Cognitive Mechanisms Underlying Implicit Negative Self Concept in Dysphoria

Gal Sheppes and Nachshon Meiran
Ben-Gurion University of the Negev

Eva Gilboa-Schechtman
Bar-Ilan University and Gonda Brain Sciences Center

Golan Shahar
Ben-Gurion University of the Negev and Yale University School of Medicine

Whereas explicit measures of the self-concept typically demonstrate a negative bias in depressed individuals, implicit measures such as the Implicit Association Test (IAT), revealed an opposite, positive bias. To address this inconsistent pattern, the authors examined, using a novel paradigm, mental set maintenance (i.e., the difficulty of maintaining active a required mental set) and set operation (the efficiency of executing the mental set while it is maintained). Dysphoric ($N = 33$) and nondysphoric ($N = 30$) participants alternated between an IAT focusing on self reference and a matched neutral task. Nondysphorics had greater difficulty in maintaining a negative self reference task compared to a neutral task. Conversely, dysphorics did not exhibit such difficulty, and they maintained a negative self-reference task more easily than nondysphorics. No group differences were evinced in smoothness of set operation. These results suggest that the shield protecting nondysphorics from maintaining negative mental sets is absent in dysphorics.

Keywords: Implicit Association Test, dysphoria, task switching, self-concept, attentional control

There is a widespread agreement that dysphoric individuals possess a more negative self-concept than nondysphoric individuals. Indeed, such a negative self concept is a corner stone in major theoretical and clinical descriptions of depression and dysphoria (e.g., Beck, Rush, Shaw, & Emery, 1979; Abramson, Seligman, & Teasdale, 1978; *DSM-IV-TR*, 2000). For example, in Beck's theory of depression (1976) one vertex of the cognitive triad involves negative thoughts about the self. Similarly, hopelessness theory of depression (e.g., Abramson, Metalsky, & Alloy, 1989) highlights construing negative events as implying that the self is unworthy or deficient. Support for these theoretical views was clearly demonstrated in research using self-report measures (see Ingram, Miranda, & Segal, 1998, for review).

Surprisingly, in studies using implicit measures of self concept, depressed and nondepressed individuals alike showed a *positive* bias in their self concept. For example, Gemar, Segal, Sagrati, and Kennedy (2001) employed the Implicit Association Test (IAT, Greenwald & Farnham, 2000; Greenwald, McGhee, & Schwartz,

1998) to study the implicit self concept of a depressed, a formerly depressed, and a control group. In this version of the IAT, participants used the same two response keys to classify self-relevant words (e.g., participant's first name) and non self-relevant words (e.g., someone's else first name), as well as positively (e.g., "happy") and negatively (e.g., "sad") valenced words. Two conditions were compared: (1) a "congruent" condition, in which a single key was associated with both self and positive words and a another key was associated with nonself and negative words; and (2) an "incongruent" condition, in which negative and self words, and positive and nonself were respectively paired. Gemar et al. (2001) found faster responses in the congruent condition as compared to the incongruent condition, for depressed, formerly depressed, and control participants. According to the rationale of the IAT, these findings indicate that depressed, like nondepressed individuals, associate the self with positive attributes, a conclusion that stands in sharp contrast to the self report findings. Recently, De Raedt, Schacht, Franck, and De Houwer (2006) obtained similar results using the IAT and other implicit measures, and proposed that depressed individuals might have latent positive self-schemas.

Resolving the Paradox: Set Maintenance Versus Set Operation

Herein we propose a different approach to this inconsistent pattern of results. Utilizing insights from the task-switching literature, we rely on the notion of "mental sets," defined as a complex cognitive configuration which readies the system to perform a particular task (see Logan, 2003; Monsell, 2003, for reviews). The task-switching literature distinguishes between *set maintenance*, difficulties in initiating and maintaining mental sets, as opposed to *set operation*, effi-

ciency of operation of the mental set while it is maintained (e.g., Monsell, 2003; Rubinstein, Meyer, & Evans, 2001). In nontechnical terms, the distinction is between how difficult it is to start a task or keep doing the task (set maintenance) on the one hand and how successful task performance is once the task is being executed (set operation). Consider, for example, two situations in which you have recently reviewed a paper for a journal and also submitted a paper for publication. You receive two action letters in the mail. Both of them indicate blunt rejection and both of them contain information beside the final decision such as points of strength and weakness. These points are simple to comprehend, short, and to the point. Set operation, in this case, refers to how successful you would be in comprehending the information in the letters (which, as said, is pretty straightforward and simple in both cases). Set maintenance refers to how difficult it is to force yourself to read all the details of the letters (which, in this case, is likely to be considerably harder if you read the letter rejecting your manuscript).

To enable the separate assessment of set maintenance and set operation, researchers let participants switch between simple cognitive tasks. A common experimental design involves switching between runs of trials, each involving a different task. For example, with two Tasks A and B, a trial sequence with run-length = 3 would be AAA-BBB-AAA-BBB and so forth. According to this literature, set maintenance difficulties are mostly evident in the first trial in the run following a previous run, called a switch trial. This trial is typically associated with slower and more error prone responses as compared to the remaining trials in the run, called "nonswitch trials." The difference in performance between switch and nonswitch trials indexes "switching cost." Note that the typical design of task-switching experiments involves trivially easy tasks, so that the source of difficulty is not in the task itself (set operation) but in set maintenance.

Somewhat counterintuitively, the literature suggests that a major determinant of the size of the switching cost is the maintenance difficulty of the trials which *preceded* the task switch. In the trial sequence above, switch trials in Task B (BBB) index the difficulty of maintaining the task set that is required to perform Task A which preceded it. This notion is called Task Set Inertia (TSI, Allport, Styles, & Hsieh, 1994). The idea here is that mental sets that are difficult to maintain require a heightened state of activation for a substantial amount of time. Consequently, if the next trial requires one to adopt a different mental set from the previous trial, the lingering activation of the previous mental set causes performance in the following task to deteriorate.

The TSI notion is supported by experiments demonstrating that the degree of conflict in the preswitch task dictates the switch cost (Allport et al., 1994; Allport & Wylie, 2000). An especially convincing demonstration comes from a recent neuro-imaging study by Yeung, Nystrom, Aronson, and Cohen (2006). In that study, participants switched between a color task and a face task, which are known to involve topographically different regions in the temporal lobe of the cortex. Importantly, individual differences in the behavioral switch costs in each task (e.g., in the face task) were strongly correlated with the blood oxygenation levels in the region associated with the alternative task (e.g., color). To conclude, three conditions must be met in order to evaluate set maintenance difficulty. First, the paradigm must involve frequent switching between two tasks. This condition is a prerequisite for the second and third conditions which include: (b) the separate

evaluation of the first trial in the run relative to the remaining trials in the run and (c) the difficulty of maintaining a set is reflected in the response slowing observed in the first trial of the next task.

Taking this distinction into account, consider again the IAT which (broadly described) involves several serial conditions: (a) practice of the self = negative condition (self is mapped with negative attributes), (b) performing the self = negative condition, (c) practice of the self = positive condition (self is mapped with positive attributes), (d) performing the self = positive condition. As can be clearly seen, this design does not involve task switching between the two conditions, aside from the single change which takes place in the middle of the experiment (for a further discussion of IAT in the context of task switching, see Klauer & Mierke, 2005). Consequently, the IAT does not meet the perquisite condition of frequent task switching and as such it cannot assess set maintenance. Furthermore, the IAT measure is derived from mean performance in each condition (in which the first trial's performance is diluted by the remaining trials), and practice trials precede each condition performed. Thus, the IAT mainly represents the efficiency of executing the mental set while it is maintained (set operation) rather than the effort associated with set maintenance.

As a result, we note that previous research using the IAT (Gemar et al., 2001; De Raedt et al., 2006), evaluated what we describe as set operation but not set maintenance. In other words, though previous research showed that dysphorics and nondysphorics demonstrate a similar positive self bias in measures that evaluate set operation, it is currently unknown whether dysphoria is associated with maladaptive positive/negative set maintenance. We argue that dysphorics' negative bias would be evident in set maintenance and not in smoothness of operation. This hypothesis is buttressed by two lines of evidence. One line shows that dysphorics are differentiated from nondysphorics in cognitive tasks that involve effortful but not "simple" processing (see Hartlage, Alloy, Va'zquez, & Dykman, 1993, for review). The cognitive literature regards processes as "effortful" when they engage central control and "simple" when they require fewer attentional resources (Shiffrin & Schneider, 1977). Mental set maintenance is widely regarded as a paradigmatic case of executive processing whereas set operation is not (Logan, 2003; Monsell, 2003). The second line of evidence concerns differences between dysphorics and nondysphorics in negative thoughts that become accessible in mildly depressed mood (e.g., Ingram, Miranda, & Segal, 1998; Persons & Miranda, 1992; Teasdale, 1983, 1988). These theories hold that whereas both dysphorics and nondysphorics may share quite similar negative thoughts, only dysphorics easily focus their attention on them. Accordingly, as was described above, the IAT task examines the relative strength of implicit associations between certain contents (e.g., how strong are negative attributes associated with the self, see De Houwer, 2002) and not attentional biases toward these associations (e.g., focusing on negative attributes). Consequently, if dysphoria is related to attentional bias, it is more likely to be reflected in set maintenance (which reflects the relative ease of attending to something) and not in set operation (which reflects implicit content).

The Present Study

To test this prediction, we designed a task which allows a distinction between set maintenance and set operation. Specifically, we developed a new paradigm involving frequent task switching between a self-reference based IAT task and a closely matched neutral task. A

further novel feature of our design is that we can separately assess which mental set is difficult/easy for dysphorics and nondysphorics to maintain. The addition of a neutral task, closely matched to the self-reference based IAT in terms of its nonself related processes (including perceptual and response-related processing demands), constitutes a baseline for the IAT maintenance effects. Specifically, in some blocks participants switched between a self = positive mental set (the IAT condition where self is paired with positive attributes) and a neutral mental set. In other blocks participants alternated between a self = negative mental set and a neutral mental set. According to the aforementioned logic, this made it possible to assess the difficulty of maintaining the positive and the negative sets separately. This difficulty is indexed by the response slowing observed in the first trial of the neutral task relative to the remaining trials in that task. To avoid confusion, we label the first trial slowing observed in the neutral task "IAT maintenance cost" and the first trial slowing observed in the self-relevant task "Neutral maintenance cost." The estimates that we used for set operation were based on the level of performance observed when participants execute a single task (i.e., "traditional" IAT) and not when they switch between tasks (i.e., switching between a self and a neutral task in the same block).[1]

Therefore, in the present study, differences between positive and negative biases in the self concepts of dysphoric and nondysphoric individuals were evaluated separately. Simply put, differences between groups could be evident in negative self reference biases or in positive self reference biases. If the differences are in negative self reference biases, they could be due to *nondysphorics* finding it *difficult* to maintain a negative attitude toward the self, or to *dysphorics* finding it *easy* to maintain a negative stance. If the differences are in positive self reference biases, they could exhibit by *nondysphorics* finding it *easy* to maintain a positive stance, or by *dysphorics* finding it *difficult* to maintain a positive stance.

Thus, there are two general hypotheses, each subdivided into two specific hypotheses.

Hypothesis 1: The group differences are in negative self reference biases.

Hypothesis 1a: Nondysphoric individuals will find it relatively difficult to maintain a self = negative mental set. This will be reflected in an increased IAT maintenance cost relative to the neutral task maintenance cost ("Impaired Negativity Hypothesis").

Hypothesis 1b: Dysphoric will maintain a self = negative mental set more easily than the neutral set. This will be reflected in a reduced IAT maintenance cost relative to the neutral task maintenance cost ("Enhanced Negativity Hypothesis").

Hypothesis 2: The group differences are in positive self reference biases.

Hypothesis 2a: Nondysphoric individuals will find it relatively easy to maintain a self = positive mental set. This will be reflected in a reduced IAT maintenance cost relative to the neutral task maintenance cost ("Enhanced Positivity Hypothesis").

Hypothesis 2b: Dysphorics will exhibit a greater difficulty in maintaining a self = positive mental set compared to the neutral set ("Impaired Positivity Hypothesis"). This will be reflected in a larger IAT maintenance cost in the self = positive condition relative to the neutral task maintenance cost.

In addition to examining these hypotheses, we also expected to replicate Gemar et al. (2001) and De Raedt et al. (2006) findings in the IAT, showing that dysphoric and nondysphoric individuals alike would exhibit faster responses in the self = positive condition as compared to the self = negative condition (when performing this task solely), thus demonstrating similarly smooth mental set operation in these two groups. Furthermore, we also sought to show that the neutral task maintenance score was indeed neutral by our definitions. This was important since we used the neutral maintenance score as a baseline reference point to which we compared the negative and positive maintenance costs. By showing that this measure would not be affected by the other experimental variables, we could establish the neutrality of this measure in the present study.

Method

Participants and Procedure

Sixty-three undergraduate students received either course credit or monetary compensation (the Israeli currency equivalent of $6 US) for their participation. All participants were administered the IAT-Task switching paradigm (henceforth IAT-TS) followed by the administration of the self report depression scale (CES-D, Radloff, 1977). We used 16 as a cutoff to identify dysphorics, as previous research suggested it correctly classifies over 90% of depressed patients (see Joormann & Siemer, 2004; Windle, 1992). This resulted in a group of 33 dysphoric and 30 nondysphoric individuals (see Table 1 for group characteristics).

Measures and Tasks

Experimental task. Two tasks form the IAT-TS paradigm. The self task involved a new variant of the IAT (Wigboldus, 2001; see also Karpinski & Steinman, 2006, for a discussion on the advantage of this measure over the classic IAT in evaluating implicit self esteem). It included one target concept (self) and two attributes (positive and negative attributes). The self target concept was composed of four stimuli: participant's first name, last name, and the Hebrew words for "I" and "mine." The positive and negative attributes consisted of four participant-generated attributes describing positive or negative character traits.

The neutral task was closely matched to the self task. It also included one target concept (shape words) and two attributes (dark and light color words). The shape target concept included four stimuli (two shapes that the participants chose, and the Hebrew words for "ellipse" and "rhombus"). The dark and light

---

[1] Apart from the fact that using the IAT as a measure for set operation provides a replication to previous studies, in the task switching literature it was shown that nonswitch trials in mixed blocks (i.e., blocks that involve switching between two tasks) are not equivalent to single task performance (i.e., only one condition is performed at a time, like the IAT). In fact "mixing cost" is an effect commonly found in which non switch trials in mixed blocks are slower than single task performance (see Rubin & Meiran, 2005, for a review). This effect denotes the ongoing state of readiness which is important for mixed blocks but not for single task performance.

Table 1
Group Characteristics

Variable	Group	
	Dysphoric	Non-dysphoric
Gender (female/male)	28/33	17/30
CESD*	24.75	9.20

* $p < .00001$.

color stimuli consisted of four words chosen by each participant. The target concept "shape" was mapped with either the dark or light color words in separate blocks of trials.

Note that the neutral task was matched to the self task in: (a) the number of stimuli, (b) number of participant-generated stimuli, (c) usage of same two response keys, (d) the semantic relations between attributes (dark and light color attributes being two extreme points in a hue axis; positive and negative attributes being two extreme points in a valence axis), and (e) alternating mapping between target and attributes ("shape" words were mapped to "dark" color words or "light" color words in different blocks; "self" words were mapped to "positive" or "negative" words in different blocks).

In an effort to simplify the complicated design of the experimental blocks for the reader, the description of this design is divided into four hierarchical levels. These levels include: the general design, block, run, and trial levels (see also Figure 1).

General design and block levels. The general design included two parts: the IAT-TS, followed by a "traditional" IAT. The
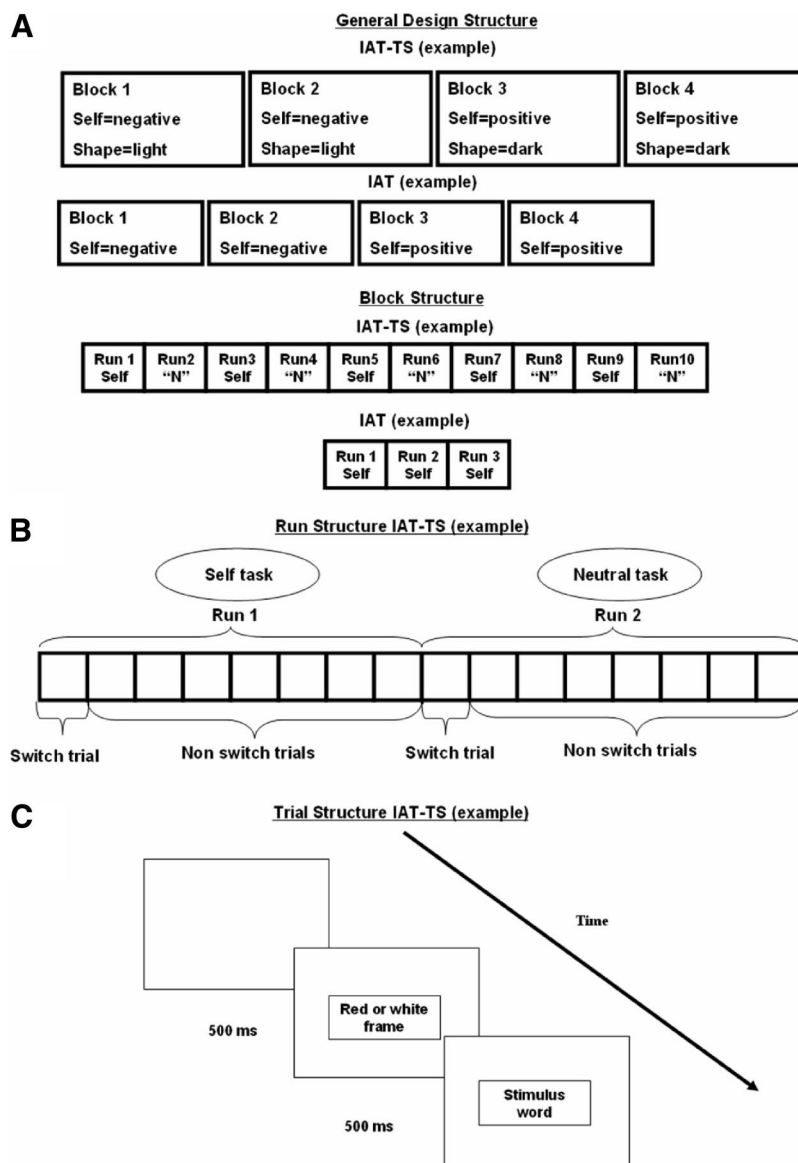

Figure 1. (a) general design and block structure of the experimental task.; (b) two runs in the IAT-TS part; (c) trial sequence IAT-TS.

IAT-TS consisted of four mixed blocks of trials in which participants alternated between the self task, and the neutral task. In each of the four blocks only one condition of the self mapping was performed (i.e., either the self = positive mapping or self = negative mapping). To compare between tasks, the neutral task also had two mappings (i.e., shape = dark colors or shape = light colors) each performed in separate blocks. Therefore, the total of four blocks included performing two blocks for each condition. In the IAT-TS part, a 16 trial practice block for each task was administered before the first experimental block, and every time the condition changed (i.e., self = negative changing to self = positive). Order of key mappings in both tasks, and order of tasks were counterbalanced across participants. In addition a short break was given to participants after each block.

After performing the IAT-TS part, participants completed the "traditional" IAT part which included four single IAT blocks. This part was included in order to evaluate the set operation effects when performing the IAT task only (see Figure 1a).

Run level. In the IAT-TS, participants alternated between runs of 8 trials involving the self-reference based IAT and the neutral task, respectively, using the same two keys in both tasks. Notably, *switch trials* were the first trials in each run, whereas *nonswitch* trials were the remaining trials in the run (see Figure 1b).

Trial level. The onset of each trial was signaled by a 500-ms (ms) of a blank screen, followed by a 500-ms presentation of the task cue. This cue was a white or red frame, signaling the self and neutral tasks, respectively. This frame was aligned with the center of the screen. The stimulus word appeared inside the frame, and the entire display remained visible until the response was given or after 4 sec. elapsed (see Figure 1c).

Self report depression scale (CES-D, Radloff, 1977). Respondents rated the level to which they experienced 20 symptoms over the past 2 weeks (e.g., "I am sad"). Items were scored on a 4-point scale, 0 (*rarely*) to 3 (*most of the time*). Participants' scores were computed as the sum of their responses to the items. The possible score range is 0–60.

Results

Mental Set Maintenance Effects

To test our four hypotheses we conducted a $2 \times (2 \times 2 \times 2 \times 2)$ Analysis of Variance with Dysphoria (nondysphoric vs. dysphorics) serving as a between-subjects variable, and Mapping (self = positive vs. self = negative), Task (self vs. neutral), Switch (switch - performance in the first trial in the run vs. nonswitch - the remaining trials) and Stimulus-type[2] (target vs. attribute) served as within-subject variables. In addition to reporting conventional *p* values, we report Prep-values (Killeen, 2005), which is the post hoc probability of replicating a non-null effect.

Focusing on the proportion of errors as an outcome, we found only a significant Switch effect, $F(1, 61) = 21.82, p < .0001$, partial $\eta^2 = .26$, Prep > .99, indicating that participants made more errors on switch trials than on nonswitch trials. Since there was no evidence for speed–accuracy trade-off, the remaining analyses focus on RT.

In the present design there were Mapping, Switch and Task main effects (all F's > 14.04, p's < .001), Switch by Dysphoria, Mapping by Switch, and Task by Switch 2-way interactions (all F's > 9.54, p's < .01), and Task by Switch by Dysphoria, and Mapping by Task by Switch 3-way interactions (all F's > 5.87, p's < .02). However, these effects were qualified by the expected 4-way interaction between Dysphoria, Task, Switch, and Mapping $F(1, 59) = 4.89, p < .05$, partial $\eta^2 = .08$, Prep = .91. In line with the four hypotheses, mean RTs for correct responses for each condition are presented in Figure 2a and mean maintenance costs are depicted in Figure 2b. Probing this interaction was done by employing theoretically meaningful contrasts according to the four aforementioned hypotheses. The difference between groups was obtained for the self = negative condition. Specifically, the pattern of this interaction was found to be consistent with Hypothesis 1a, which posits that nondysphoric individuals have a difficulty in maintaining a mental set involving negative self reference. The simple-simple, Task by Switch, interaction at IAT Mapping self = negative was significant for nondysphoric individuals, $F(1, 59) = 30.03, p < .00001$, partial $\eta^2 = .34$, Prep > .99. This result indicates that nondysphoric individuals show an increased IAT maintenance cost relative to the neutral task maintenance cost, when the IAT mapping is self = negative. Importantly, this pattern was significantly different between nondysphoric and dysphoric individuals $F(1, 59) = 12.42, p < .001$, partial $\eta^2 = .17$, Prep = .99, with the dysphoric group showing similar maintenance costs for the IAT and neutral tasks when self = negative $F(1, 59) < 1$.[3]

Similar contrasts testing the three remaining hypotheses were not significant, Hypothesis 1b: $F(1,59) < 1$; Hypothesis 2a: $F(1, 59) = 3.05, p = .09$, with a trend in the hypothesis disconfirming direction; Hypothesis 2b: $F(1, 59) < 1$.

Validating the Neutrality of the Neutral Task Maintenance Score

The performance of the neutral task was not affected by the other experimental variables. As predicted, the neutral task maintenance score was not affected by Dysphoria $F(1, 59) < 1$; IAT Mapping $F(1, 59) = 2.78, p = .10$, nor was it affected by their interaction $F(1, 59) < 1$. These results constitute an important validation of this measure as being neutral with respect to our variables of interest.

---

[2] The results we report are not qualified by any interactions involving Stimulus-Type variable (target—self and shape stimuli which had one key mapping throughout the experiment vs. attribute—positive, negative, dark and light colors stimuli which had 2 keyboard mappings that changed between blocks). Therefore this variable is not discussed any further.

[3] Due to a different male/female ratio in the dysphoric and nondysphoric groups, we wanted to check whether the results obtained concerning hypothesis 1 were affected by gender. Therefore, we conducted a separate ANOVA in which Gender (male, female) was added as a between participant variable. A marginal 5-way interaction involving Gender, Dysphoria, Mapping, Task and Switch was found, $F(1, 57) = 3.64, p = .06$. Nonetheless, this marginal interaction does not compromise our conclusions regarding the difference between groups in negative self reference. Specifically, the contrast corresponding to Hypothesis 1a was significant both for male nondysphoric participants $F(1, 57) = 8.95, p < .001$, and female nondysphoric participants $F(1, 57) = 21.09, p < .00001$, and the effect was statistically comparable in size for both genders, $F(1, 57) < 1$. Complementary to this result, the nonsignificant trend that we have found for the dysphoric group was evident in a non significant result both for male dysphoric participants $F(1, 57) < 1.02$ and female dysphoric participants $F(1, 57) < 1$.
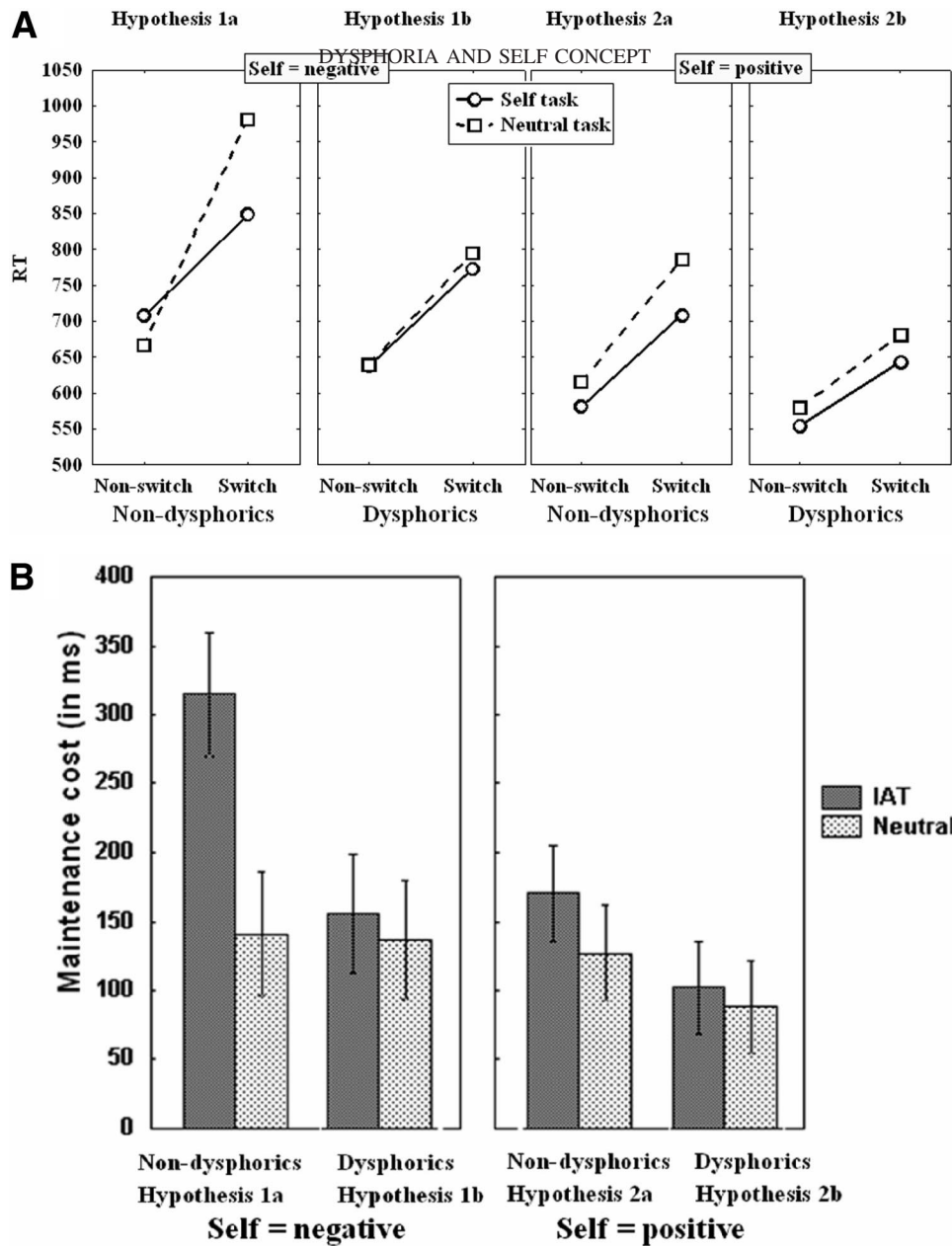
Figure 2. (a) RT According to Dysphoria, Mapping, Task and Switch; (b) Maintenance Costs According to Dysphoria, Task and Mapping (see text for details). Bars represent Confidence Intervals.

Mental Set Operation Effects

We conducted a two-way mixed model ANOVA on participants' mean correct RTs when performing the self task only, with Mapping (self = positive, self = negative) as a within-participants independent variable, and Dysphoria (nondysphoric, dysphoric) as a between-participants independent variable. As predicted, only a significant main effect for Mapping was found, $F(1, 61) = 39.31$, $p < .0001$, partial $\eta^2 = .39$, *Prep* $> .99$ (all other F's <1.67 *ns*) showing that RTs were faster when self = positive than when self = negative. The two groups showed a similar bias, as indicated by the nonsignificant interaction, $F(1, 61) < 1$ between Dysphoria and IAT Mapping, replicating results from former studies (De Raedt et al., 2006; Gemar et al., 2001). The IAT mapping effect was 62 and 76 ms for dysphoric and nondysphoric individ-

uals, respectively. This is consistent with our expectation that dysphorics and nondysphorics will not differ in their set operation characteristics in the IAT task.

Discussion

In this study, we addressed a puzzling inconsistency between a negative bias in the self concept of depressed individuals found using explicit measures, and an opposite, positive bias found using implicit measures. We adopted the distinction between set maintenance (how difficult it is to execute a task) and set operation (how successful is task performance once the task is being executed). We reasoned that the IAT effect taps only set operation. Moreover, we hypothesized that differences between dysphorics and nondysphorics would be found in attentional biases, and

should be reflected in set maintenance and not set operation in the case of the IAT. Applying a task switching approach, we disentangled effortful mental set maintenance from set operation. The principal novel finding in this study was that differences between dysphorics and nondysphorics were indeed limited to set maintenance indices and were not found for set operation indices. Specifically, the results were that the groups differed in their negative self reference biases so that nondysphorics exhibited a difficulty in maintaining active an (implicit) negative self-reference mental set, whereas dysphorics did not. In other words, our indices show that nondysphorics experienced a difficulty forcing themselves to execute the task in which their selves was implicitly associated with negative attributes (Hypothesis 1a).

However, we did not find that for dysphorics it was easy to maintain a negative self reference mental set (Hypothesis 1b). Furthermore, no differences between groups appeared in positive self reference: nondysphoric individuals did not exhibit an exceptional ease in maintaining a positive self referent mental set (Hypothesis 2a) and dysphoric individuals did not have an exceptional difficulty to maintain a positive self reference mental set (Hypothesis 2b), Our findings supporting the impaired negativity hypothesis in nondysphorics is especially important, as it stands against the background of a general positive bias exhibited in the IAT, an effect replicated in our findings.

The distinction between effortful set maintenance and simple set operation seems central in unveiling differences in the implicit self concept of dysphoric and nondysphoric individuals. According to the task set inertia rationale (Allport et al., 1994), switching from the self reference task to the neutral task denotes the set maintenance difficulty of the self reference task. Consequently, for nondysphorics, the negative self reference mental set is difficult to maintain, and hence, requires a heightened activation. This heightened state of activation persists into the next (neutral) task and results in large switching cost. By contrast, for dysphorics the negative self reference mental set is easily adopted and does not require a heightened state of activation. This leads to a reduced maintenance cost when switching to the neutral mental set. This finding implies that dysphorics more easily and perhaps habitually activate a negative self reference relative to nondysphorics.

Our results may seem counterintuitive since they suggest that dysphorics do not exhibit an impairment in their attentional biases. However, several other studies showed that dysphoric individuals do not demonstrate impaired performance on tasks in which participants are asked explicitly to attend to emotional aspects of the stimuli (Matt, Vazquez, & Campbell, 1992; Siegle, Ingram, & Matt, 2002). Note that our results may be interpreted as suggesting that dysphorics have better executive control (smaller maintenance costs) relative to nondysphorics. This conclusion is unwarranted, however, because the exaggerated maintenance costs among nondysphorics were found only when the self was mapped to negative attributes, while the groups did not differ from one another in this respect when the self was associated with positive attributes. This latter finding is in line with Roberts, Gotlib, and Gilboa (1998) who proposed that positive and negative self-schema may be distinct constructs, and that dysphoria is related to the greater accessibility of negative information, but does not erode a positive self-schema. In line with this view, Joormann (2004) showed that dysphoric individuals did not differ from nondysphorics in their ability to inhibit positive distractors.

In showing that dysphoria is associated with set maintenance and not set operation, our results comply with a gradual shift in recent years in the understanding of cognitive vulnerability to depression. Whereas the traditional account viewed vulnerability to depression as depending on an underlying dysfunctional attitudes or contents (e.g., Beck, Epstein, & Harrison, 1983), a newer view holds that vulnerability to depression and relapse depends on the enhanced accessibility to negative thoughts during mildly depressed mood (e.g., Ingram, Miranda, & Segal, 1998; Persons & Miranda, 1992; Teasdale, 1983, 1988). This theoretical shift is also evident in corresponding changes in the cognitive treatment of depression and the prevention of relapse introducing mindfulness technique (see Bishop et al., 2004, for a definition). Unlike the more conventional cognitive–behavioral therapy, which focuses on changing contents or specific meanings of negative automatic thoughts, mindfulness techniques involve deactivation of the negative schemata via the improvement of attentional control and the change in attentional habits and preferences (e.g., Teasdale, Segal, Williams, Ridgeway, Soulsby, & Lau, 2000).

The extant literature suggests that dysphoria and depression are characterized by an attentional bias to negative information (e.g., Bradley, Mogg, & Lee, 1997; Gotlib, Krasnoperova, Yue, & Joormann, 2004; Joormann & Gotlib, 2007; Koster, De Raedt, Goeleven, Franck, & Crombez, 2005; Mogg, Bradley, & Williams, 1995). Notice that most of these studies have used some form of the dot probe task which measures spatial selective attention to general negative and positive contents. Our results contribute to these studies in showing direct effects of attentional control within the self concept of dysphoric individuals.

The design of the IAT-TS paradigm raises two major questions of theoretical relevance. First, why did we choose a neutral task as an alternating task to the self task rather than create a paradigm which includes alternating between self = positive and self = negative conditions? The use of a neutral task had two significant advantages. One is that it enabled us to validate (within our paradigm) the task set inertia account. Alternating between two tasks that are conceptually independent (self task and a neutral task) enabled us to conclude that the maintenance costs observed in the neutral task derive exclusively from difficulties in maintaining the self task that preceded it as the task set inertia hypothesis predicts (see especially Yeung et al., 2006). Had we alternated between two self tasks (i.e., switching between self = positive and self = negative), which are by definition conceptually dependent, such a conclusion could not have been unequivocally reached. The second advantage of using a neutral task is in providing a baseline to the positive and negative conditions. This enabled us to expose the origin of the effect as was formalized in our study in deciding between Hypotheses 1a and 1b and between Hypotheses 2a and 2b.

A second question is why we found maintenance costs at all? This question seems relevant as the task switching literature is divided on the conditions needed to obtain maintenance/switching costs. According to one view, switching costs are found only when both tasks use the same stimuli (Allport & Wylie, 2000; Wylie, Javitt, & Foxe, 2004). For example using the same set of stimuli (i.e., 1–9 numbers) switching costs will be discovered when alternating between two tasks: (a) an odd/even task (i.e., 7 is odd) and (b) larger/smaller than 5 task (i.e., 7 is larger than 5). According to this view, we should not have observed any switching effects. However, the present results show that stimulus overlap is not a

necessary condition to observe switch costs contrary to Wylie et al.'s predictions. In fact, one of us (Meiran, 2000b, Experiment 1) showed that the overlap in *responses* is a contributing factor to switch costs, and in the same study (Experiment 2) showed that switch costs are found with tasks not sharing the stimulus set but sharing the responses. According to Meiran's (2000a) model the reason why switch costs are found is related to the need to change the meaning associated with each key press in the two tasks (see also Brass et al., 2003; Koch et al., 2003; Philipp et al., in press; Schuch & Koch, 2003). In the present paradigm, the participants used the same set of response keys in both tasks. In fact, the critical IAT-TS manipulation referred to the overlap of response meanings. Therefore, when performing the self task following the neutral task one has to change the meaning from right key = dark color to right key = negative attributes. The fact that switch costs were found provides another demonstration that the sharing of stimuli is not necessary for such costs to be found. Furthermore, Schuch and Koch (2003) have shown that task set inhibition (Mayr & Keele, 2000), which is an essential component of task set maintenance according to current theorizing, is related to the level of response-interpretation representation. Meiran, Levine, Meiran, and Henik (2000, Experiment 3, which was conducted on normal young adults) have additionally shown that the dissipation of the previous task set (which is the cause of TSI) is tightly linked to response overlap.

Several limitations of the present research should be noted. First, the IAT-TS paradigm is complex and takes more than 40 min to complete. We are currently in the process of developing simpler and shorter tasks. Nevertheless, it seems unlikely that in this study dysphoric's performance was compromised due to task complexity, since dysphoric individuals showed smaller maintenance costs than nondysphoric ones. Second, in the set operation analysis of the IAT we did not use the improved scoring algoritm suggested by Greenwald, Nosek, and Banaji (2003). This algorithm provides more statistical power and has greater internal and external validity than our approach. However, in order to the use this algorithm several design modifications were needed, which could have affected differently the dysphoric and nondysphoric groups. Note that in our design the IAT was administered following the IAT-TS part. Since this task was already long and hence could possibly affect performance of the dysphoric group, we shortened the practice and test phases of this task relative to Greenwald et al. (2003). Furthermore, the scoring algorithm includes providing an error feedback and penalty. As has recently been shown (Holmes & Pizzagalli, 2007) dysphorics's performance is adversely affected by negative performance feedback. Last, notwithstanding differences in mental set maintenance between dysphorics and nondysphorics, it is still unclear whether these differences stem from a transient mood state as opposed to enduring vulnerability.

How then can our results resolve the paradox described above? Our findings seem to agree with a recent paper which integrates the role of implicit and explicit measures in predicting vulnerability to depression (Haeffel, Abramson, Brazy, Shah, Teachman, & Nosek, 2007; see also Beevers, 2005). Relying on the dual process theory (Chaiken & Trope, 1999), the authors view implicit processes as occurring without effort and not taxing cognitive resources and explicit processes as effortful and resource demanding. Accordingly, we have argued that set maintenance is a paradigmatic case involving effortful executive control whereas

set operation (the standard IAT) is not. Therefore, it seems reasonable that our set maintenance results and not set operation results converge with explicit measures to show dysphorics's (relative to nondysphorics) increased attentional focus on negative reference to the self.

To conclude, our findings suggest that nondysphorics find it more difficult to maintain a negative-self mental set than a neutral set, whereas dysphorics find the two mental sets equally difficult to maintain. These findings highlight two important distinctions: (a) between set maintenance and set operation, and (b) between positive and negative aspects of the self concepts in the context of dysphoria. It appears that the shifts in and out of moods is related not only to the accumulation of previous positive and negative experiences but also to the ability to attend to certain self aspects so as to elaborate, differentiate, and integrate these experiences.

References

Abramson, L. Y., Metalsky, G. I., & Alloy, L. B. (1989). Hopelessness depression: A theory-based subtype of depression. *Psychological Review, 96,* 358–372.

Abramson, L. Y., Seligman, M. E. P., & Teasdale, J. (1978). Learned helplessness in humans: Critique and reformulation. *Journal of Abnormal Psychology, 87,* 49–74.

Allport, A., Styles, E. A., & Hsieh, S. (1994). Shifting intentional set: Exploring the dynamic control of tasks. In C. Umila & M. Moscovitch (Eds.), *Attention and performance XV* (pp. 421–452). Hillsdale, NJ: Erlbaum.

Allport, A., & Wylie, G. (2000). 'Task-switching', stimulus-response bindings and negative priming. In S. Monsell & J. Driver (Eds.), *Attention and performance XVIII*: *Control of cognitive processes* (pp. 35–70). Cambridge, MA: MIT Press.

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., text revision). Washington, DC: Author.

Beck, A. T. (1976). *Cognitive therapy and the emotional disorders.* New York, NY: International Universities Press.

Beck, A. T., Epstein, N., & Harrison, R. (1983). Cognitions, attitudes and personality dimensions in depression. *British Journal of Cognitive Psychotherapy, 1,* 1–16.

Beck, A. T., Rush, A. J., Shaw, B. F., & Emery, G. (1979). *Cognitive therapy of depression.* New York, NY: Guilford Press.

Beevers, C. G. (2005). Cognitive vulnerability to depression: A dual process model. *Clinical Psychology Review, 25,* 975–1002.

Bishop, S. R., Lau, M., Shapiro, S., Carlson, L., Anderson, N. D., Carmody, J., et al. (2004). Mindfulness: A proposed operational definition. *Clinical Psychology: Science and Practice, 11,* 230–241.

Bradley, B. P., Mogg, K., & Lee, S. C. (1997). Attentional biases for negative information in induced and naturally occurring dysphoria. *Behaviour Research and Therapy, 35,* 911–927.

Brass, M., Ruge, H., Meiran, N., Rubin, O., Koch, I., Zysset, S., et al. (2003). When the same response has different meaning: Recoding the response meaning in the lateral prefrontal cortex. *NeuroImage, 20,* 1026–1031.

Chaiken, S., & Trope, Y. (1999). *Dual-process theories in social psychology.* New York, NY: Guilford Press.

De Houwer, J. (2002). The implicit association test as a tool for studying dysfunctional associations in psychopathology: Strengths and limitations. *Journal of Behavior Therapy and Experimental Psychiatry, 33,* 115–133.

De Raedt, R., Schacht, R., Franck, E., & De Hower, J. (2006). Self-esteem and depression revisited: Implicit positive self-esteem in depressed patients? *Behaviour Research and Therapy, 44,* 1017–1028.

Gemar, M. C., Segal, Z. V., Sagrati, S., & Kennedy, S. J. (2001). Mood-

induced changes on the implicit association test in recovered depressed patients. *Journal of Abnormal Psychology, 110,* 282–289.

Gotlib, I. H., Krasnoperova, E., Yue, D. N., & Joormann, J. (2004). Attentional biases for negative interpersonal stimuli in clinical depression. *Journal of Abnormal Psychology, 113,* 127–135.

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.

Haeffel, G. J., Abramson, L. Y., Brazy, P. C., Shah, J. Y., Teachman, B. A., & Nosek, B. A. (2007). Explicit and implicit cognition: A preliminary test of a dual-process theory of cognitive vulnerability to depression. *Behaviour Research and Therapy, 45,* 1155–1167.

Hartlage, S., Alloy, L. B., V'azquez, C., & Dykman, B. (1993). Automatic and effortful processing in depression. *Psychological Bulletin, 113,* 247–278.

Holmes, A. J., & Pizzagalli, D. A. (2007). Task feedback effects on conflict monitoring and executive control: Relationship to sub-clinical measures of depression. *Emotion, 7,* 68–76.

Ingram, R. E., Miranda, J., & Segal, Z. V. (1988). *Cognitive vulnerability to depression.* New York, NY: Guilford Press.

Joormann, J. (2004). Selective attention in dysphoria: The role of inhibitory processes. *Cognition and Emotion, 18,* 125–147.

Joormann, J., & Gotlib, I. H. (2007). Selective attention to emotional faces following recovery from depression. *Journal of Abnormal Psychology, 116,* 80–85.

Joormann, J., & Siemer, M. (2004). Memory accessibility, mood regulation, and dysphoria: Difficulties in repairing sad mood with happy memories? *Journal of Abnormal Psychology, 113,* 179–188.

Karpinski, A., & Steinman, R. B. (2006). The single category implicit association test as a measure of implicit social cognition. *Journal of Personality and Social Psychology, 91,* 16–32.

Killeen, P. R. (2005). An alternative to null-hypothesis significance tests. *Psychological Science, 16,* 345–353.

Klauer, K. C., & Mierke, J. (2005). Task-set inertia, attitude accessibility, and compatibility-order effects: New evidence for a task-set switching account of the implicit association effect. *Personality and Social Psychology Bulletin, 31,* 208–217.

Koch, I., Ruge, H., Brass, M., Rubin, O., Meiran, N., & Prinz, W. (2003). Equivalence of cognitive processes in brain-imaging and behavioral studies: Evidence from task switching. *NeuroImage, 20,* 572–577.

Koster, E. H. W., De Raedt, R., Goeleven, E., Franck, E., & Crombez, G. (2005). Mood-congruent attentional bias in dysphoria: Maintained attention to and impaired disengagement from negative information. *Emotion, 5,* 446–455.

Logan, G. D. (2003). Executive control of thought and action: In search of the wild homunculus. *Current Directions in Psychological Science, 12,* 45–48.

Matt, G. E., Vazquez, C., & Campbell, W. K. (1992). Mood-congruent recall of affectively toned stimuli: A meta-analytic review. *Clinical Psychology Review, 12,* 227–255.

Mayr, U., & Keele, S. W. (2000). Changing internal constraints on action: The role of backward inhibition. *Journal of Experimental Psychology: General, 129,* 4–26.

Meiran, N. (2000a). Modeling cognitive control in task-switching. *Psychological Research, 63,* 234–249.

Meiran, N. (2000b). The reconfiguration of the stimulus task-set and the response task set during task switching. In S. Monsell & J. Driver (Eds.), *Attention and performance XVIII: Control of cognitive processes* (pp. 377–400). Cambridge, MA: MIT Press.

Meiran, N., Levine, J., Meiran, N., & Henik, A. (2000). Task-set switching in schizophrenia. *Neuropsychology, 14,* 471–482.

Mogg, K., Bradley, B. P., & Williams, R. (1995). Attentional bias in anxiety and depression: The role of awareness. *British Journal of Clinical Psychology, 34,* 17–36.

Monsell, S. (2003). Task switching. *Trends on Cognitive Sciences, 7,* 134–140.

Persons, J. B., & Miranda, J. (1992). Cognitive theories of vulnerability to depression: Reconciling negative evidence. *Cognitive Therapy and Research, 16,* 485–502.

Philipp, A. M., Jolicoeur, P., Falkenstein, M., & Koch, I. (2007). Response selection and response execution in task switching: Evidence from a go-signal paradigm. *Journal of Experimental Psychology: Human Perception and Performance, 33,* 1062–1075.

Radloff, L. (1977). The CES-D scale: A self-report depression scale for research in the general population. *Applied Psychosocial Measurement, 1,* 385–401.

Roberts, J. E., Gilboa, E., & Gotlib, I. H. (1998). Ruminative response style and vulnerability to episodes of dysphoria: gender, neuroticism, and episode duration. *Cognitive Therapy and Research, 22,* 401–423.

Rubin, O., & Meiran, N. (2005). On the origins of the task mixing cost in the cuing task switching paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31,* 1477–1491.

Rubinstein, J. S., Meyer, D. E., & Evans, J. E. (2001). Executive control of cognitive processes in task switching. *Journal of Experimental Psychology: Human Perception and Performance, 27,* 763–797.

Schuch, S., & Koch, I. (2003). The role of response selection for inhibition of task sets in task shifting. *Journal of Experimental Psychology: Human Perception and Performance, 29,* 92–105.

Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing: N. Perceptual learning, automatic attending, and a general theory. *Psychological Review, 84,* 127–190.

Siegle, G. J., Ingram, R. E., & Matt, G. E. (2002). Affective interference: An explanation for negative attentional bias in dysphoria? *Cognitive Therapy and Research, 26,* 73–87.

Teasdale, J. D. (1983). Negative thinking in depression: Cause, effect or reciprocal relationship? *Advances in Behaviour Research and Therapy, 5,* 3–25.

Teasdale, J. D. (1988). Cognitive vulnerability to persistent depression. *Cognition and Emotion, 2,* 247–274.

Teasdale, J. D., Segal, Z. V., Williams, J. M. G., Ridgeway, V. A., Soulsby, J. M., & Lau, M. A. (2000). Prevention of relapse/recurrence in major depression by mindfulness-based cognitive therapy. *Journal of Consulting and Clinical Psychology, 68,* 615–623.

Wigboldus, D. (2001). The difference between relative and absolute implicit associations. Paper presented at the third meeting of the European Social Cognition Network, Houffalize, Belgium.

Windle, M. (1992). A longitudinal study of stress-buffering for adolescent problem behavior. *Developmental Psychology, 28,* 522–530.

Wylie, G., Javitt, D. C., & Foxe, J. (2004). The role of response requirements in task switching: Dissolving the residue. *NeuroImage, 15,* 1079–1087.

Yeung, N., Nystrom, L. E., Aronson, J. A., & Cohen, J. D. (2006). Between-task competition and cognitive control in task switching. *The Journal of Neuroscience, 26,* 1429–1438.