



Reports

Mutual influences between the implicit and explicit self-concepts: The role of memory activation and motivated reasoning[☆]

Kurt R. Peters^{*}, Bertram Gawronski

The University of Western Ontario, Canada

ARTICLE INFO

Article history:

Received 16 August 2010

Revised 19 November 2010

Available online 4 December 2010

Keywords:

Self-concept
Autobiographical memory
Implicit social cognition
Dual-process theories
Motivated reasoning

ABSTRACT

Whereas previous research has predominantly focused on dissociations between the explicit and implicit self-concepts, the current research investigates how these aspects of self-representation come into correspondence through the activation of information about the self in memory. Experiment 1 provides evidence for a “bottom-up” process of self-construal in which information activated in the implicit self-concept produces congruent changes in the explicit self-concept. Experiment 2 provides evidence for a “top-down” process of self-construal in which the motivated assertion of a propositional belief in the explicit self-concept leads, via a process of confirmatory hypothesis testing, to the activation of substantiating information in the implicit self-concept. These two processes of self-concept change are integrated within a framework that specifies how the explicit and implicit self-concepts are related within an overall, dynamic self-system. Possibilities for expanding the framework to account for self-concept dissociations are discussed.

© 2010 Elsevier Inc. All rights reserved.

Historically, the self-concept has been understood as the collection of things we believe about ourselves. The use of recently developed techniques for indirectly measuring mental contents, however, has suggested that an individual's self-concept can differ depending on how it is measured—specifically, depending on whether information about the self is *explicit* or *implicit* in behavioral responses on the measurement procedure (cf. De Houwer, 2006). Thus, measures of the “explicit” self-concept, typically assessed via self-report, have been shown to diverge from measures of the “implicit” self-concept, typically assessed via performance-based measures (e.g., Greenwald & Farnham, 2000; see Schnabel & Asendorpf, 2010, for a review). Measures of the implicit self-concept predict behavior above and beyond measures of the explicit self-concept (Asendorpf, Banse, & Mücke, 2002; Back, Schmukle, & Egloff, 2009), and this incremental validity appears to derive from a difference between the conditions under which the two types of self-information influence behavior. For example, Asendorpf et al. (2002) found evidence for a double dissociation between explicit and implicit self-concepts, such that the explicit self-concept uniquely predicted controlled behaviors and the implicit self-concept uniquely predicted spontaneous behaviors. Moreover, it has been shown that discrepancies between the explicit and implicit self-concepts on a particular

dimension (e.g., shyness) uniquely predict behaviors intended to reduce these discrepancies (Briñol, Petty, & Wheeler, 2006).

Based on these findings, it appears that measures of explicit and implicit self-concepts do tap different types of information about the self, and that these different self-conceptions can become dissociated. It is currently less clear, however, how these two self-conceptions may correspond. Based on theoretical perspectives that conceive of the self as an integrated system for facilitating adaptive behavior (Cross & Markus, 1990; Steele, 1988), it makes sense to expect these conceptions to be related. In fact, measures of the explicit and implicit self-concepts are typically correlated, suggesting a significant degree of correspondence. For example, Asendorpf et al. (2002), Briñol et al. (2006), and Back et al. (2009) observed correlations in the range of .30 to .40 between measures of the explicit and implicit self-concepts. A meta-analysis of correlations between self-report measures and the Implicit Association Test (IAT; Greenwald, McGhee, & Schwartz, 1998) also found that these assessments of the explicit and implicit self-concepts correlated at .21, more strongly than assessments of self-esteem and roughly equal to the overall correlation between self-report measures and the IAT across research domains (Hofmann, Gawronski, Gschwendner, Le, & Schmitt, 2005). Thus, although explicit and implicit self-conceptions appear to be distinct, it is equally clear that they can correspond, raising the possibility that these two aspects of self-representation are related through processes of mutual influence.

The aim of the present research is to provide a framework for understanding these mutual influences by viewing the explicit and implicit self-concepts as distinct but interacting aspects of an individual's self-representation. Within this framework, measures of the implicit self-concept are assumed to reflect the momentary

[☆] This research was supported by the Ontario Graduate Scholarship Program [000130773] and the Social Sciences and Humanities Research Council of Canada [410-2008-2247].

^{*} Corresponding author. Department of Psychology, The University of Western Ontario, Social Science Centre, London, Ontario, Canada N6A 5C2.

E-mail address: kpeter3@uwo.ca (K.R. Peters).

activation of specific self-associations in memory. Measures of the explicit self-concept are assumed to reflect validated self-beliefs, which are descriptive propositions about the self based on activated self-associations that are regarded as *true* by the individual. From this perspective, the explicit self-concept can be considered a “working” self-concept, in that it constitutes a continuously maintained network of beliefs about the self (Markus & Wurf, 1987). The implicit self-concept provides an online, context-sensitive source of activated information that substantiates, and potentially informs the revision of, this network of self-beliefs. The construction and maintenance of the working self-concept—a process we refer to as *self-construal*—is thus understood as a fundamentally epistemic enterprise involving the generation and validation of self-beliefs (Kruglanski, 1989; Quine & Ullian, 1970).¹

This framework for relating the explicit and implicit self-concepts suggests two specific routes of influence underlying the general process of self-construal. First, a bottom-up, “data-driven” process of self-construal can occur when self-associations in memory are activated without the intention to revise the explicit, working self-concept. The increased accessibility of newly activated self-knowledge will then promote its incorporation into the explicit, working self-concept. Second, a top-down, “hypothesis-driven” process of self-construal can occur when the explicit, working self-concept is intentionally revised, which involves asserting the validity of a new propositional belief about the self (e.g., “I am extraverted”). To test this hypothesis about the self, autobiographical memory can be searched for relevant evidence; however, it is expected that this search will be biased toward activating confirmatory information that substantiates the asserted self-belief (Klayman & Ha, 1987; Kunda, 1990). In either of these cases, the process of self-construal should produce correspondence between the explicit and implicit self-concepts. The key difference is that during bottom-up self-construal, change in the implicit self-concept is expected to mediate change in the explicit self-concept, whereas during top-down self-construal, the reverse mediation is expected.

The following two experiments were designed to test these predictions concerning the mutual influences between the explicit and implicit self-concepts. In Experiment 1, self-associations in memory were activated independently of the intention to revise the explicit, working self-concept as a test of bottom-up self-construal. In Experiment 2, participants were motivated to revise their working self-concepts directly as a test of top-down self-construal. By relating the explicit and implicit self-concepts together within an overall framework of self-construal, this research promises to clarify the connection between these two aspects of self-representation.

Experiment 1

The first experiment tested the proposed account of bottom-up self-construal by asking participants to recall autobiographical memories relating to a specific personality trait (ostensibly as part of a study investigating the relation between personality styles and autobiographical memory). The recall task was intended to activate specific self-associations in memory independent of the intention to revise the working self-concept, thus initiating a process of bottom-up self-construal. To ensure that most participants would possess relevant memories and that the revision of the working self-concept implied by activated self-associations would not be resisted, it was necessary to manipulate a relatively broad and malleable domain of self-knowledge. Toward that end, the trait dimension of extraversion–

introversion was identified as sufficiently fluid to ensure that most participants would be willing and able to see themselves as more or less extraverted (Sanitioso, Kunda, & Fong, 1990).

The current experiment tested three specific predictions derived from the framework of self-construal outlined above. First, the measure of the implicit self-concept (in this case, a self-concept IAT; Greenwald & Farnham, 2000) was expected to reflect the selective activation of self-associations congruent with the personality trait (i.e., extraversion or introversion) targeted in the memory recall task. Second, the measure of the explicit, working self-concept was expected to reflect the revision of beliefs about the self in line with the recalled memories, such that participants recalling extraverted (or introverted) memories would report more (or fewer) extraversion-related self-beliefs. Finally, it was predicted that these changes in the explicit, working self-concept would be mediated by changes in the activation of self-associations in the implicit self-concept, consistent with the proposed account of bottom-up self-construal.

Method

Sample and design

A total of 118 undergraduate students (80 women and 38 men) participated in a study on personality and autobiographical memory for course credit. The experimental design consisted of a single between-subjects factor with two conditions (Recalled Trait: Extraversion vs. Introversion). Order of the two dependent measures was counterbalanced across participants.

Memory recall task

Upon entering the lab, participants were seated at individual computer carrels and given informed consent documents to sign. Participants then began the memory recall task, which guided them through the process of sequentially recalling and describing two memories of their past behavior that they considered to be extraverted or introverted, according to the experimental condition. To encourage recalled behaviors to be interpreted as arising from the self rather than situational influences, participants were instructed to recall each memory using an observer’s (as opposed to an actor’s) visual perspective (Libby, Eibach, & Gilovich, 2005). For each memory, once the participant indicated that the requested memory had been recalled, a series of brief questions was asked to increase its vividness (e.g., “Can you see what your facial expression was?”; Libby et al., 2005). Participants were then asked to describe the recalled memory briefly in writing, again using an observer’s visual perspective.

Measurement of implicit self-concept

A “self/extravert” IAT was used to assess the selective activation of trait-related self-knowledge following the memory recall task (see Appendix A for stimuli). The IAT compares reaction times to responses that pair a target (e.g., *me*) with an attribute (e.g., *extraverted*) against responses that pair the same target (e.g., *me*) with a complementary attribute (e.g., *introverted*). The resulting difference score provides a sample-relative index of the degree to which target–attribute associations are activated in memory. In the first block of the IAT, “me” and “not me” words had to be assigned to the categories *Me* (right) and *Not Me* (left). In the second block, extraversion and introversion words had to be assigned to the categories *Extravert* (right) and *Introvert* (left). In the third block, target and attribute trials were presented in alternating order, with “me” and extraversion words on the right and “not me” and introversion words on the left. In the fourth block, participants practiced categorizing only extraversion and introversion words with key assignments reversed. In the fifth block, target and attribute trials were again combined, with “me” and introversion words on the right and “not me” and extraversion words on the left. Blocks 1, 2, and 4 consisted of 20 trials, and blocks 3 and 5 consisted of 80 trials. The inter-trial interval was 250 ms. Following

¹ In line with the broader use of the term *construal* in the social-cognitive literature, we use the term *self-construal* to refer to a general process of constructing beliefs based on momentarily accessible information. Previous use of the term *self-construal* to refer to the influence of culture on self-definition (e.g., Markus & Kitayama, 1991) can thus be understood within the present framework as self-construal in a particular content domain, whereby cultural factors influence the accessibility and desirability of specific self-information (e.g., independent versus interdependent self-characteristics).

incorrect responses the word “ERROR!” was presented in the center of the screen for 1000 ms.

Measurement of explicit self-concept

A self-report rating scale was used to measure participants' perceptions of their own personality traits, which are assumed to reflect the self-beliefs constituting the explicit, working self-concept. To mitigate demand effects following the memory recall task, participants were told that the researchers were “also interested in how a variety of personality dimensions influence recalled memories” and the scale was therefore presented as a general personality assessment. The scale consisted of six items relating to extraversion and six items relating to introversion (identical to the stimuli used in the IAT), along with six positively valenced filler items and six negatively valenced filler items (see Appendix B). The items were presented in an a priori randomized order and were rated on a 7-point scale.

All participants were debriefed at the completion of the experiment. None indicated suspicion of a link between the memory recall task and either of the dependent measures.

Results

Data preparation

An index of extraversion-related (vs. introversion-related) self-associations was calculated from responses in the IAT following Greenwald, Nosek, and Banaji's (2003) *D*-600 algorithm (Cronbach's $\alpha = .84$). Scores were calculated such that higher values reflect stronger associations between the self and extraversion (compared to introversion) in the implicit self-concept. An index of extraversion-related (vs. introversion-related) self-beliefs was calculated from the self-report scale by reverse-coding the six introversion-related items and computing the combined mean of the six extraversion-related items with the six reverse-coded introversion items (Cronbach's $\alpha = .92$). Higher scores therefore reflect a more extraverted (compared to introverted) explicit, working self-concept. The index of activated self-knowledge and the index of self-beliefs were significantly correlated, $r = .53$, $p < .001$.

Effects of memory recall task

Inspection of participants' written descriptions of recalled memories suggested that they complied with instructions to recall the requested extraversion- or introversion-related memories. Means and standard deviations for the two self-concept measures are presented in Table 1. As predicted, participants who recalled extraversion-related memories revealed significantly stronger associations between the self and extraversion (relative to introversion) on the IAT compared to participants who recalled introversion-related memories, $t(116) = 2.54$, $p = .01$, $d = .47$. Similarly, and again in line with predictions, participants who recalled extraversion-related memories reported significantly more extraverted (relative to introverted) self-beliefs than participants who recalled introversion-related memories, $t(116) = 1.97$, $p = .05$, $d = .37$.

Mediation analysis

The third prediction tested in the current experiment was that the activation of trait-related self-associations in memory would mediate

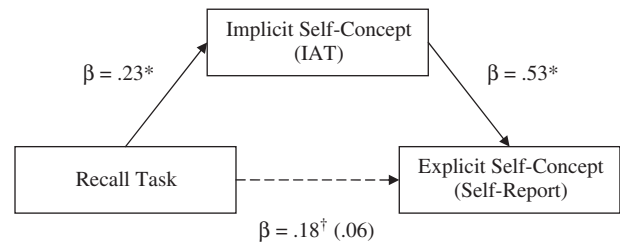


Fig. 1. Mediation model tested in Experiment 1 (on the basis of Baron & Kenny, 1986). The indirect effect of the recall task on the explicit self-concept (mediated by the implicit self-concept) is statistically significant, Sobel's $Z = 2.38$, $p = .02$. $^{\dagger}p = .05$; $^*p < .05$.

the effect of the recall task on the explicit, working self-concept, in line with the proposed account of bottom-up self-construal (see Fig. 1). To test this prediction, self-report scores were simultaneously regressed on both the memory recall task and IAT scores (Baron & Kenny, 1986). The relation between IAT scores and self-report scores remained significant, $\beta = .51$, $t(115) = 6.31$, $p < .001$, but the effect of the recall task became non-significant, $\beta = .06$, $t(115) = .77$, $p = .45$. Thus, changes in the implicit self-concept fully accounted for changes in the explicit self-concept. The indirect effect of the memory recall task on self-report scores when IAT scores were included as a mediator was significant, Sobel's $Z = 2.38$, $p = .02$.

To rule out the alternative possibility of a top-down influence, we also tested the reverse mediation model, in which IAT scores were simultaneously regressed on both the memory recall task and self-report scores. In this mediation model, changes in self-report scores failed to account for the obtained effect on IAT scores, in that the memory recall task still had a marginally significant effect on IAT scores after controlling for self-report scores, $\beta = .14$, $t(115) = 1.76$, $p = .08$. Thus, changes in the implicit self-concept fully accounted for changes in the explicit self-concept, but not the other way around.

Discussion

The results of Experiment 1 provided support for all three predictions concerning bottom-up self-construal. The memory recall task selectively activated trait-related self-associations in the implicit self-concept and also led to congruent revision of the explicit, working self-concept. The mediation analysis supported the prediction that the effect of the recall task on the explicit self-concept would be mediated by activation of self-associations in the implicit self-concept. Consistent with this prediction, changes in self-associations fully accounted for changes in explicit self-beliefs, but not the other way around. Taken together, these findings suggest that differences in participants' explicit, working self-concepts between the memory recall conditions were due to the bottom-up integration of the self-beliefs implied by activated self-associations in memory.

Experiment 2

The second experiment was designed to test the proposed account of top-down self-construal. In contrast to bottom-up self-construal, in which the explicit, working self-concept is influenced independent of

Table 1

Means and standard deviations by condition for measures of explicit and implicit self-concepts in Experiments 1 and 2.

Measure	Experiment 1				Experiment 2			
	Recall extraversion		Recall introversion		Pro-Extraversion		Pro-Introversion	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Explicit self-concept	4.64	1.10	4.24	1.09	4.98	1.04	4.53	1.04
Implicit self-concept	.61	.56	.35	.53	.61	.45	.44	.52

the intention to revise it, top-down self-construal begins with the intentional revision of the working self-concept. Thus, participants in the current experiment were asked to generate explanations for (fabricated) scientific findings linking either extraversion or introversion to successful life outcomes. By making a specific personality trait desirable, this manipulation was intended to motivate participants to assert the validity of the corresponding propositional belief about themselves (i.e., “I am extraverted” or “I am introverted”). Having marked this propositional belief as valid, participants were expected to treat it as a hypothesis to be tested by searching autobiographical memory for relevant evidence. Due to the confirmatory nature of hypothesis testing (Klayman & Ha, 1987), however, participants were expected to selectively activate self-associations in memory that would substantiate the asserted propositional belief (Sanitioso et al., 1990), leading to corresponding effects on the explicit and implicit self-concepts.

The current experiment tested three specific predictions derived from the proposed framework of self-construal. First, the measure of the explicit, working self-concept was expected to reflect the revision of beliefs about the self in line with the motivation manipulation, such that participants motivated to see themselves as extraverted (or introverted) would report more (or fewer) extraversion-related self-beliefs. Second, the measure of the implicit self-concept (a self-concept IAT) was expected to reflect the selective activation of self-associations congruent with the personality trait (i.e., extraversion or introversion) that participants were motivated to believe they possessed. Finally, it was predicted that these changes in the implicit self-concept would be mediated by changes in the explicit, working self-concept, consistent with the proposed account of top-down self-construal.

Method

Sample and design

A total of 148 undergraduate students (111 women and 37 men) participated in a study on personality and explanation styles for course credit. The experimental design consisted of a single between-subjects factor with two conditions (Desired Trait: Pro-Extraversion vs. Pro-Introversion). Order of the two dependent measures was counterbalanced across participants.

Motivation induction task

Upon entering the lab, participants were seated at individual computer carrels and given informed consent documents to sign. Participants then began the motivation induction task (adapted from Sanitioso et al., 1990). The task was framed as an investigation of how people generate explanations for scientific findings. As with Experiment 1, the current experiment manipulated self-perceptions along the extraversion–introversion dimension of personality. Participants were presented with a fabricated newspaper clipping reporting the findings of a recent scientific study comparing the benefits of extraverted personality traits with introverted personality traits (Appendix C). The clipping briefly described a study that found that extraversion leads to more academic and job success than introversion (or vice versa in the Pro-Introvert condition). After reading the clipping, participants were asked to write down two brief explanations for the observed relationship between extraversion (or introversion) and positive life outcomes. Generating the two explanations was intended to strengthen the manipulation and reinforce the cover story. Because most people presumably desire to see themselves as successful in life, the newspaper clipping was expected to motivate participants to hypothesize that they themselves possessed the personality trait that contributed to positive life outcomes, thereby initiating a process of top-down self-construal.

Dependent measures

The measures of the implicit and explicit self-concepts were identical to those used in Experiment 1.

Control measure and manipulation check

A manipulation check was included to ensure that the motivation induction task influenced the desirability of extraverted and introverted personality traits. For three of the extraversion-related and three of the introversion-related trait words used in the measurement of the explicit self-concept, participants were asked to indicate on a 7-point scale how much that trait contributed to success after university.

All participants were debriefed upon completion of the experiment. None indicated suspicion of a link between the motivation induction task and either of the dependent measures.

Results

Data preparation

Indices of extraversion-related (vs. introversion-related) self-associations (Cronbach's $\alpha = .76$) and of extraversion-related (vs. introversion-related) self-beliefs (Cronbach's $\alpha = .91$) were calculated as described in Experiment 1. The two indices were significantly correlated, $r = .46$, $p < .001$. For the manipulation check, an index of the degree to which extraversion vs. introversion contributes to positive life outcomes was calculated by reverse-coding the three introversion-related items and computing the combined mean with the three extraversion-related items (Cronbach's $\alpha = .85$). Higher scores thus reflect increased desirability of extraversion compared to introversion.

Manipulation check

The motivation induction task led participants in the Pro-Extravert condition to report that extraversion was more desirable ($M = 5.82$, $SD = .75$) than participants in the Pro-Introvert condition ($M = 5.10$, $SD = .78$), $t(146) = 5.75$, $p < .001$, $d = .95$. Thus, given that most people desire positive life outcomes for themselves, it is reasonable to assume that the induction task indeed motivated participants to perceive themselves as possessing more extraverted or introverted qualities, according to the experimental condition.

Effects of motivation induction task

Means and standard deviations for the primary measures are presented in Table 1. Participants in the Pro-Extravert condition revealed significantly stronger associations between the self and extraversion (relative to introversion) on the IAT compared to participants in the Pro-Introvert condition, $t(146) = 2.10$, $p = .04$, $d = .35$. Similarly, and again in line with predictions, participants in the Pro-Extravert condition reported significantly more extraverted (relative to introverted) self-beliefs than participants in the Pro-Introvert condition, $t(146) = 2.65$, $p = .01$, $d = .44$.

Mediation analysis

The third prediction tested in the current experiment was that asserting the validity of a propositional belief within the explicit, working self-concept would initiate a biased search through memory to activate substantiating self-associations, thereby mediating the effect of the motivation induction on the implicit self-concept (see Fig. 2). To test this prediction, IAT scores were simultaneously regressed on both the motivation induction and self-report scores (Baron & Kenny, 1986). The relation between self-report scores and IAT scores remained significant, $\beta = .45$, $t(145) = 5.97$, $p < .001$, but the effect of the motivation induction became non-significant, $\beta = .08$, $t(145) = 1.00$, $p = .32$. Thus, changes in the explicit self-concept fully accounted for changes in the implicit self-concept. The indirect effect

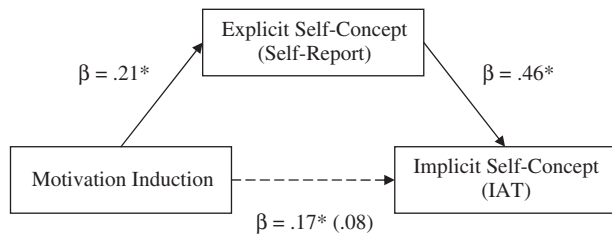


Fig. 2. Mediation model tested in Experiment 2 (on the basis of Baron & Kenny, 1986). The indirect effect of the motivation induction on the implicit self-concept (mediated by the explicit self-concept) is statistically significant, Sobel's $Z = 2.44$, $p = .02$. * $p < .05$.

of the motivation induction on IAT scores when self-report scores were included as a mediator was significant, Sobel's $Z = 2.44$, $p = .02$.

To rule out the alternative possibility of a bottom-up influence, we also tested the reverse mediation model, in which self-report scores were simultaneously regressed on both the motivation induction and the IAT scores. In this mediation model, changes in IAT scores failed to account for the obtained effect on self-report scores, in that the motivation induction still had a marginally significant effect on self-report scores after controlling for IAT scores, $\beta = .14$, $t(115) = 1.88$, $p = .06$. Thus, changes in the explicit self-concept fully accounted for changes in the implicit self-concept, but not the other way around.

Discussion

The results of Experiment 2 provided support for all three predictions concerning top-down self-construal. The motivation induction led to revision of the explicit, working self-concept in line with the desired trait and selectively activated congruent trait-related self-associations in the implicit self-concept. The mediation analysis supported the prediction that the effect of the motivation induction on the implicit self-concept would be mediated by changes in the explicit self-concept. Consistent with this prediction, changes in explicit self-beliefs fully accounted for changes in self-associations, but not the other way around. Taken together, these findings suggest that differences in participants' implicit self-concepts between motivation induction conditions were due to the top-down, intentional activation of self-associations in memory to substantiate the assertion of a propositional belief within the explicit, working self-concept.

General discussion

The current experiments were designed to test a framework specifying how the explicit and implicit self-concepts are related through processes of mutual influence. The results of these two experiments provide converging evidence for the predictions derived from this framework regarding the roles of memory activation and motivated reasoning in achieving correspondence between these two aspects of self-representation. In Experiment 1, participants recalled specific autobiographical memories independent of the intention to revise the explicit, working self-concept. The recall task produced congruent changes in the implicit and explicit self-concepts and, consistent with the proposed account of bottom-up self-construal, changes in the implicit self-concept fully mediated changes in the explicit self-concept. In Experiment 2, participants were motivated to revise their explicit, working self-concepts directly by asserting the validity of a propositional self-belief. The induced motivation produced congruent changes in the explicit and implicit self-concepts; however, in this case—consistent with the proposed account of top-down self-construal—changes in the explicit self-concept fully mediated changes in the implicit self-concept. These results together support the claim that the explicit and implicit self-concepts are integrated, interacting aspects of a dynamic self-system.

Although the mediation analyses generally confirmed our predictions about bottom-up and top-down self-construal, a potential concern is that the reverse mediation models in both experiments revealed evidence for partial mediation (for similar findings, see Gawronski & Walther, 2008; Whitfield & Jordan, 2009). Specifically, the reverse mediation models showed simultaneous direct and indirect effects that were close to or at statistical significance.² These data patterns reflect an inherent limitation of correlation-based approaches to mediation, in which mediation is established on the basis of the shared covariance between two measured variables and an independent variable (cf. Zhao, Lynch, & Chen, 2010). The possibility of partial mediation, however, becomes theoretically implausible when examined alongside evidence from the predicted mediation models. In Experiment 1, the effect of the recall task on self-associations fully accounted for changes in explicit self-beliefs. Likewise, in Experiment 2, the effect of the motivation induction on explicit self-beliefs fully accounted for changes in self-associations. If the current manipulations influenced our dependent measures through processes of partial mediation, the proposed mediators in the two experiments would be unable to fully account for changes in the proposed distal outcomes. Rather, there should still be a direct effect on the distal outcome in the predicted mediation model after controlling for the proposed mediator. For instance, if the data in Experiment 1 reflected the operation of a direct influence on self-associations and a simultaneous indirect influence on self-associations mediated by a direct influence on explicit self-beliefs, the obtained effect on explicit self-beliefs should remain significant after controlling for self-associations. Similarly, if the data in Experiment 2 reflected the operation of a direct influence on self-beliefs and a simultaneous indirect influence on self-beliefs mediated by a direct influence on self-associations, the obtained effect on self-associations should remain significant after controlling for self-beliefs. This was not, however, the case. Instead, changes in self-associations fully accounted for the obtained effect on self-beliefs in Experiment 1, and changes in self-beliefs fully accounted for the obtained effect on self-associations in Experiment 2. These results are consistent with the current hypotheses of bottom-up and top-down construal, but they are inconsistent with the alternative possibility of partial mediation.³

Correspondence vs. dissociation

The present research emphasized the correspondence between measures of the explicit and implicit self-concepts, in contrast to previous research that has emphasized their dissociation (Schnabel & Asendorpf, 2010). An obvious question for the proposed framework of self-construal, then, is how to account for such dissociations. To begin with, the results of the current experiments provide evidence for mutual influences between the explicit and implicit self-concepts via a *knowledge-activation* process (Förster & Liberman, 2007). All else being equal, the activation of self-knowledge, whether occurring via a process of bottom-up or top-down self-construal, should increase correspondence between the explicit and implicit self-concepts. This correspondence may break down during either of these processes,

² In Experiment 1, the reverse mediation model revealed a marginally significant direct effect, $\beta = .14$, $t(115) = 1.76$, $p = .08$, and a marginally significant indirect effect, Sobel's $Z = 1.89$, $p = .06$; in Experiment 2, the reverse mediation model revealed a marginally significant direct effect, $\beta = .14$, $t(115) = 1.88$, $p = .06$, and a significant indirect effect, Sobel's $Z = 1.99$, $p = .05$.

³ Further evidence for our mediation hypotheses could be obtained through experimental approaches that do not rely on simple covariations between the mediator and the distal outcome (Spencer, Zanna, & Fong, 2005). One option is to experimentally manipulate the effect of the proposed mediator on the distal outcome (see Gawronski & LeBel, 2008, for an example). To the extent that the effect of the mediator on the distal outcome can be disrupted, the effect of the original manipulation (e.g., the motivation induction) should remain intact for the mediator (e.g., explicit self-concept), but it should disappear for the distal outcome (e.g., implicit self-concept).

however, resulting in a dissociation. First, in the case of bottom-up self-construal, the influence of self-associations on explicit self-beliefs is likely moderated by a *belief-validation* process. Following Gawronski and Bodenhausen (2006, in press), the belief-validation process is expected to operate according to principles of cognitive consistency, such that activated self-associations that are inconsistent with other (subjectively valid) self-beliefs may be rejected as invalid information about the self. Whereas validation of activated self-associations should increase the correspondence between the explicit and implicit self-concepts, invalidation should result in a dissociation within the relevant domain of self-knowledge. Second, in the case of top-down self-construal, dissociations may arise when new beliefs are asserted as valid within the explicit, working self-concept, but are not substantiated via selective activation of confirmatory self-associations. Thus, whereas selective activation of confirmatory self-associations should increase the correspondence between the explicit and implicit self-concepts, disrupting the process of confirmatory information search should lead to a dissociation.

Accounting for both correspondence and dissociation between the explicit and implicit self-concepts suggests a more comprehensive framework for understanding self-construal as an epistemic enterprise, characterized in terms of the basic principles of knowledge-activation and belief-validation (Gawronski, LeBel, & Peters, 2007). Such a framework has the potential to clarify both how the explicit and implicit self-concepts correspond and how they become dissociated.

Future directions

Based on the above discussion, an important next step in the development of this framework is to investigate the proposed account of self-concept dissociations. In particular, the framework predicts that the overall self-system comprising the explicit and implicit self-concepts can become “unbalanced” when the processes that maintain correspondence break down. On the one hand, inconsistent beliefs implied by self-associations activated within the implicit self-concept may not be validated for incorporation into the explicit, working self-concept (though self-associations may nevertheless influence spontaneous behaviors; Asendorpf et al., 2002). On the other hand, propositional beliefs asserted within the explicit, working self-concept may remain unsubstantiated if the activation of confirmatory self-associations in memory is interrupted. The resulting discrepancies may promote uncertainty in self-definition (Briñol et al., 2006) and compensatory behaviors intended to substantiate the asserted self-beliefs (e.g., Wicklund & Gollwitzer, 1981). Thus, the current framework not only offers specific predictions about the mutual influences between these two aspects of self-representation, but also integrates earlier findings on the dynamics of the explicit and implicit self-concepts, providing intriguing directions for future research.

Appendix A. Implicit Association Test stimuli

The following tables list the stimuli used in the Implicit Association Test (IAT) for the target (“me” vs. “not me”) and attribute (“extravert” vs. “introvert”) categories in Experiments 1 and 2.

Target words		Attribute words	
“Me”	“Not me”	“Extravert”	“Introvert”
I	Few	Active	Passive
Me	Some	Talkative	Quiet
My	Any	Sociable	Withdrawn
Mine	It	Outgoing	Private
Self	Other	Assertive	Reserved

Appendix B. Self-report scale of self-perceived personality traits

For each of the 24 items in the scale, subjects rated the statement “I am X,” where X was one of the personality traits below, on a 7-point agree/disagree scale. The same scale was used in Experiments 1 and 2.

Extraversion	Introversion	Filler (positive)	Filler (negative)
Active	Passive	Curious	Anxious
Talkative	Quiet	Disciplined	Impulsive
Sociable	Withdrawn	Generous	Selfish
Outgoing	Reserved	Humorous	Dishonest
Assertive	Private	Optimistic	Cynical
Extraverted	Introverted	Rational	Superstitious

Appendix C. Manipulation of motivation in Experiment 2

The following text was presented in the form of a fake newspaper clipping (adapted from Sanitioso et al., 1990). The text used in the Pro-Extravert condition is shown here; in the Pro-Introvert condition, all references to extraversion and introversion were switched.

Extraverts get ahead

A recent study at Stanford University investigating the effects of extraverted personality on academic and job success has concluded that outgoing people are more successful than their less talkative peers. Dr. Brian Carswell, the lead investigator, reports that although roughly equal numbers of extraverts and introverts exist in the population, being extraverted appears to predict success in these settings to a high degree. In particular, Carswell and his colleagues found that extraverts tend to receive higher grades in school and are more likely to earn graduate and professional degrees compared to introverts. Carswell also reports that extraverted individuals are more likely to end up in successful, high-paying careers. “Extraversion appears to confer distinct advantages in the modern world,” Carswell said.

References

Asendorpf, J. B., Banse, R., & Mücke, D. (2002). Double dissociation between explicit and implicit personality self-concept: The case of shy behavior. *Journal of Personality and Social Psychology, 83*, 380–393.

Back, M. D., Schmukle, S. C., & Egloff, B. (2009). Predicting actual behavior from the explicit and implicit self-concept of personality. *Journal of Personality and Social Psychology, 97*, 533–548.

Baron, R. M., & Kenny, D. A. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology, 51*, 1173–1182.

Briñol, P., Petty, R. E., & Wheeler, S. C. (2006). Discrepancies between explicit and implicit self-concepts: Consequences for information processing. *Journal of Personality and Social Psychology, 91*, 154–170.

Cross, S. E., & Markus, H. R. (1990). The willful self. *Personality and Social Psychology Bulletin, 16*, 726–742.

De Houwer, J. (2006). What are implicit measures and why are we using them? In R. Wiers, & A. Stacy (Eds.), *Handbook of implicit cognition and addiction* (pp. 11–28). Thousand Oaks, CA: Sage.

Förster, J., & Liberman, N. (2007). Knowledge activation. In A. W. Kruglanski, & E. T. Higgins (Eds.), *Social psychology: Handbook of basic principles* (pp. 201–231). (2nd ed.). New York: Guilford Press.

Gawronski, B., & Bodenhausen, G. V. (2006). Associative and propositional processes in evaluation: An integrative review of implicit and explicit attitude change. *Psychological Bulletin, 132*, 692–731.

Gawronski, B., & Bodenhausen, G. V. (in press). The associative-propositional evaluation model: Theory, evidence, and open questions. *Advances in Experimental Social Psychology*.

Gawronski, B., & LeBel, E. P. (2008). Understanding patterns of attitude change: When implicit measures show change, but explicit measures do not. *Journal of Experimental Social Psychology, 44*, 1355–1361.

Gawronski, B., LeBel, E. P., & Peters, K. R. (2007). What do implicit measures tell us? Scrutinizing the validity of three common assumptions. *Perspectives on Psychological Science, 2*, 181–193.

- Gawronski, B., & Walther, E. (2008). The TAR effect: When the ones who dislike become the ones who are disliked. *Personality and Social Psychology Bulletin*, 34, 1276–1289.
- Greenwald, A. G., & Farnham, S. D. (2000). Using the implicit association test to measure self-esteem and self-concept. *Journal of Personality and Social Psychology*, 79, 1022–1038.
- Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology*, 74, 1464–1480.
- Greenwald, A. G., Nosek, B. A., & Banaji, M. R. (2003). Understanding and using the Implicit Association Test: I. An improved scoring algorithm. *Journal of Personality and Social Psychology*, 85, 197–216.
- Hofmann, W., Gawronski, B., Gschwendner, T., Le, H., & Schmitt, M. (2005). A meta-analysis on the correlation between the Implicit Association Test and explicit self-report measures. *Personality and Social Psychology Bulletin*, 31, 1369–1385.
- Klayman, J., & Ha, Y. -W. (1987). Confirmation, disconfirmation, and information in hypothesis testing. *Psychological Review*, 94, 211–228.
- Kruglanski, A. (1989). *Lay epistemics and human knowledge: Cognitive and motivational bases*. New York: Plenum Press.
- Kunda, Z. (1990). The case for motivated reasoning. *Psychological Bulletin*, 108, 480–498.
- Libby, L. K., Eibach, R. P., & Gilovich, T. (2005). Here's looking at me: The effect of memory perspective on assessments of personal change. *Journal of Personality and Social Psychology*, 88, 50–62.
- Markus, H., & Kitayama, S. (1991). Culture and the self: Implications for cognition, emotion, and motivation. *Psychological Review*, 98, 224–253.
- Markus, H., & Wurf, E. (1987). The dynamic self-concept: A social psychological perspective. *Annual Review of Psychology*, 38, 299–337.
- Quine, W. V. O., & Ullian, J. S. (1970). *The web of belief*. New York: Random House.
- Sanitioso, R., Kunda, Z., & Fong, G. T. (1990). Motivated recruitment of autobiographical memories. *Journal of Personality and Social Psychology*, 59, 229–241.
- Schnabel, K., & Asendorpf, J. B. (2010). The self-concept: New insights from implicit measurement procedures. In B. Gawronski, & B. K. Payne (Eds.), *Handbook of implicit social cognition: Measurement, theory, and applications* (pp. 408–425). New York: Guilford Press.
- Spencer, S. J., Zanna, M. P., & Fong, G. T. (2005). Establishing a causal chain: Why experiments are often more effective than meditational analyses in examining psychological processes. *Journal of Personality and Social Psychology*, 89, 845–851.
- Steele, C. M. (1988). The psychology of self-affirmation: Sustaining the integrity of the self. In L. Berkowitz (Ed.), *Advances in experimental social psychology*, Vol. 21. (pp. 261–302). San Diego, CA: Academic Press.
- Whitfield, M., & Jordan, C. H. (2009). Mutual influences of explicit and implicit attitudes. *Journal of Experimental Social Psychology*, 45, 748–759.
- Wicklund, R. A., & Gollwitzer, P. M. (1981). Symbolic self-completion, attempted influence, and self-deprecation. *Basic and Applied Social Psychology*, 2, 89–114.
- Zhao, X., Lynch, J. G., & Chen, Q. (2010). Reconsidering Baron and Kenny: Myths and truths about mediation analysis. *Journal of Consumer Research*, 37, 197–206.