection	
	Low	High
Negative		
Difference (favorable-unfavorable)	+3.09	+1.08
Favorable	+2.18	+2.25
Unfavorable	-0.91	+1.17
Positive		
Difference (favorable-unfavorable)	+3.33	+4.00
Favorable	+3.50	+3.36
Unfavorable	+0.17	-0.64

Note. $n = 10$. Higher numbers indicate higher rated accuracy. Range = -5 to +5.

for example, revealed a reliable interaction between reflection and self-concept, $F(1, 42) = 5.177, p < .05$. Simple effects analyses revealed that participants with negative self-concepts in the high reflective condition were more inclined than participants in each of the other three cells to embrace the unfavorable feedback (relative to the favorable feedback) as accurate (all $F_s > 6.58$, all $p_s < .02$). Moreover, within the other three cells, none were reliably different from the others (all $F_s < 1$).

The foregoing results suggest that participants with negative self-concepts in the high reflective condition were less apt to distinguish the two evaluations than participants in any of the other conditions. Subsequent analyses revealed that participants with negative self-views in the high reflective condition failed to differentiate the two types of feedback, $t(11) = 1.78, p > .10$,² whereas participants in the other three conditions assigned greater accuracy ratings to the favorable as compared with the unfavorable feedback (all $t_s > 5.4$, all $p_s < .001$).

In summary, our findings supported our predictions in that participants who had opportunity to reflect on themselves were more likely to endorse the accuracy of self-descriptive feedback relative to self-discrepant feedback. Participants who had available a paucity of mental resources simply embraced favorable feedback regardless of their self-conceptions. The systematic preference exhibited by participants in the low reflective condition indicates that they did indeed discriminate the favorable feedback from the unfavorable feedback but simply failed to access their self-concepts.

Experiment 2

The second experiment incorporated three methodological refinements designed to illuminate the strength and generalizability of the phenomenon demonstrated in Experiment 1. First, we adopted a very different criterion to gauge the validity of self-insight. Specifically, rather than having participants evaluate the self-descriptiveness of feedback, we gauged the validity of their self-assessments by comparing their self-ratings with their friends' ratings of them. Second, we focused on participants' conceptions of four self-attributes in addition to their sociability. Third, rather than depriving participants of the op-

portunity to self-reflect by inducing them to rate themselves quickly, we had them rate themselves while performing a simultaneous cognitive task.

In summary, then, participants in Experiment 2 either were or were not deprived of the opportunity to reflect while they rated themselves on five self-attributes. We then compared participants' self-ratings with their friends' ratings of them. We expected that (a) the amount of covariation between participants' self-ratings and their friend's ratings of them would be higher when participants had ample opportunity to reflect than when they did not and (b) as in Experiment 1, participants who had relatively little opportunity to reflect would rate themselves more positively than those who had ample opportunity to reflect.

Method

Participants

Forty-five women who were enrolled in introductory psychology at the University of Texas at Austin participated in this experiment for either course credit or payment of \$3. Each participant brought along a female friend who received \$3 for her participation.

Aware of the tendency for most people to evaluate themselves favorably (e.g., Swann, 1987; Taylor & Brown, 1988) and anxious to ensure that at least some participants would have negative self-views (to avoid difficulties with restriction of range), we recruited only participants who scored in the lowest 30% on the short form of the Self Attributes Questionnaire (SAQ; Pelham & Swann, 1989) during a pretest. The SAQ is a measure of five self-views central to self-worth: intellectual capability, physical attractiveness, athletic ability, social skills, and aptitude for arts and music. For each attribute, participants rate themselves relative to other college students their own age on graduated interval scales that range from 1 (bottom 5%) to 10 (top 5%). The SAQ has been shown to be stable over a period of 4 months, test-retest $r(50) = .77$.

Procedure

On their arrival, a male experimenter led the participant and her friend to separate cubicles. The experimenter introduced the study as an investigation of computer diagnosis of personality attributes. He explained that the participant's task would be to answer some questions on a computer. He told her that each question (e.g., "How intelligent are you?") would appear on the screen above a scale that ranged from 0 (*not very much so*) to 9 (*very much so*). The participant was told to answer each question by pressing the appropriate key and that her response would prompt the computer to proceed to the next question. The participant was given 20 s to answer each question.

When the participant indicated that she understood the procedure, the experimenter introduced the self-reflection manipulation. The experimenter explained that regulations regarding confidentiality of experimental data required that each participant have a special access code. Before starting the computer program, the experimenter provided the participant with an 8-digit access number. In the high reflective condition, the experimenter provided this number on a piece of paper. In the low reflective condition, he gave the access number ver-

² As we note in the General Discussion, the absolute values of these mean difference scores depend on many factors that are irrelevant to our concerns here; we accordingly urge readers to focus on the pattern of means between conditions.

bally and asked the participant to remember it until the computer prompted her to enter it (earlier the experimenter had taken steps to ensure that the participant had no writing implements available to her). He then started the computer program and departed. The computer asked participants to rate themselves on the five SAQ attributes as well as two filler dimensions (tan and wealthy) and then requested them to enter the access number.

While the participant was completing the self-rating task, the experimenter approached the participant's friend and asked her to answer a brief set of questions about her friend (i.e., the participant). The questions were identical to those that the participant answered about herself, except that we reworded the questions so that they assessed the friend's view of the participant (e.g., "How intelligent is your friend?"), and we excluded the filler items. The experimenter stressed that the friend's responses would not be disclosed to the participant. When the participant and the participant's friend were finished, each was separately debriefed, given course credit or paid, and thanked for her participation.

Results and Discussion

We expected that the average correlation between each participant's self-ratings and her friend's ratings of her across the five SAQ attributes would be lower for participants in the low reflective as compared with the high reflective condition. To test this prediction, we computed a correlation coefficient for each participant-friend pair, performed Fisher's r -to- z transformation, and then converted the z scores back into transformed correlations to be used in analyses. The results supported our prediction. That is, the mean transformed correlation was larger in the high reflective condition ($r = .71$) as compared with the low reflective condition ($r = .43$), $F(1, 43) = 6.43$, $p < .05$.³

One potential ambiguity associated with the foregoing correlations is that participants in the high reflective condition may have achieved higher correlations by making ratings in line with a common or prototypic notion of the typical person's SAQ profile. One could argue that participants in the low reflective condition achieved lower correlations because they failed to access or use the prototype in making their self-ratings (rather than because their self-ratings were relatively "accurate" or "inaccurate"; see discussions of *stereotypic accuracy* in Cronbach, 1955; Funder, 1980). To explore this possibility, we constructed a prototypic SAQ profile by averaging the ratings of 1,468 participants who completed the SAQ during the pretest. The results did not support the rival hypothesis. First, our participants did not endorse a widely shared prototype; in fact, the ratings of only 1 of our 45 participants matched the prototype constructed from the pretest scores. In addition, we found that even if we relaxed the criterion for agreement, the self-ratings of participants in the high reflective condition were not more strongly correlated with the prototype than the ratings of low reflective participants ($F < 1$, ns). Apparently, the tendency for participants in the high reflective condition to show higher participant-friend correlations than participants in the low reflective condition reflected greater access to idiosyncratic self-knowledge rather than greater access to a common prototype.

We also anticipated that the positivity of participants' self-ratings (relative to their friends' ratings of them) would be greater for participants in the low as compared with the high reflective condition. To examine this possibility, we computed

the discrepancy between friends' ratings of participants and participants' self-ratings such that higher numbers reflect greater participant positivity. The expected pattern occurred for both the social skills-competence variable, $F(1, 43) = 4.87$, $p < .05$ (mean difference scores: high reflective participants = -0.83 , low reflective participants = $+0.32$), and the artistic-creative variable, $F(1, 43) = 4.64$, $p < .05$, (mean difference scores: high reflective participants = -1.00 , low reflective participants = $+0.68$). The opportunity to reflect did not cause participants to inflate their self-ratings relative to their friends' ratings of them on the dimensions of intelligence, athletic ability, and attractiveness (all F s < 1 , ns).

Curtailling participants' opportunity to reflect caused them to inflate their ratings on the sociability dimension. This replicates the results of Experiment 1, which also focused on self-perceived sociability. It is less clear why reflection influenced sociability and artistic self-ratings but had no impact on intelligence, athletic ability, and physical attractiveness ratings. We suspect that our participants somehow developed an implicit sense that they could inflate their self-perceived social and artistic ability because there are relatively few objective markers that would provide a grounds for refutation of such self-views. In contrast, the relatively clear objective criteria associated with intelligence, athletic ability, and physical attractiveness (e.g., Scholastic Achievement Test scores, trophies, and appearance, respectively) mean that it is hazardous to entertain overly favorable self-estimates on these dimensions (e.g., Dunning, Meyerowitz, & Holzberg, 1989). Further research, of course, will be needed to test this proposition directly.

In summary, the results of Experiment 2 indicate that self-reflection promotes knowledge of the relations among multiple self-concepts. Taken together with the results of Experiment 1, these findings reinforce the notion that reflecting on the self can indeed promote knowledge of the self.

Experiment 3

Why did we find that reflection fosters self-insight, whereas others (e.g., Millar and Tesser and Wilson and his colleagues) found that reflection undermined self-insight? The answer to this question may lie in the type of reflection being investigated. For example, type of reflective focus seems to influence insight into one's affect-based attitudes. Triggering reflection by asking *why* (e.g., "Why do you feel this way?"; "List the reasons why you hold your attitude about this object.") undermines insight such that reported attitudes are virtually unrelated to behavior, but triggering reflection by focusing on *what* (e.g., "Focus on what you feel about this object.") tends, if anything, to foster insight into one's attitudes (Millar & Tesser, 1986; Wilson & Dunn, 1986). Presumably, focusing on what makes the actual components of attitudes more accessible and thus promotes insight, whereas focusing on why calls to mind shared

³ We also expected that the reflection manipulation might diminish the correlation between participants' self-ratings during the experiment and their self-ratings during the pretest. The high test-retest reliability of SAQ scores, however, meant that this correlation was only slightly lower in the low reflective condition ($r = .69$) as compared with the high reflective condition ($r = .74$).

theories, plausible accounts, and rationalizations that may bear little resemblance to one's actual attitudes (Nisbett & Wilson, 1977), thus undermining insight.

Experiments 1 and 2 have focused on the consequences of asking what with regard to the self (e.g., in Experiment 1: "What do you think about your sociability vis-à-vis these evaluations?"; in Experiment 2: "What is your level of intelligence?"). The results have suggested that reflection promotes self-insight. However, just as reflecting on why undermines insight into one's attitudes, it may also undermine insight into one's self. Experiment 3 explored the possibility that these two questions trigger self-reflection processes that have unique properties and different consequences.

As in Experiment 1, participants reflected on a favorable and an unfavorable evaluation for either 10 s (low reflective) or 45 s (high reflective) before estimating the accuracy of each evaluation. In addition, we manipulated the nature of the reflective focus by instructing some participants to focus on "what you are" during their reflection time and by instructing others to focus on "why you are the way you are" during their reflection time. All participants in this experiment had negative self-concepts.

We anticipated that the *what* focus would facilitate self-insight when participants had sufficient time to reflect. Accordingly, among participants in the *what* focus condition, we expected that low reflective participants would show little self-insight and would judge the favorable evaluation as more accurate than the unfavorable evaluation but that high reflective participants would access their negative self-concepts and would therefore show a greater tendency to judge the unfavorable evaluation as more accurate than the favorable one. On the other hand, we anticipated that the *why* focus would disrupt self-insight even among those with ample time to reflect. Accordingly, we expected that participants in the *why* focus conditions would judge the favorable evaluation as more accurate than the unfavorable evaluation regardless of whether they had ample time to reflect.

Method

Participants

Forty-one women enrolled in introductory psychology at the University of Texas at Austin participated in partial fulfillment of a course requirement. All participants had scored in the lowest 20% on the TSBI in a pretest administered at the beginning of the semester. We omitted the data from 2 participants who expressed skepticism regarding the source and nature of the evaluations, leaving 39 participants in the final analyses.

Procedure

The procedures for Experiment 3 were similar in most respects to those of Experiment 1. A female experimenter led each participant to a private cubicle. The experimenter told the participant that she was interested in the ability of clinical trainees to discern a person's true characteristics on the basis of personality profiles. The experimenter further indicated that two graduate students in clinical psychology had prepared evaluations of the participant on the basis of the participant's responses on pretest questionnaires administered earlier in the semes-

ter. The participant learned that she would be asked to estimate the accuracy of each evaluation.

As in Experiment 1, the experimenter then indicated that time constraints would limit the participant's time for reviewing the evaluations. Consequently, the experimenter familiarized each participant with a blank evaluation form. Participants in the low reflective condition learned that they would have 10 s to review the evaluations; participants in the high reflective condition learned that they would have 45 s. The experimenter then introduced the focus manipulation by claiming that giving participants a focus question would help them in assessing the accuracy of the evaluations. Participants assigned to the *what* focus condition learned that they should think about the question "What kind of person are you in terms of sociability, likability, and interestingness?"; participants in the *why* focus condition learned that they should think about the question "Why are you the kind of person you are in terms of sociability, likability, and interestingness?" The experimenter then presented the participant with evaluations that had ostensibly been prepared by two independent raters. Each evaluation sheet contained ratings of the participant's sociability, likability, and interestingness on scales that ranged from 0 to 10. As in Experiment 1, one set of ratings was relatively favorable ($M = 8.33$), and the other was relatively unfavorable ($M = 5.0$).

The experimenter placed the evaluations on a table in front of the participant. For some participants, the favorable evaluation was placed on the left and the unfavorable on the right; for other participants the order was reversed (spatial position had no effect on the results). After either 10 s or 45 s, the experimenter asked the participant to assess the accuracy of the evaluations on scales that ranged from -5 (*not at all accurate*) to $+5$ (*extremely accurate*). When the participant was finished, the experimenter thanked and debriefed her.

Results and Discussion

We expected that all participants in the *why* focus condition, and low reflective participants in the *what* focus condition, would judge the favorable evaluation as more accurate than the unfavorable evaluation. In contrast, we expected that high reflective participants in the *what* focus condition would be more inclined to judge the unfavorable evaluation as more accurate than the favorable one.

As in Experiment 1, we focused our attention on the difference scores, and Table 2 shows results consistent with our predictions. An analysis of variance revealed the predicted interaction between self-reflection and reflective focus, $F(1, 35) = 6.653$, $p < .05$. Simple effects analyses showed that this interaction was of the predicted form. Among participants in the *why* focus condition, self-reflection had no impact on participants' tendency to judge the favorable evaluation as more accurate than the unfavorable one, $F(1, 17) = 1.92$, $p > .19$. In contrast, among participants in the *what* focus condition, self-reflection had a substantial impact on accuracy ratings assigned to the favorable as compared with the unfavorable evaluation, $F(1, 18) = 5.17$, $p < .05$.

An examination of each group's accuracy judgments of the favorable and unfavorable evaluations sheds further light on these results. Participants who focused on *why* judged the favorable evaluation as more accurate than the unfavorable one, even though this tendency was more reliable for high reflective participants, $t(8) = 3.49$, $p < .01$, than for low reflective participants, $t(9) = 1.50$, $p = .17$. Among participants who focused on *what*, however, those in the low reflective condition judged the

Table 2
Impact of Self-Reflection and Reflective Focus on Rated Accuracy of Favorable and Unfavorable Evaluations (Experiment 3)

Reflective focus	Self-reflection	
	Low	High
What ^a		
Difference (favorable–unfavorable)	+2.00	–1.00
Favorable	+2.73	+1.11
Unfavorable	+0.73	+2.11
Why ^b		
Difference (favorable–unfavorable)	+0.80	+2.67
Favorable	+2.10	+2.33
Unfavorable	+1.30	–0.34

Note. Higher numbers indicate higher rated accuracy. Range = –5 to +5.

^a $n = 11$ for low self-reflective group; $n = 9$ for high self-reflective group. ^b $n = 10$ for low self-reflective group; $n = 9$ for high self-reflective group.

favorable evaluation as more accurate than the unfavorable one, $t(10) = 3.03$, $p = .01$, whereas their high reflective counterparts showed no reliable tendency whatsoever, $t(8) < 1$, *ns*.

These results replicate and extend the findings from Experiments 1 and 2. As in the first two experiments the opportunity to introspect promoted self-insight, but only among those in the *what* focus condition. In conjunction with earlier work showing that people activate relevant self-concepts only when sufficient cognitive resources are available (Swann et al., 1990, Experiment 1), this finding offers converging evidence that the beneficial effects of reflection stem from fostering greater access to self-knowledge (i.e., “what” one is). At the same time, results from the *why* focus condition qualify our earlier findings by indicating that not all types of introspection promote self-insight. That is, focusing on why disrupts insight into one’s self-knowledge. Presumably, participants who focused on why used their reflection time to rationalize, justify, and explain away the negative information that forms the basis of their self-concepts. Indeed, our findings suggest that thinking about why one is the way one is may be no better than not thinking about one’s self at all.

Experiment 4

One important difference between our research and that of earlier researchers is the amount of time people were given to introspect. Whereas earlier researchers gave their participants several minutes to reflect and found that reflection undermined insight, thus far we have limited our participants to less than 1 min. Conceivably, under some circumstances the relation between reflection and self-insight may be curvilinear: Just as a moderate amount of reflection allows people to access relevant self-knowledge, a little more may cause them to access additional information that bears on their subsequent behavioral choices.

Under what circumstances might this be true? Swann (1990) has proposed that when the epistemic and pragmatic consequences of actions are minimal, increments in reflection time

beyond that needed to access self-concepts may actually diminish people’s tendency to behave in accord with their self-concepts. Consider, for example, the relative impact of reflection on people involved in marital relationships versus people involved in dating relationships. Reflection may lead people involved in marital relationships to prefer partners whose appraisals confirm their self-views. After all, the appraisals of spouses are usually highly credible (thus raising the epistemic consequences of appraisal-related choices), and incongruent appraisals can be highly disruptive to the relationship (thus raising the pragmatic consequences). In contrast, reflection may cause people in dating relationships to realize that the epistemic and pragmatic consequences of associating with an inappropriately favorable partner are minimal and that the costs of receiving unfavorable appraisals are high. That is, because dating relationships are “trial” relationships, the very survival of the relationship turns on the positivity of the appraisals of both parties—the subjective accuracy of those appraisals is relatively unimportant. Consistent with this reasoning, Swann, Hixon, and De La Ronde (1992) found that people involved in marital relationships prefer their spouses to see them as they see themselves, whereas people involved in dating relationships, regardless of their own self-views, prefer their partners to view them positively.

The foregoing analysis suggests that when people find themselves in contexts in which the epistemic and pragmatic implications of their choices are minimal, an extended period of reflection will cause them to seek feedback that is positive rather than self-verifying. In Experiment 4 we tested this possibility by varying the amount of time that participants had to decide whether they wanted to engage in a brief interaction with a stranger who had appraised them favorably or with one who had appraised them unfavorably. Specifically, participants had either 10 s (low reflective), 45 s (high reflective), or 3 min (extremely reflective) to reflect before indicating how much they preferred to interact with a favorable or unfavorable evaluator in a later phase of the experiment. We minimized the epistemic and pragmatic consequences of participants’ choices by stressing that the evaluators were their peers (other students) and that the interaction would be a simple 15–20-min chat. All participants had negative self-concepts.

We expected that low reflective and extremely reflective participants would prefer to interact with the favorable evaluator more than the unfavorable one—the former because limited time would prevent access to self-knowledge, the latter because an extended period of time would lead them to realize that the match between their self-concepts and their chosen partner’s appraisal would be irrelevant in this minimally consequential interaction. We expected that high reflective participants would show a greater tendency to prefer the unfavorable evaluator relative to the favorable one because only they would have sufficient time to access their self-knowledge without having time to recognize the limited epistemic and pragmatic consequences of their choice.

Method

Participants

Forty men enrolled in introductory psychology at the University of Texas at Austin participated in partial fulfillment of a course require-

ment. All participants scored in the lowest 20% on the TSBI in a pre-test administered at the beginning of the semester.

Procedure

A male experimenter ushered each participant to a private cubicle. The experimenter explained that the study involved a comparison of two modes of forming impressions: examining personality profiles and meeting face to face. In the initial phase of the study, the experimenter continued, 2 other students had evaluated the participant on the basis of questionnaires he had completed in a mass pretesting session at the beginning of the semester. The experimenter then asked if the participant would be willing to have a 15–20 min getting-acquainted chat with one of the evaluators at a later session. When the participant agreed (all did), the experimenter told him that he would be allowed to see the two evaluations, indicate how much he preferred to interact with each evaluator, and in the later session meet with the evaluator he preferred more.

The experimenter explained that, because of time constraints, the participant would have a limited amount of time, either 10 s (low reflective), 45 s (high reflective), or 3 min (extremely reflective), to review the evaluations before indicating his preferences. The experimenter then presented the two evaluations. On each sheet, one of the two evaluators had rated the participant's sociability, likability, and interest-ness on 10-point scales. As in Experiments 1 and 3, one evaluation was favorable (mean rating = 8.33), and one was relatively unfavorable (mean rating = 5). The spatial position of the favorable and unfavorable evaluations was randomized (spatial position had no effect on the results). After either 10 s, 45 s, or 3 min had elapsed, the participant indicated on 10-point scales how much he wanted to interact with each of the 2 evaluators in the later phase of the experiment. When the participant was finished, the experimenter announced that the experiment was over and then debriefed and thanked the participant.

Results and Discussion

We expected that low reflective and extremely reflective participants would prefer to interact with the favorable evaluator more than the unfavorable one, but that high reflective participants would show a greater tendency to prefer the unfavorable evaluator. As Table 3 shows, the results generally conformed to our predictions.

A planned comparison on the difference scores revealed the predicted curvilinear trend, $F(1, 37) = 6.35, p < .05$. Simple effects tests indicated that high reflective participants were marginally more inclined than low reflective participants to prefer the unfavorable evaluator relative to the favorable one,

$F(1, 24) = 3.13, p < .09$, and were reliably more inclined than extremely reflective participants to do so, $F(1, 25) = 6.15, p < .05$. Further analyses indicated that high reflective participants displayed no difference in their preference to interact with the favorable and unfavorable evaluators, $t(12) < 1, ns$, but that low reflective and extremely reflective participants indicated a greater preference to interact with the favorable evaluator than the unfavorable one: $t(12) = 2.91, p < .05$; $t(13) = 5.68, p < .01$, respectively.⁴ These findings therefore paralleled those of Experiment 1.

The results of Experiment 4 extend our earlier findings by showing that just as some reflection promotes choices that are consistent with underlying dispositions, too much reflection can result in choices that are inconsistent with dispositions. In contrast to low reflective participants, high reflective participants were more likely to prefer an interaction partner whose evaluation of the participant was consistent with the participant's own self-evaluation. However, extremely reflective participants' preferences were very much like those of their low reflective counterparts. When the epistemic and pragmatic consequences of choices are minimal, lots of thinking can yield the same results as hardly thinking at all.

General Discussion

Does self-contemplation promote or undermine self-insight? Whereas recent research has suggested that introspection may do little to promote self-insight, our findings suggest that self-reflection may be an essential ingredient in self-knowledge. Specifically, Experiments 1 and 2 demonstrated that the opportunity to reflect on oneself promotes accurate assessments of the self-descriptiveness of social feedback (Experiment 1) and is associated with greater agreement between people's self-ratings and the appraisals of others (Experiment 2). Experiment 3 clarified the results of Experiments 1 and 2 by showing that reflection promotes self-insight only when one reflects on what one is; reflection on why one is as one is does not promote self-insight. Finally, Experiment 4 offered evidence that the effects of reflection are not necessarily linear—even though a modest amount of reflection promoted choices consistent with one's knowledge of one's self, a longer period of reflection did not.

We have claimed that modest amounts of self-reflection foster self-insight. It may thus seem surprising that, even at their most insightful, participants with negative self-concepts did not consistently endorse unfavorable evaluations or evaluators over favorable ones (i.e., there were positive mean difference scores in the high reflective conditions of Experiments 1 and 4 and a negative mean difference score in the high reflective *what* condition in Experiment 3, none of which were statistically different from zero). Numerous factors, however, affect the absolute nature of participants' responses to evaluations, such as the specific self-concept involved, the precise format and wording of the evaluation, and the response (see Swann, 1990, for a dis-

Table 3
Impact of Self-Reflection on Preference to Interact With Favorable and Unfavorable Evaluators (Experiment 4)

Evaluator	Self-reflection		
	Low ^a	High ^a	Extremely ^b
Contrast weights	+1	-2	+1
Difference (favorable-unfavorable)	1.46	0.15	1.64
Favorable	7.38	6.46	7.86
Unfavorable	5.92	6.31	6.21

Note. Higher numbers indicate greater preference. Range = 0 to 10.
^a $n = 13$. ^b $n = 14$.

⁴ We also ran a companion investigation in which epistemic and pragmatic consequences were moderate. Rather than the curvilinear effect shown here, the companion study revealed a simple (albeit unreliable) linear effect such that greater amounts of reflection led to stronger preferences for the unfavorable evaluator.

cussion). If we had focused on a different self-concept, used different experimental procedures, or used slightly different evaluations we might well have found that high reflective participants with negative self-concepts would in an absolute sense endorse an unfavorable evaluation or evaluator over a favorable one (i.e., have a negative difference score). Of course, even though such alterations might affect difference scores within each condition, we wouldn't expect the relations of scores between conditions to change at all because those alterations would have been the same in all conditions of the experiment. We focus extensively on comparisons between conditions because our claims about the impact of self-reflection on self-insight are theoretically based on such comparisons. We have focused considerably less attention on difference scores within particular experimental conditions because they depend on factors that were not systematically investigated here and that logically should not affect our critical comparisons between conditions.

In all four experiments we found that participants who were deprived of cognitive resources (low reflective participants) embraced particularly favorable appraisals. This finding provides further support for the model of the interplay of positivity and self-verification strivings proposed by Swann et al. (1990). These investigators suggested that when people encounter self-relevant stimuli, a two-stage process unfolds. First, people characterize the stimuli as favorable or unfavorable. This characterization process produces an immediate "affective" reaction or positive tropism (e.g., Zajonc, 1980) that encourages the recipient to embrace favorable feedback and avoid unfavorable feedback. Second, if sufficient cognitive resources are available, people go on to access a self-view relevant to the stimulus and compare that self-view with the feedback. This comparison process prompts people to embrace self-verifying feedback and eschew self-discrepant feedback. Consistent with this model, Swann et al. (1990) showed that resource-deprived people tended to choose self-enhancing interaction partners, and those who were not so deprived tended to choose self-verifying partners. Experiments 1–3 complement this earlier work by showing that depriving people of cognitive resources causes them to endorse relatively favorable descriptions of themselves.⁵ Consistent with the three-stage model recently proposed by Swann (1990), Experiment 4 suggests the presence of an additional processing phase in which one weighs the consequences of one's interpersonal choices.

Our analysis offers a new perspective on recent evidence that increasing arousal (Paulhus & Levitt, 1987) or placing people under cognitive load (Paulhus, Graf, & Van Selst, 1989) causes them to endorse relatively positive adjectives as self-descriptive. Previous researchers have interpreted these findings as evidence that arousal and cognitive load reduce cognitive complexity and thus invite self-deception. We suggest instead that such manipulations do not encourage an active process of self-deception but instead prevent people from engaging in the comparison process that allows them to identify overly favorable self-conceptions as such.

We have claimed that modest amounts of reflection foster self-insight, a claim that on the surface may seem to clash with the suggestion that introspection undermines self-insight (e.g., Millar & Tesser, 1986, 1989; Nisbett & Wilson, 1977; Wilson &

Dunn, 1986; Wilson et al., 1984; Wilson, Dunn, Kraft, & Lisle, 1989). In part, the explanation for this difference may lie in the type of internal cues that we examined. Whereas earlier researchers concentrated on the consequences of reflecting on relatively obscure and transitory internal cues, we focused on relatively salient and chronic cues (i.e., self-views). Research and theorizing in the domain of attitudes suggest that thinking leads to greater polarization of attitudes when they are relatively salient and based on well-developed knowledge than when they are not (e.g., Tesser & Leone, 1977) because prior knowledge guides the recruiting of new thoughts (see Tesser, Martin, & Mendolia, in press, for a review). In a similar fashion, when the object of thought is the self, reflection appears to channel thinking and reveal a clearer picture of who and what one is. From this perspective, our findings offer support for the speculation of earlier theorists (e.g., Bem, 1972; Nisbett & Wilson, 1977) that strong and unambiguous internal cues may foster rather than undermine self-knowledge.

A caveat is in order here, however. Even though a modest amount of reflection may make strong and unambiguous self cues accessible, focusing on why one is as one is (Experiment 3) or reflecting for a very long period of time (Experiment 4) may undermine self-insight by causing people to reflect on information that has little to do with their self-concepts. From this vantage point, the relation of reflection to self-insight is by no means a simple one and hinges on at least three important parameters. First, the nature of the self-knowledge is important; reflection on strong and unambiguous pieces of self-knowledge should be particularly apt to yield insight into the self. Second, the nature of the reflective activity is important; considering what one is should foster self-insight, whereas considering why one is as one is may prove misleading. Third, the amount of reflection is important; although some reflection may be a prerequisite for self-insight, a bit more reflection may sometimes cause people to behave in ways that belie their self-concepts.

⁵ There is a potential alternative explanation for our data that we would like to address. One reviewer suggested that the process of accessing a self-concept and matching it to incoming feedback might not be inherently time consuming (as we have claimed) but rather is simply done faster for positive self-concepts than for negative ones. Such an account would be quite plausible in situations that involve multiple self-concepts but is a much less plausible account of our results in Experiments 1, 3, and 4, which focused on a single self-concept (sociability). When minimally reflective participants with negative social self-concepts endorsed a favorable evaluation or evaluator in those experiments, it was not because they only had time to access a positive self-concept and match it to the feedback—there was no such thing to access.

References

- Bem, D. J. (1972). Self-perception theory. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 6). San Diego, CA: Academic Press.
- Brooks, L. R. (1978). Nonanalytic concept formation and memory for instances. In E. Rosch & B. B. Lloyd (Eds.), *Cognition and categorization*. San Diego, CA: Academic Press.
- Cronbach, L. J. (1955). Processes affecting scores on "understanding

- of others" and "assumed similarity." *Psychological Bulletin*, 52, 177-193.
- Dunning, D., Meyerowitz, J. A., & Holzberg, A. D. (1989). Ambiguity and self-evaluation: The role of idiosyncratic trait definitions in self-serving assessments of ability. *Journal of Personality and Social Psychology*, 57, 1082-1090.
- Ericsson, K. A., & Simon, H. A. (1980). Verbal reports as data. *Psychological Review*, 87, 215-251.
- Funder, D. C. (1980). On seeing ourselves as others see us: Self-other agreement and discrepancy in personality ratings. *Journal of Personality and Social Psychology*, 48, 473-493.
- Helmreich, R., Spence, J. T., & Stapp, J. (1974). Short form of the Texas Social Behavior Inventory, an objective measure of self-esteem. *Bulletin of the Psychonomic Society*, 4, 473-475.
- Howard, J. H., & Ballas, J. A. (1980). Syntactic and semantic factors in the classification of nonspeech transient patterns. *Perception and Psychophysics*, 28, 431-439.
- Kihlstrom, J. F., Cantor, N., Albright, J. S., Chew, B. S., Klein, S. B., & Niedenthal, P. M. (1988). Information processing and the study of self. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 21, pp. 145-177). San Diego, CA: Academic Press.
- Klein, S. B., & Loftus, J. (1988). The nature of self-referent encoding: The contributions of elaborative and organizational processes. *Journal of Personality and Social Psychology*, 55, 5-11.
- Kraut, R. E., & Lewis, S. H. (1982). Person perception and self-awareness: Knowledge of one's influences on one's own judgments. *Journal of Personality and Social Psychology*, 42, 448-460.
- Millar, M. G., & Tesser, A. (1986). Effects of affective and cognitive focus on the attitude-behavior relationship. *Journal of Personality and Social Psychology*, 51, 270-276.
- Millar, M. G., & Tesser, A. (1989). The effects of affective-cognitive consistency and thought on attitude behavior relations. *Journal of Experimental Social Psychology*, 25, 189-202.
- Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review*, 84, 231-259.
- Paulhus, D. L., Graf, P., & Van Selst, M. (1989). *Attention and positivity of self-presentation*. Unpublished manuscript, University of British Columbia, Vancouver, British Columbia, Canada.
- Paulhus, D. L., & Levitt, K. (1987). Desirable responding triggered by affect: Automatic egotism? *Journal of Personality and Social Psychology*, 52, 245-259.
- Pelham, B. W., & Swann, W. B., Jr. (1989). From self-conceptions to self-worth: The sources and structure of self-esteem. *Journal of Personality and Social Psychology*, 57, 672-680.
- Reber, A. S. (1976). Implicit learning of synthetic languages: The role of instructional set. *Journal of Experimental Psychology: Human Learning and Memory*, 2, 88-94.
- Reber, A. S. (1989). Implicit learning and tacit knowledge. *Journal of Experimental Psychology: General*, 118, 219-235.
- Reber, A. S., Kassin, S. M., Lewis, S., & Cantor, G. W. (1980). On the relationship between implicit and explicit modes in the learning of a complex rule structure. *Journal of Experimental Psychology: Human Learning and Memory*, 6, 492-502.
- Rogers, T. B., Kuiper, N. A., & Kirker, W. S. (1977). Self-reference and the encoding of personal information. *Journal of Personality and Social Psychology*, 35, 677-688.
- Schooler, J. W., & Engstler-Schooler, T. Y. (1990). Verbal overshadowing of visual memories: Some things are better left unsaid. *Cognitive Psychology*, 22, 36-71.
- Schooler, J. W., Foster, R. A., & Loftus, E. F. (1988). Some deleterious consequences of the act of recollection. *Memory & Cognition*, 16, 243-251.
- Smith, E. R., & Miller, F. D. (1978). Limits on perception of cognitive processes: A reply to Nisbett and Wilson. *Psychological Review*, 85, 355-362.
- Swann, W. B., Jr. (1987). Identity negotiation: Where two roads meet. *Journal of Personality and Social Psychology*, 53, 1038-1051.
- Swann, W. B., Jr. (1990). To be adored or to be known? The interplay of self-enhancement and self-verification. In R. M. Sorrentino & E. T. Higgins (Eds.), *Motivation and cognition* (Vol. 2, pp. 408-448). New York: Guilford Press.
- Swann, W. B., Jr., Hixon, J. G., & De La Ronde, C. (1992). *Dating games and marital reality: The self-concept and preferences for relationship partners*. Manuscript submitted for publication.
- Swann, W. B., Jr., Hixon, J. G., Stein-Seroussi, A., & Gilbert, D. T. (1990). The fleeting gleam of praise: Behavioral reactions to self-relevant feedback. *Journal of Personality and Social Psychology*, 59, 17-26.
- Taylor, S. E., & Brown, J. D. (1988). Illusion and well being: Some social psychological contributions to a theory of mental health. *Psychological Bulletin*, 103, 193-210.
- Tesser, A., & Leone, C. (1977). Cognitive schemas and thought as determinants of attitude change. *Journal of Experimental Social Psychology*, 13, 340-356.
- Tesser, A., Martin, L., & Mendolia, M. (in press). The role of thought in changing attitude strength. In R. E. Petty & J. A. Krosnick (Eds.), *Attitude strength: Antecedents and consequences*. Hillsdale, NJ: Erlbaum.
- Wilson, T. D., & Dunn, D. S. (1986). Effects of introspection on attitude-behavior consistency: Analyzing reasons versus focusing on feelings. *Journal of Experimental Social Psychology*, 22, 249-263.
- Wilson, T. D., Dunn, D. S., Bybee, J. A., Hyman, D. B., & Rotondo, J. A. (1984). Effects of analyzing reasons on attitude-behavior consistency. *Journal of Personality and Social Psychology*, 47, 5-16.
- Wilson, T. D., Dunn, D. S., Kraft, D., & Lisle, D. J. (1989). Introspection, attitude change, and attitude-behavior consistency: The disruptive effects of explaining why we feel the way we do. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 22, pp. 287-343). San Diego, CA: Academic Press.
- Wilson, T. D., Hodges, S. D., & Pollack, S. E. (1991). *Effects of explaining attitudes on survey responses: Moderating effects of attitude strength*. Unpublished manuscript, University of Virginia, Charlottesville.
- Wilson, T. D., Kraft, D., & Dunn, D. S. (1989). The disruptive effects of explaining attitudes: The moderating effect of knowledge about the attitude object. *Journal of Experimental Social Psychology*, 25, 379-400.
- Wilson, T. D., & Schooler, J. W. (1991). Thinking too much: Introspection can reduce the quality of preferences and decisions. *Journal of Personality and Social Psychology*, 60, 181-192.
- Zajonc, R. B. (1980). Feeling and thinking: Preferences need no inferences. *American Psychologist*, 35, 151-175.

Received January 6, 1992

Revision received June 16, 1992

Accepted June 22, 1992 ■