

## On the Validity of Idiographic and Generic Self-Concept Implicit Association Tests: A Core-Concept Model

MATTHIAS BLUEMKE<sup>1\*</sup> and MALTE FRIESE<sup>2</sup>

<sup>1</sup>Psychological Institute, University of Heidelberg, Heidelberg, Germany

<sup>2</sup>Department of Psychology, University of Basel, Basel, Switzerland

**Abstract:** We present a core-concept model (CCM) suggesting that stimulus centrality is an important factor in category representations in implicit measures. We tested the hypothesis that idiographic stimuli (first name, birthday) are more central and therefore assess self-concept in Implicit Association Tests (IATs) more validly than generic and nonspecific stimuli (me, you). Superior validity of the idiographic variant emerged across three different domains of self-concept. First, an idiographic self-esteem IAT displayed higher correlations than a generic IAT with self-assessments and observer-assessments of self-esteem. Second, an idiographic body scheme-IAT predicted subjective ratings of body image and objective body-mass index. Third, an idiographic aggressiveness-IAT had higher incremental validity for unprovoked aggression when interacting with explicit measures of aggressiveness. We conclude that idiographic stimuli focus participants' attention on the core features of the self, hence, tapping into self-related associations to a stronger degree than generic stimuli. Copyright © 2011 John Wiley & Sons, Ltd.

**Key words:** implicit association; self-concept; centrality; idiographic; nomothetic

Would you pay rather attention when someone called 'hey, you' or when someone called your name, after spotting you in the crowd? Apparently, whether we address each other by unspecific or specific labels (such as pronouns or names) can have a tremendous effect on what we focus on, or how easily we pay attention. Can the different ways of addressing the self be exploited to improve on the assessment of self-concept? The *self-concept* refers to a very refined knowledge structure about the person we know best (Symons & Johnson, 1997). It plays a key role in memory when we filter information, construct knowledge, or evaluate objects and people including ourselves (Rogers, 1959; Markus, 1977; Greenwald & Pratkanis, 1984; Mussweiler & Strack, 2000). It is a dynamic associative structure with manifold semantic and evaluative links (Collins & Loftus, 1975; Markus & Wurf, 1987). In the present research, we investigate the hypothesis that the stimulus setup can influence the validity of implicit measures. The Implicit Association Test (IAT; Greenwald, McGhee, & Schwartz, 1998) is one of the most popular implicit measures to target the self-concept (Greenwald & Farnham, 2000). Previous findings have shown that the specific stimulus selection shapes the outcomes of the IAT procedure (e.g. Bluemke & Friese, 2006; Govan & Williams, 2004). Across three psychological domains, we will show that idiographic representations by self-relevant stimuli (e.g. BRAD or BOB) in the IAT are likely to be more sensitive to self-related associations than unspecific, generic pronouns (e.g. ME or YOU).

### THEORETICAL CONSIDERATIONS: A CORE-CONCEPT MODEL

To explain the impact of stimuli on the validity of self-concept implicit measures, we propose a *core-concept model* (CCM) that is based on existing research on semantic network models and includes self-schemas as cognitive-affective structures (Markus, 1977). Object-related associations are typically depicted in social knowledge structures such as Greenwald et al.'s (2002) 'unified theory.' The model incorporates triangular associations between attitudes (object-valence relations), stereotypes (object-attribute relations), self-esteem (self as object-valence relations), and self-concept (self as object-attribute relations). The CCM makes additional qualifications and assumptions about the *centrality* of the stimuli required for the representation of a concept. Centrality has been defined as 'the relative importance that various aspects bear in a person's conceptualization of objects in a given cognitive domain' (Scott, Kline, Faguy-Coté, & Peterson, 1980, p.12).

Obviously, not all aspects are equally central in defining who one is. Trait centrality refers to the idea that some personality traits are closer to the core identity than alternative traits (Sedikides & Skowronski, 1993). For example, age may be more important for some people than for others. Or consider the idea that people's global self-esteem can be understood only to the extent that one acknowledges the esteem attached to self-components, which themselves are regarded by persons as more or less central (Rosenberg, 1979). In its strictest sense, centrality refers to invariant characteristics of the concepts. In the case of self-concept, this is a person's *core*.

Take an exemplary participant (including a social comparison standard) as depicted in Figure 1. Centrality can be

\*Correspondence to: Matthias Bluemke, Psychological Institute, University of Heidelberg, Hauptstrasse 47–51, Heidelberg D-69117, Germany. E-mail: matthias.bluemke@psychologie.uni-heidelberg.de

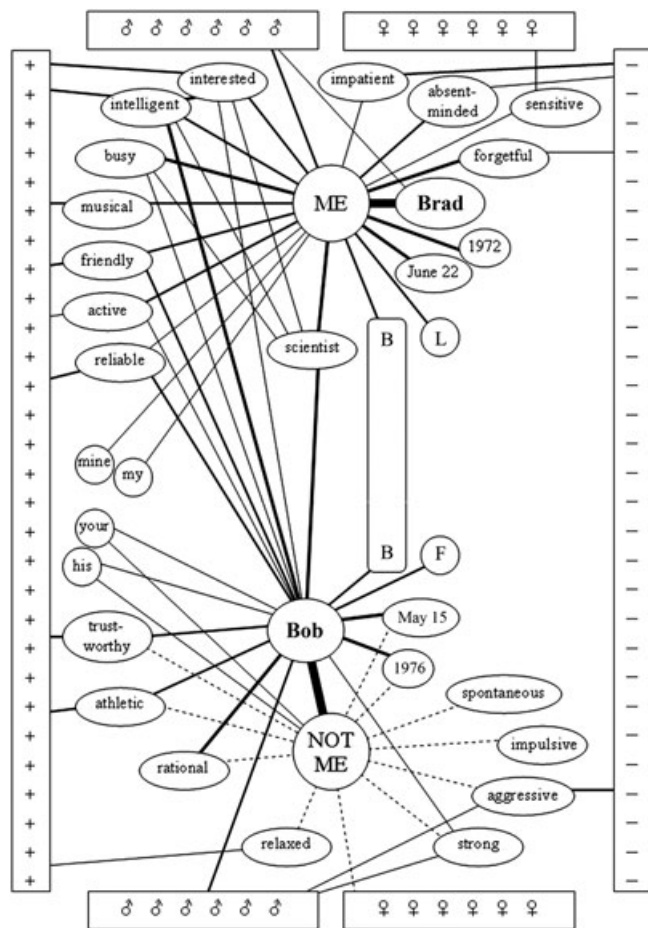


Figure 1. A social knowledge structure (adapted from Greenwald et al., 2002) according to the core-concept model (CCM) depicting the self of an exemplary person including a friend as a comparison standard: Nodes represent semantic concepts, and links represent associations between concepts that can vary in strength (line thickness). In contrast to stereotypes and attitudes, self-concept and self-esteem always involve direct links to ME. Depending on the comparison standard, aspects relating to NOT-ME can be included in the self too, either directly or indirectly (as mediated by links). ME and NOT-ME can overlap in social categories (gender, professional roles), trait attributes (e.g. intelligent), and first name initials (B). Alternatively, ME and NOT-ME can uniquely relate to attributes (e.g. forgetful, athletic), pronouns (my, your), and family name initials (L, F). Aspects central to the self are represented closer to ME, resulting in the self being linked to pronouns, initials, birthdays, and other personal identifiers with increasing centrality.

illustrated by the relative strengths of links, or spatial distances, between hierarchically related concepts in a semantic network model. Some stimuli are closer to the concepts or have stronger ties than others. The more central a stimulus is, the more conceptually relevant the associative echo from the stimulus prompt will be. Stimulus centrality helps a participant focus on the concept in question during a measurement procedure. Consequently, when targeting the self-concept with central stimuli, the mental representation will be centred more strictly on the core self. In the IAT, this will influence the ease with which a stimulus can be validly associated with *me*, thus influencing which aspects during the measurement procedure constitute compatible and incompatible category alignments, ultimately reducing nuisance variance in the measure.

To qualify the CCM more formally, we assume that for accurate inferences about people's associations, the concepts

in implicit measures need to be tapped into *as centrally as possible*. This can be achieved with stimulus selections that are (a) *representative* for and (b) *proximal* to the core concept. The stimulus selection to represent self-concept will determine how centrally represented a person's self-concept is. Importantly, these stimulus selection rules apply to any kind of stimulus type—regardless whether letters, numbers, action verbs, state verbs, adjectives, social categories, nouns, or names are used to address the self. Moreover, we assume that the CCM principles apply regardless whether the self-concept is assessed by category-based measures (that use category labels) such as the IAT or exemplar-based measures (that do not rely on self-related categories) such as the affective priming task (Wentura, Kulfanek, & Greve, 2005) or the name-letter-task (NLT; Nuttin, 1985). Although the CCM is meant to be applicable to a range of measurement situations, we test the basic idea of tapping into core concepts by exploring the centrality of *self-related aspects*. And although the model is applicable to various kinds of implicit measures, in the following experiments, we examine its utility with respect to self-concept IATs.

With regard to the first requirement, *representativeness*, the CCM appreciates that different cues increase the accessibility of different representations. When the concept representation within an implicit measurement procedure is not represented in a balanced manner or otherwise deviates from the concept as it is encountered in a specific social context, the validity of the measure is reduced markedly (Blair, 2002; Barden, Maddux, Petty, & Brewer, 2004; Rydell & Gawronski, 2009). Therefore, the stimulus selection should always be representative of the concept. If the selection targets the core concept, then the stimuli should be superior to stimuli that merely elicit peripheral and fluctuating mental representations. For if peculiarities prevent the mental construal of the intended concept, the measurement outcomes will be less indicative of the conceptual associations. In support of this, implicit–explicit (I–E) as well as implicit–implicit (I–I) correlations of an IAT that employed synonyms of the category labels as stimuli were higher than that of a (typical) stimulus-based IAT that allowed fluctuating mental construal because of peculiar stimulus connotations (Steffens, Kirschbaum, & Glados, 2008). With regard to self-concept measures then, ideally, those stimuli should be chosen that reduce fluctuations and allow access to the *invariant* features representing a person, unless one is explicitly interested in a context-dependent self construal (see McConnell, 2011).

With regard to the second aspect of the CCM, *proximity*, one and the same concept can be represented by stimuli that reside at different locations in the semantic network (Rosch, 1978; Rosch, Mervis, Gray, Johnson, & Boyes-Braem, 1976). Although two stimuli may both have ties with a concept, the semantic distance between these stimuli and the concept may differ. For instance, a researcher may choose EINSTEIN to represent the target concept *math* in an IAT. The semantic links from EINSTEIN to the concept *math* can be construed as such, but they are weak from the perspective of what characterises math at heart. By contrast, ALGEBRA, NUMBER, or EQUATION resonate stronger

with math. Whenever a researcher incorporates stimuli to target at conceptual associations, proximal stimuli that resonate strongly are to be preferred to distal ones that resonate only weakly with the concept. Again, this pertains to the self as well, and this is where the distinction between idiographic and generic stimuli becomes crucial. Stimulus sets that represent self- and other-concept in implicit measures can vary in their degree of idiography. Generic and idiographic stimulus selections reflect the *nomothetic* and the *idiographic* approach to person description—a distinction developed by later Heidelberg philosopher Windelband (1894/1998), introduced into psychology by Münsterberg (1899/1994), and popularised by Allport (1937). The nomothetic approach draws on normative, standardised procedures to compare individuals, mostly using quantitative data. The idiographic approach relies on person-specific profiling, often using qualitative data. The latter approach takes many degrees of freedom to reach a full understanding, but at the risk of volatile data and low chances for exact replication. Neither of these approaches is *a priori* more legitimate than the other (cf. Molenaar, 2004; Tuerlinckx, 2004).

Thus, pronouns such as I or MINE are used as normative stimuli, in line with a nomothetic approach to measurement. In this case, the concepts and the measurement procedure can be held constant, applied to anyone, and still bear somehow on individual characteristics. At the downside, generic pronouns are distal and hardly descriptive of people's cores. (Likewise, nomothetic stimuli such as THEY or THEIR can represent the concept *others* or *not-me*, but they remain rather distant and vague.) By contrast, idiographic stimuli, such as name and birthday, convey highly person-specific aspects, yield vivid mental representations, and typically imply temporally invariant characteristics. Therefore, idiographic aspects are closer to the core of the self. At an intermediate level between idiographic and generic stimulus selections, sociographic stimuli (e.g. FEMALE or STUDENT) can be used to represent social categories such as gender or other social groups (Richetin, Richardson, & Mason, 2010). Some of these social group memberships can change, others cannot; but sociographic stimuli always differ from idiographic ones by their degree of person specificity. Thus, the proximity of sociographic stimuli to the core self—on average—tends to be lower than that of idiographic stimuli.

The CCM suggests that even *within* a level of relatively low proximity, some stimuli might still be more central than others. The self can be tapped via semantic routes at levels even lower than that of pronouns, provided they are idiographic, pre-meditated, and accessible. The name-letter task exploits this logic in using the liking of name letters as indication of self-esteem (Nuttin, 1985). The effect is strongest for participants' initials commonly used for acronyms (Hoorens & Todorova, 1988; Stieger, Voracek, & Formann, 2011). Birthday numbers are likewise associated with the self and better liked than other numbers (Kitayama & Karasawa, 1997; Koole, Dijksterhuis, & Van Knippenberg, 2001).

Taken together, the CCM extends Greenwald et al.'s (2002) unified theory to incorporate the *centrality* of items and to clarify the semantic overlap and relative hierarchical distance between items and concepts. Centrality considerations

apply to all sorts of concepts, and misspecification of categories by choosing suboptimal stimuli runs the risk of introducing nuisance variance. A specific methodological consequence following from the CCM is the hitherto untested hypothesis that IATs employing idiographic stimuli should have higher validity than IATs employing generic stimuli. There is some evidence to support this assumption. Greenwald and Farnham (2000) used both a generic and an idiographic self-concept IAT. The generic IAT consisted of preselected self-concept items (pronouns) whereas a second IAT comprised idiographic items, that is, participants themselves chose target stimuli that were descriptive for the individuals (e.g. their names). The researchers found stronger I–E correlations for the idiographic gender-self-concept IAT as compared with the generic IAT,  $|r| = .33$  vs.  $.22$ , despite a high correlation between both types of IAT ( $r = .68$ ). The authors hesitated to infer the equivalence of the two IAT types—the idiographic IAT 'better defined that latent variable' in a confirmatory factor analysis—yet they suggested that the generic format was 'likely to be the more efficient', as the idiographic format requires collecting participant-specific information (pp. 1030–1031).

This initial finding was later corroborated by a meta-analysis on the relationship between IATs and explicit measures (Hofmann, Gawronski, Gschwendner, Le, & Schmitt, 2005). The study found higher I–E consistency for idiographic than generic IATs ( $r = .32$  vs.  $.15$ ), but this analysis rested on a low number of idiographic studies ( $k = 6$ ). In addition, the meta-analysis compared IAT types post hoc and across different study contents, samples, and explicit measures. Therefore, it does not allow for causal inferences and can merely hint to generalisability *if* causality is established independently. In sum, these data suggest that idiographic IATs may assess self-relevant associations better than generic IATs.

To date, possible differences between generic and idiographic IATs have neither been theoretically explained nor systematically investigated. The current research therefore (1) suggests a theoretical model, the CCM, to explain differences between idiographic and generic self-concept measures, and (2) tests the CCM-derived prediction that idiographic stimuli outperform generic stimuli in self-concept IATs *across* topics, *within* samples (repeated measures), while using the *identical* validation criteria.

## OVERVIEW OF EXPERIMENTS AND MAIN HYPOTHESIS

From the CCM logic, it follows that idiographic stimuli should raise the focus on the core self more than generic stimuli do. Using pronouns can be sufficient to tap into the self-concept, but an idiographic representation of IAT categories should reliably favor a central and more resonant route to the self, resulting in higher validity of assessed self-associations (Perugini, O'Gorman, & Prestwich, 2007). Whereas the general theoretical hypothesis is that tapping into core concepts increases validity, the main empirical hypothesis for the present research is: Idiographic self-representations reduce nuisance variance in IATs, as they narrow the nomothetic–idiographic gap, and this will be evident in

stronger relationships of idiographic than generic implicit measures with domain-specific criteria.

We compared idiographic IATs (henceforth idio-IAT) that applied individual descriptions of self/others (e.g. first name, date of birth) to generic IATs (gen-IAT) that used pronouns (e.g. I, YOU; cf. Bosson, Swann, & Pennebaker, 2000; Greenwald & Farnham, 2000). To limit the between-participant variation, our selection procedure for idiographic stimuli restricted which person the participants could choose to represent the *other*-category (Karpinski, 2004). Depending on a participant's choice of the other-category, self and other might have been confounded with positivity differences, or involved a gender-mismatch. Therefore, we instructed participants to choose a well-known person such as a close friend or a relative of the same sex and of similar age. Similar approaches have been used by other researchers (e.g. Yamaguchi et al., 2007; Franck, De Raedt, Dereu, & Van den Abbeele, 2007). Moreover, we avoided fluctuating items that might only be vaguely and non-permanently related to the self, such as telephone number or student ID. To increase statistical power, cross-sectional within-subject designs required that participants took both IATs on the same occasion. We tested the better quality of idiographic IATs for self-evaluations ('implicit self-esteem'; Experiment 1), body-schema ('implicit weight identity'; Experiment 2), and the 'automatic aggressive self-concept' (Experiment 3).

Which criteria can be used to establish the higher validity of any IAT? Implicit and explicit self-concepts are assessed with implicit and explicit measures. Neither of these is process pure, so, the conceptual level must not be equated with the measurement level. For each topic, we will discuss whether the utility of idio-IAT will be evident in direct or interactive relations with explicit measures (Perugini, 2005; Perugini, Richetin, & Zogmaister, 2010). Generally, implicit and explicit measures at least partly depend on the same associative basis and can influence each other (Gawronski & Bodenhausen, 2006; Strack & Deutsch, 2004; Whitfield & Jordan, 2009); therefore, a complete I-E dissociation is unlikely in the first place (Nosek, 2007). Research has also shown that increasing self-accessibility results in stronger implicit-explicit correlations (LeBel, 2010; Perugini et al., 2007); this should make I-E consistency more likely for idiographic IATs. If one IAT variant consistently results in higher I-E correlations in a head-to-head comparison—equal reliability provided—we will conclude that lower correlations indicate lower validity (Hofmann, Gschwendner, Nosek, & Schmitt, 2005). To reduce same-source bias, we will also employ objective indicators such as observer assessment, physical properties, and objective behaviour.

## EXPERIMENT 1

In the first study, we looked at implicit self-esteem. To our knowledge, idiographic self-esteem IATs have hardly ever been used, hence, idiographic and generic self-esteem IATs have not been compared within participants or between samples drawn from the same population. Powerful within-sample

comparisons should reveal validity differences, if they exist. Following the CCM, we hypothesised that idio-IAT would show stronger signs of validity than gen-IAT with regard to the following criteria: First, we expected higher I-E consistency for IATs that apply idiographic rather than generic stimuli (Hofmann, Gawronski, et al., 2005). Despite some non-significant I-E self-esteem correlations in the literature (e.g. Baccus, Baldwin, & Packer, 2004; Jordan, Spencer, Zanna, Hoshino-Browne, & Correll, 2003; Koole et al., 2001), low but significant I-E relationships for self-esteem have repeatedly been found before (e.g. Conner & Barrett, 2005; Gebauer, Riketta, Broemer, & Maio, 2008; Greenwald & Farnham, 2000; Jordan, Whitfield, & Zeigler-Hill, 2007; Karpinski, 2004; Krizan, 2008; Krizan & Suls, 2008; LeBel, 2010; Olson, Fazio, & Hermann, 2007; Wentura et al., 2005). Second, implicit measures have predicted non-verbal behaviour indicative of personality aspects that people are unable to report (e.g. Asendorpf, Banse, & Mücke, 2002), so, we expected higher correlations of the idio-IAT than gen-IAT with experimenters' ratings of participants' self-esteem. Third, we expected higher correlations of idio-IAT than gen-IAT with sense of coherence (SOC; Antonovsky, 1979). SOC acts as a moderator of healthy (or unhealthy) responses to stress and serves as a proxy for physical and mental health (Eriksson & Lindström, 2006). It is a global orientation that entails (a) how good an individual is at comprehending things that occur to her, (b) how skilled an individual is to deal with challenges throughout life, and (c) how meaningful life is experienced despite obstacles. As such, SOC should be positively related to self-esteem (Johnson, 2004).

## Method

### *Procedure and materials*

*Participants.* One-hundred and forty-four (29 male) psychology undergraduates at Heidelberg University ( $M_{age} = 23.55$  years;  $SD = 7.17$ ) took part in a lab study on computer-based assessment in exchange for course credit. Participants worked in groups of three to six, yet individually, in cubicles, first on gen-IAT, then idio-IAT, and finally, the questionnaires. Once they started working on the tasks, the experimenter rated the self-esteem of participants. After the examination had ended, participants were debriefed and compensated.

*Procedure.* The order of (a) measures, (b) IAT blocks (within the measures), and (c) IAT stimuli (in the blocks) remained constant across participants because we were primarily interested in correlations, not mean IAT effects. We were guided by the following reasoning: First of all, counterbalancing the design factors would have introduced error variance that is usually detrimental for correlations (Egloff & Schmukle, 2002). Second, although meta-analytic findings indicate that—on average and across all attitudinal domains—no systematic differences for I-E-order could be observed (Hofmann, Gawronski, et al., 2005), specifically with regard to self-concept, stronger I-E consistency has emerged when the explicit measure preceded the implicit measure (e.g. Bosson et al., 2000; Glen & Banse, 2004; Krizan & Suls, 2008). As prior reflection might influence the

subsequent assessment of implicit self-esteem, resulting in biased I–E consistency owing to a methodological factor, we decided that explicit measures should follow implicit ones. Third, idio-IAT might invoke a person's self-concept more vividly than gen-IAT, so an idio-IAT–gen-IAT order might have induced unwanted carry-over effects (LeBel, 2010). Finally, as most previous studies relied exclusively on gen-IAT, we kept its position comparable to the one in previous studies, thereby putting idio-IAT at a potential disadvantage as later IATs tend to suffer from fatigue or strategic effects, error proneness, and reduced reliability (Greenwald, Nosek, & Banaji, 2003; Nosek, Greenwald, & Banaji, 2007). In sum, a standardised order conserves the properties of typical gen-IATs, while putting the presumably higher validity of idio-IATs to a stringent test.

#### Observer assessment of self-esteem

After introducing themselves and after having seated participants at the computers, approximately 1 minute after arrival, the experimenters rated participants' self-esteem on a single 11-point rating scale (0 = no self-esteem at all; 10 = very high self-esteem). Experimenters were unacquainted with the participants and unaware of name, date of birth, or explicit and implicit self-esteem scores (cf. Jordan et al., 2007; Ranganath, Smith, & Nosek, 2008). Therefore, they had to rely on spontaneous gut reactions for their judgments. Prior to the data collection, experimenters underwent a coding training and agreed on the usage of the scale. By identifying people they all knew and by agreeing on which answering categories to assign to them, they anchored the meaning of the scale. Training highlighted the importance of paraverbal and nonverbal cues (speech parameters, facial expression, gaze avoidance, or head position; cf. Hofmann, Gschwendner, & Schmitt, 2009). Experimenters formed an overall impression of each participant during the initial interaction at the beginning of the study (less than a minute). We restricted coding to a *global overall rating* because extensive coding activity on several participants at the same time was infeasible and would have distracted and possibly raised suspicion. Research has shown that single-item measures can be reliable measures of *global* concepts (Wanous & Reichers, 1996; Wanous, Reichers, & Hudy, 1997) and that single-item measures can be good proxies for *global self-esteem* (Robins, Hendin, & Trzesniewski, 2001).

#### Implicit measures: self-esteem-IATs

Gen-IAT and idio-IAT were applied as a nine-block procedure that required approximately 10 minutes (404 trials; see Table 1).

Category labels and procedural parameters (e.g. inter-trial interval of 150 minutes) were identical. Exceptions to the rule are (a) the stimuli that represented the self-categories and other-categories, and (b) the missing single discrimination of positive–negative attributes for the second IAT. Every stimulus was practised and presented at least once in the respective practice and combined blocks. Generic target stimuli for self were ICH (I), SELBST (myself), MIR (me), MEINE (mine), EIGEN (own). The other-category was represented by ANDERE (others), EUCH (you [dative plural]), IHR (her or you [vocative plural]), EURE (your [genitive plural]), FREMD (foreign [German for not-me/strange/not familiar with]). Idiographic target stimuli for self were the first name, family name, place of birth, year of birth, and month and day of birth (e.g. 22 June). Equivalent information was obtained for a close and well-known person (e.g. a friend or a relative, such as a cousin) of roughly the same age and identical sex. To not confuse self- and other-category, items had to be different.

For both IATs, the valence categories contained 18 stimuli each to reflect the multiple dimensions of the multidimensional self-esteem scale (MSES; Schütz & Sellin, 2006; see succeeding text). This was done to represent self-esteem broadly and vividly and to increase the chances of conceptual I–E consistency. Three pairs of adjectives were derived from each of the six MSES-subcales (see Appendix). The adjectives did not differ significantly in word length or number of syllables ( $t_s < 1$ ), but with regard to valence according to the average ratings of two raters who rated each item on 5-point rating scales from  $-2$  (very negative) to  $+2$  (very positive) (inter-rater reliability: Krippendorff's (1970)  $\alpha$  [ordinal] = .81),  $M_s = 1.19$  vs  $-1.11$  ( $SD_s = 0.41$  vs  $0.72$ ),  $t(34) = 11.81$ ,  $p < .001$  (corrected for unequal variances), Cohen's  $d = 3.94$ . We computed IAT effects as  $D_5$ -scores (see Greenwald et al., 2003). Higher scores represent more positive self-esteem. Cronbach's  $\alpha$  of gen-IAT and idio-IAT amounted to .87 and .89, respectively.

#### Explicit measures: self-esteem and sense of coherence

First, a German version of the Rosenberg global self-esteem scale (RSES; Rosenberg, 1979; von Collani & Herzberg, 2003) was administered. Ten items assessed self-enhancement and self-derogation as positive and negative aspects of global self-esteem on 4-point rating scales ( $\alpha = .86$ ). RSES is not domain specific and taps into a person's overall global self-worth. As a more specific measure, participants filled in the multidimensional self-esteem scale (MSES; Schütz & Sellin, 2006).

Table 1. Procedure of Generic and Idiographic Implicit Association Test (Experiment 1; Analogous Setup for Exp. 2–3)

| Block no. | Left response    | Right response   | Target stimuli | No. of trials |
|-----------|------------------|------------------|----------------|---------------|
| 1         | Self             | Other            | Generic        | 20            |
| 2         | Positive         | Negative         | .              | 36            |
| 3         | Self + positive  | Other + negative | Generic        | 72            |
| 4         | Other            | Self             | Generic        | 20            |
| 5         | Other + positive | Self + negative  | Generic        | 72            |
| 6         | Self             | Other            | Idiographic    | 20            |
| 7         | Self + positive  | Other + negative | Idiographic    | 72            |
| 8         | Other            | Self             | Idiographic    | 20            |
| 9         | Other + positive | Self + negative  | Idiographic    | 72            |

The 32 MSES-items were answered on 7-point rating scales ( $\alpha = .93$ ). The MSES features (1) emotional self-esteem, (2) social self-esteem during social contact, (3) capability to deal with criticism, (4) performance-related self-worth, (5) self-esteem owing to physical attractiveness, and (6) self-esteem owing to athleticism. The overall self-esteem score, based on all the items, served as a validity criterion. Finally, the short form of Antonovsky's Sense of Coherence scale (SOC-13; Antonovsky, 1993, 1997) assessed the extent to which participants express confidence in life as being predictable and explicable (comprehensibility), their availability of resources to meet the demands throughout life (manageability), and the general significance of life in terms of challenges and worthiness of investment (meaningfulness). We used all answers on the 7-point rating scales to form an aggregate scale score ( $\alpha = .78$ ).

#### Data analysis

Unless indicated otherwise, potential univariate and bivariate outlier values (Mahalanobis distance) were not detrimental to analyses and conclusions. When contrasting IAT correlations, within-sample planned comparisons with one-tailed testing were used (Steiger, 1980).

### Results and discussion

As regards convergent validity, global self-esteem (RSES) and multidimensional self-esteem (MSES) showed high overlap (cf. Table 2). By comparison, the relationship between the IATs was significant, but relatively weak, given that two

equally reliable measures targeted at implicit self-esteem. Both IAT variants correlated significantly and to a similar extent with RSES. Only idio-IAT showed a significant relationship with MSES, resulting in a marginally significant difference in correlations between the two IAT variants, Steiger's  $Z_{df=141} = 1.26$ ,  $p = .10$ . The weak correlation between the two IATs reduced the power for this planned comparison (Steiger, 1980). [Our sample size was in line with power considerations based on Greenwald and Farnham's (2000) finding of  $r = .68$  between the IAT variants. Had our IATs correlated at this level, or even only as low as .40, the  $p$ -value would have been significant.] In addition, idio-IAT correlated higher than gen-IAT with the experimenters' ratings of participants' self-esteem ( $Z_{df=141} = 1.64$ ,  $p = .05$ ), and descriptively so with sense of coherence ( $Z_{df=141} = 1.25$ ,  $p = .11$ ). Observers' intuitively reported *summary evaluations* of global self-esteem were significantly, although weakly, related to global self-esteem according to RSES (cf. Robins et al., 2001). These global observer ratings may not qualify as a comprehensive assessment of participants' self-esteem-based behaviour, but they may capture less controlled aspects of a participant's facial mimics, gaze, voice, body posture, and interaction style (cf. Hofmann, Friese, & Strack, 2009; Hofmann, Gschwendner, et al., 2009) that leak through from self-esteem during the brief encounter. In the midst of well-established multiple-item measures of self-esteem, idio-IAT was the strongest predictor of observer ratings of participants' self-esteem.

In sum, Experiment 1 supported the usefulness of idiographic IATs in line with the expectations derived from the CCM. When implicit self-esteem was measured with idiographic items rather than generic items, the IAT showed

Table 2. Descriptives and Interrelations (Experiments 1–3)

|  | <i>M</i> | <i>SD</i> | 2       | 3      | 4                 | 5                 | 6      |
|--|----------|-----------|---------|--------|-------------------|-------------------|--------|
| <b>Experiment 1 (<i>N</i> = 144)</b>         |          |           |         |        |                   |                   |        |
| 1. Rosenberg self-esteem (RSES)              | 2.33     | 0.50      | 0.74*** | 0.18*  | 0.51***           | 0.21*             | 0.20*  |
| 2. Multidimensional self-esteem (MSES)       | 4.70     | 0.84      |         | 0.14   | 0.57***           | 0.09              | 0.22** |
| 3. Observer rating of self-esteem (ORSE)     | 5.72     | 1.09      |         |        | 0.15 <sup>†</sup> | 0.07              | 0.24** |
| 4. Sense of coherence (SOC)                  | 4.65     | 0.80      |         |        |                   | 0.04              | 0.17*  |
| 5. Generic IAT                               | 0.66     | 0.46      |         |        |                   |                   | 0.22** |
| 6. Idiographic IAT                           | 0.25     | 0.36      |         |        |                   |                   | -      |
| <b>Experiment 2 (<i>N</i> = 123)</b>         |          |           |         |        |                   |                   |        |
| 1. Body Mass Index                           | 23.01    | 3.11      | .76***  | .43*** | .55***            | .10               | .14    |
| 2. Body Image                                | 4.20     | 0.92      |         | .56*** | .74***            | .03               | .18*   |
| 3. Body Ideal                                | 3.77     | 0.62      |         |        | -.15              | .08               | .10    |
| 4. Body Dissatisfaction                      | 0.43     | 0.77      |         |        |                   | -.03              | .14    |
| 5. Generic IAT                               | -0.66    | 0.47      |         |        |                   |                   | .48*** |
| 6. Idiographic IAT                           | -0.11    | 0.34      |         |        |                   |                   | -      |
| <b>Experiment 3 (<i>N</i> = 125)</b>         |          |           |         |        |                   |                   |        |
| 1. Caprara Irritability Scale                | 3.04     | 0.60      | .51***  | .69*** | .04               | .13               | -.01   |
| 2. Gladue Aggression Inventory - Impulsivity | 2.55     | 0.56      |         | .50*** | -.07              | .04               | -.06   |
| 3. Buss & Perry Aggression Questionnaire     | 2.34     | 0.46      |         |        | .12               | .11 <sup>a)</sup> | .03    |
| 4. Aggression: Neg. Pictures (%)             | 0.30     | 0.21      |         |        |                   | .05               | .05    |
| 5. Generic IAT                               | -0.78    | 0.40      |         |        |                   |                   | .56*** |
| 6. Idiographic IAT                           | -0.24    | 0.36      |         |        |                   |                   | -      |

#### Note:

<sup>†</sup> $p < .10$ ,

\* $p < .05$ ,

\*\* $p < .01$ ,

\* $p < .001$ ;

<sup>a)</sup> $N = 123$  (due to outliers; cf. Fn.1).

stronger relations with several criteria such as multi-dimensional self-esteem, observer assessments of participants' self-esteem, and sense of coherence.

## EXPERIMENT 2

Subjective body schemes play an important role in the self-concept (Markus, Hamill, & Sentis, 1987). Whereas body image refers to the mental representation of one's own body shape, body ideal represents an endorsed standard and is conceptually independent of one's current body shape. (Empirically though, when people at the more extreme ends of the weight spectrum are included in a sample, actual body shape exerts some influence on what body shape is considered to be ideal; e.g. Woodman & Steer, 2011). Insofar as discrepancies between body image and body ideal exist, they will lead to body dissatisfaction (e.g. Williamson, Gleaves, Watkins, & Schlundt, 1993). To the extent that participants spontaneously associate themselves differently with thinness and thickness, they show 'implicit weight identity' (Grover, Keel, & Mitchell, 2003). With previous findings, implicit weight identity should mirror to some extent the explicitly reported current body shape and objective physical properties such as body-mass index (BMI; cf. Grover et al., 2003). By contrast, implicit weight identity should not reflect one's own or society-based standards such as body ideal. In line with the CCM, idio-IAT should reflect implicit weight identity better than gen-IAT.

## Method

### Procedure and materials

**Participants.** One-hundred and twenty-three students of the University of Heidelberg (43 male) took part in the experiment in exchange for course credit or a snack bar and a soft-drink for the 20-minute session. Their mean age was 22.51 years ( $SD=7.33$ ).

**Procedure.** First, participants completed implicit measures of weight concept and provided socio-demographic data including self-reported size and weight. Next, explicit body image and body ideal were assessed, as well as objective physical properties.

### Implicit measures: weight-identity-IATs

The setup largely followed Experiment 1, with the exception of the attribute categories *thin* and *thick*. Trait words that formed 10 pairs of opposite content were chosen as attribute items, under the constraint that stimulus valence and word length were matched (see Appendix). IAT scores were computed as before (subtracting *self+thick* from *self+thin*). The higher an IAT effect, the stronger was implicit weight identity related to thickness. The IATs (both  $\alpha_s = .87$ ) correlated at  $r = .48$ .

### Explicit measures and body mass index

Participants were presented with a sex-specific figure rating scale (Stunkard, Sorensen, & Schulsinger, 1983). These visual analogue scales, which apply body silhouettes from thin to fat along a fine-graded ruler (1 = smallest ectomorph; 9 = largest

endomorph), asked for the current body image as well as the subjective body ideal (Rand, Resnick, & Seldman, 1997). Higher ratings indicate a robust body shape, or a personal preference toward robust forms. We also obtained self-reports and factual data on weight and size, yet exclusively the latter ones were used for the analysis (cf. Nyholm et al., 2007). BMI scores were computed as weight (in kilograms) divided by the squared size (in meters). On average, the sample was neither underweight (BMI < 18.5), nor overweight (BMI > 25),  $M = 23.01$  ( $SD = 3.11$ ).

## Results and discussion

Implicit weight identity showed significant convergence with body image only when assessed with idio-IAT (cf. Table 2), making it a stronger predictor than gen-IAT, Steiger's  $Z_{df=120} = 1.62$  ( $p = .05$ ). As expected, no significant correlations of implicit weight identity with body ideal emerged, which can be taken as indication of discriminant validity. Surprisingly, there was no direct relationship with BMI. To account for unsolicited influences of sex and age, which shared variance with body schema variables, we regressed an IAT simultaneously with control variables in multiple regression analyses (cf. Table 3). Idio-IAT, not gen-IAT, correlated uniquely with body image, but not with body ideal, attesting to the convergent and discriminant validity of the idio-IAT once more. After controlling for sex and age, idio-IAT correlated significantly with BMI too, whereas gen-IAT did not. These data support

Table 3. Convergent and Discriminant Validity (Experiment 2): Multiple Regressions of Body Image, Body Ideal, and Body Mass Index on Age and Sex as well as Idiographic (Model 1a) and Generic (Model 1b) Implicit Measures

|                        | <i>R</i> | <i>R</i> <sup>2</sup> | <i>F</i> | $\beta$ | <i>p</i> |
|------------------------|----------|-----------------------|----------|---------|----------|
| <b>Body Image</b>      |          |                       |          |         |          |
| Model 1a               | .43      | .18                   | 8.78     |         | < .001   |
| Sex                    |          |                       |          | .30     | .001     |
| Age                    |          |                       |          | .20     | .017     |
| Idiographic IAT        |          |                       |          | .19     | .022     |
| Model 1b               | .38      | .15                   | 6.76     |         | < .001   |
| Sex                    |          |                       |          | .29     | .001     |
| Age                    |          |                       |          | .20     | .021     |
| Generic IAT            |          |                       |          | .04     | .617     |
| <b>Body Ideal</b>      |          |                       |          |         |          |
| Model 1a               | .58      | .33                   | 19.68    |         | < .001   |
| Sex                    |          |                       |          | .13     | .088     |
| Age                    |          |                       |          | .53     | < .001   |
| Idiographic IAT        |          |                       |          | .09     | .217     |
| Model 1b               | .57      | .32                   | 18.99    |         | < .001   |
| Sex                    |          |                       |          | .13     | .098     |
| Age                    |          |                       |          | .53     | < .001   |
| Generic IAT            |          |                       |          | .03     | .698     |
| <b>Body mass index</b> |          |                       |          |         |          |
| Model 1a               | .54      | .29                   | 16.49    |         | < .001   |
| Sex                    |          |                       |          | .36     | < .001   |
| Age                    |          |                       |          | .32     | < .001   |
| Idiographic IAT        |          |                       |          | .16     | .045     |
| Model 1b               | .53      | .28                   | 15.40    |         | < .001   |
| Sex                    |          |                       |          | .37     | < .001   |
| Age                    |          |                       |          | .31     | < .001   |
| Generic IAT            |          |                       |          | .10     | .193     |

Note:  $N = 123$ ,  $df = 1, 119$ ;  $\beta$  = standardised beta coefficient.

the superiority of idiographic stimuli over generic stimuli for establishing the validity of implicit weight identity.

### EXPERIMENT 3

The 'automatic aggressive self-concept' has been successfully used to predict rough and irregular behaviour in sports, although this behaviour is difficult to foretell by self-reports (Banse & Fischer, 2002). It also reflects long-term and short-term consequences of playing violent computer games (Uhlmann & Swanson, 2004; Bluemke, Friedrich, & Zumbach, 2010) as well as treatment effects of social skills trainings (Gollwitzer, Banse, Eisenbach, & Naumann, 2007). Because of typically low I-E consistency in aggression research (Bluemke & Zumbach, 2007), deviating from our predictions in Experiments 1 and 2, we expected neither idio-IAT nor gen-IAT to correlate substantially with explicit measures. Instead, we expected incremental validity when predicting aggressive behaviour.

Implicit measures are usually good at predicting uncontrolled and nonverbal behaviours (e.g. Asendorpf et al., 2002). As participants can usually monitor and control their behaviour, the usefulness of implicit measures strongly depends on participants' self-control (Friese, Hofmann, & Schmitt, 2008). To begin with, we used an unobtrusive picture-choice task that reduces the likelihood that the behaviour *is* actually monitored and controlled (Webb, Campbell, Schwartz, & Sechrest, 1966). Yet, the incremental validity of aggressiveness-IATs should become particularly obvious when explicit measures indicate low self-control (Friese & Hofmann, 2009; Hofmann, Friese, et al., 2009). Participants low in trait self-control are known to respond more aggressively when provoked (Stucke & Baumeister, 2006; DeWall, Baumeister, Stillman, & Gailliot, 2007). Even in the absence of provocation, people differ in how they construe an ambivalent situation (irritability), the amount of behavioural control they exert (impulsivity), and their overall likelihood to control their anger or show hostility (trait aggressiveness). An individual's propensity for aggression may then be predicted on grounds of impulsive precursors (as indicated by IAT scores) when self-control is low, in other words, when irritability, impulsivity, and trait aggressiveness are high. Instead of a direct relationship, or incremental validity beyond explicit measures, we hypothesized an I × E interaction (Perugini, 2005; Perugini et al., 2010), and this interaction should be more pronounced for idio-IAT.

### Method

#### *Procedure and materials*

**Participants.** Students of various majors at the University of Heidelberg took part in a lab study on personality and reaction times in exchange for course credit or a snack and a soft-drink. Data sets from  $N=125$  (43 male;  $M=23.38$  years,  $SD=6.90$ ) were available.<sup>1</sup> Because of the aversive nature of the

<sup>1</sup>Data from two participants had to be dropped exclusively from the correlation/regression analysis involving gen-IAT and BPAQ, because when including them two bivariate (MAHAL) outliers distorted this specific relationship,  $r=.19$  ( $p=.03$ ), and the I × E interaction,  $\Delta R^2=.04$ ,  $\beta=.20$  ( $p=.03$ ). We deem these unpredicted results unreliable, because they are not supported by the other analyses, and because their significance hinges on a mere two data points. For all other analyses, the full number of participants was retained.

aggression measure, we took care that nobody suffered from depressive or suicidal thoughts or underwent psychotherapeutic treatment at the time.

**Procedure.** Before the participants could become aware of the study topic, we covertly assessed aggressive behaviour. Upon arrival, the experimenter asked the participant for help with an additional 1-minute task before the main examination, namely to select 10 out of a set of 30 pictures. Ostensibly, this random selection of pictures was needed for another participant in an unrelated study. Given the many experiments run at the institute at the time, the cover story was plausible. After this measure of aggressive behaviour, the session continued with the implicit measures and paper-and-pencil questionnaires. Finally, the experimenter debriefed and compensated the participants. Funneled debriefing confirmed that none of the participants suspected the picture selection to be related to the experiment.

#### *Behavioural measure of aggression*

Following Mussweiler and Förster's (2000) procedure, the participants saw 30 photos from the International Affective Picture System (Lang, Bradley, & Cuthbert, 2005). Participants were left for themselves to select the pictures they considered most suitable for an approaching fellow student. The lack of any explicit selection rules underscored the randomness of the choice. Half of the randomly ordered photos (15 x 20 cm) displayed positive content; the other half displayed negative, emotionally disturbing content. Behavioural aggression was indexed by the ratio of negative to total number of the pictures selected by a participant. To the extent that participants choose negative pictures they violate the implicit social norm to not inflict pain on others. When prompted, participants are well aware of the psychological consequences of the procedure and dislike receiving negative pictures themselves (Denzler, Förster, & Liberman, 2009). As the images are aversive and create discomfort in the recipient, selecting negative pictures and exposing others to them constitutes *unprovoked* aggression (Mussweiler & Förster, 2000).

#### *Implicit measures: aggression-IATs*

The only change in setup involved the categories *peaceful* and *aggressive*. Two pairs of complementary action verbs, matched in word length, mirrored each of the subscales of the Buss and Perry (1992) aggression questionnaire (physical aggression, verbal aggression, anger, and hostility; see Appendix). Higher IAT scores indicate stronger associations of the self with aggressive relative to peaceful. Both gen-IAT and idio-IAT were reliable ( $\alpha s=.82$  and  $.85$ ) and correlated at  $r=.56$ .

#### *Explicit measures: irritability, impulsivity, and aggressiveness*

Participants completed three questionnaires that became increasingly blatant in how they asked about the target construct, aggressiveness. First, a German version of the 20-item Caprara Irritability Scale (CIS; Caprara et al., 1985; Bluemke & Steinmayr, 2008), assessing the tendency to aggress impulsively in response to mild or even no provocation, hence *irritability* (6-point rating scales, e.g. 'I can't help being a little rude to people I don't like',  $\alpha=.82$ ). Next, a German translation of the Gladue



Aggression Inventory followed (GAI; Gladue, 1991; Bluemke & Steinmayr, 2008), which provided seven items measuring (*aggressive*) *impulsivity* (GAI-imp; 5-point rating scales, e.g. 'I become easily impatient and irritable if I have to wait',  $\alpha = .51$ ). Finally, applying the same answering format, the widely-used 29-item Aggression Questionnaire (BPAQ; Buss & Perry, 1992) served as a measure of *trait aggressiveness* along the facets of physical and verbal aggression, anger, and hostility. For the German BPAQ-adaptation, only 27 of the 29 items entered the composite score to satisfy the intended factor structure (Herzberg, 2003; von Collani & Werner, 2005,  $\alpha = .82$ ).

**Results and discussion**

As expected, the explicit measures converged in the medium range, but none of them correlated significantly with the IATs, yielding the typical I-E dissociation in aggression research (Bluemke & Zumbach, 2007; see Table 2). Given that the explicit measures were substantially correlated, and to maximise reliability, we ran a Principal Component Analysis (PCA) on irritability, impulsivity, and trait aggressiveness scores. The PCA suggested a strong first component (*Eigenvalues* = 2.14, .55, and .31, explaining 71%, 18%, and 10% of the variance). All three measures loaded significantly on the first component ( $\lambda = .88, .78, \text{ and } .88$  for CIS, GAI-imp, and BPAQ, respectively). The factor scores of the one-component solution were taken as explicit aggressiveness index that was subjected to further analyses. Separate regression analyses were used to inspect the incremental validity of idio-IAT and gen-IAT. Following Aiken and West's (1991) recommendations on centering variables, we predicted aggressive behaviour by explicit aggressiveness (PCA scores), implicit measure (IAT scores), and the respective  $I \times E$  interaction effect (see Table 4). Although main effects were not significant, the interaction involving idio-IAT, but not gen-IAT, explained variance significantly. Simple slopes analyses showed that the propensity to aggress without being provoked increased with higher (+1 *SD*) idio-IAT-scores, but only for people who were high (+1 *SD*) in self-reported aggressiveness (Figure 2). In other words, when participants reported a dispositional lack of controlling aggression, the implicit measure gained predictive utility. Having high explicit and implicit scores resulted in the selection of a number of disturbing images almost twice as high as when either score was low. All explicit measures

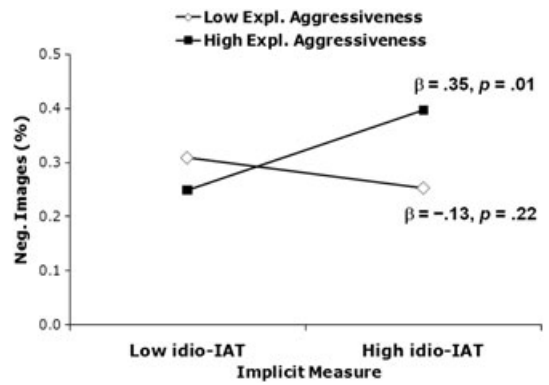


Figure 2. Prediction of unprovoked aggressive behaviour (proportion of negative images) by the implicit aggressive self-concept at different levels of explicit aggressiveness (Experiment 3).

yielded similar outcomes in separate analyses.<sup>2</sup> To summarise, neither measure predicted unprovoked aggression on its own. Instead, idio-IAT, but not gen-IAT, predicted unprovoked aggressive behaviour as a function of explicit aggressiveness. This result is in line with the CCM that predicts idio-IATs to better tap into the core self than gen-IATs.

**GENERAL DISCUSSION**

**Summary and Interpretation**

Three experiments confirmed our main hypothesis: Idiographic self-representations render IATs more suitable for the assessment of the implicit self-concept than do generic self-representations. Experiment 1 demonstrated that implicit and explicit measures of global self-esteem were related, regardless of whether gen-IAT or idio-IAT was taken as an indicator for implicit self-esteem. However, only idio-IAT captured the relations of implicit self-esteem to observer-assessed self-esteem, self-reported multidimensional self-esteem, and the salutogenetic criterion sense of coherence. Experiment 2 confirmed that implicit weight identity was related to explicit body image, yet only so when it was assessed by idio-IAT. Moreover, only for

Table 4. Incremental validity (Experiment 3): multiple regression of aggressive behavior on explicit aggressiveness (PCA-factor-scores), idiographic (Model 1a) and generic (Model 1b) implicit measures, and the respective Implicit  $\times$  Explicit interaction terms

|   | <i>R</i> | <i>R</i> <sup>2</sup> | <i>F</i> | $\beta$ | <i>p</i> |
|---|----------|-----------------------|----------|---------|----------|
| Model 1a                                  | .27      | .071                  | 3.10     |         | .029     |
| Explicit aggressiveness                   |          |                       |          | .10     | .273     |
| Idiographic IAT                           |          |                       |          | .11     | .228     |
| Idio-IAT $\times$ explicit aggressiveness |          |                       |          | .27     | .004     |
| Model 1b                                  | .16      | .024                  | 0.99     |         | .401     |
| Explicit aggressiveness                   |          |                       |          | .05     | .591     |
| Generic IAT                               |          |                       |          | .07     | .424     |
| Gen-IAT $\times$ explicit aggressiveness  |          |                       |          | .15     | .117     |

Note: *N* = 125, *df* = 1, 121;  $\beta$  = standardised beta coefficient.

<sup>2</sup>Separate models with CIS, GAI-imp, and BPAQ as predictors yielded analogous findings,  $\Delta R_{idio}^2 = .05, .04, \text{ and } .06$  ( $ps < .02$ ) vs.  $\Delta R_{gen}^2 = .01, .002, \text{ and } .01$  ( $ps > .24$ ), respectively. At this point, some readers may suggest full regression models that enter an explicit measure, both implicit measures, and their two interaction terms at once, or models that test the idio-IAT interaction as a final incremental step after the explicit measure, gen-IAT, idio-IAT, and gen-IAT  $\times$  E interaction have been controlled for. However, such regression models not only lower the ratio of data points per predictor in an unfavorable manner. To detect a small effect with sufficient statistical power, a sample size of several hundred individuals is required (Miles & Shevlin, 2001). More importantly, such models do neither provide an answer to the crucial question of incremental validity of an implicit over an explicit measure, nor to the question of incremental validity of an IAT's interaction term. Instead, they test predictors after two IAT main effects and one gen-IAT interaction have also been entered; consequently, they only test the variable *residual* after common variance among the predictors has been partialled out. Also, these models do not deal adequately with predictor multicollinearity; in our case, gen-IAT and idio-IAT were substantially correlated, and their two interaction terms are more than 50% mathematically strictly dependent. Although redundant independent variables are unlikely to emerge as significant predictors, these full regression models showed increments for the idio-IAT  $\times$  E interaction as the last predictor,  $\Delta R^2 = .05, .05, .03, \text{ and } .05$  ( $p = .02, .02, .08, \text{ and } .02$ ) for CIS, GAI-imp, BPAQ, and PCA-scores, respectively.

idio-IAT could a relationship to physical properties such as BMI be established. In Experiment 3 idio-IAT, but not gen-IAT, predicted aggression for individuals that lack self-control as indicated by their high scores in self-reported aggressiveness. Our experiments close a gap in the literature on the theoretical and empirical conclusiveness of a suggested methodological moderator of IAT validity (Hofmann, Gawronski, et al., 2005): the idiographic versus generic stimulus selection for representing the self and other. These findings are among the first to show that the stimulus selection in IATs may indeed influence the validity of these IATs, not only the magnitude of IAT effects (e.g. Bluemke & Friese, 2006; Govan & Williams, 2004).

Irrespective of the content domain, several times when gen-IAT failed to show significant direct or interactive relations with validity criteria, idio-IAT had a higher likelihood to reveal them. Remarkably, idio-IAT performed better despite being related to gen-IAT to a sizeable extent. Note that these findings were obtained although the idiographic selection of stimuli introduces variability across subjects, both with regard to the representation of the self, and with regard to the representation of the other. Thus, the chances to obtain meaningful correlations for idio-IAT were *a priori* lower than for gen-IAT, and still we never found a single instance where gen-IAT outperformed idio-IAT. We infer from this direct comparison that the chances to tap into the relevant social knowledge structures with implicit measures are higher when using an idiographic as compared with a generic approach. Had we relied on gen-IATs instead, most of the relationships with several criteria would have been obscured. Idiographic representations allowed a more optimistic view on IAT validity.

The empirical hypothesis is closely linked with the theoretical hypothesis derived from the CCM that stimulus centrality is a crucial ingredient in what concepts are activated in implicit measures. Our findings are compatible with the CCM. The relevance of idiographic stimuli can be delineated from their centrality for the concept in question. First, idiographic items allow a better representation of invariant person cores, and second, they are more proximal to the core self than generic pronouns. In combination, they ensure quick and accurate categorizations of stimuli—an essential ingredient for meaningful associations measured with speeded-classification tasks. The CCM is independent from the kind of implicit measure used, so our findings should conceptually replicate across other measures. Although it is clearly beyond the scope of this paper, the CCM also suggests further tests of the centrality assumptions for categories other than self-concept such as ethnic or gender groups.

### Limitations

With regard to the fixed order of implicit and explicit measures, one might assume that I–E consistency was better for idio-IAT than gen-IAT, because the former was taken in closer proximity to the questionnaires. Note, however, that several criteria were established before the IATs were taken (i.e., observer ratings of self-esteem, body proportions, aggressive behaviour). The proximity of the idiographic measure cannot be responsible for the higher relationships of idio-IAT with these criteria.

Logically, our findings cannot rule out that idio-IAT is a better predictor *only* when preceded by gen-IAT, as was the case in the present studies. Theoretically, the continuous activation of the self-concept by completing gen-IAT may have helped the subsequent idio-IAT to better tap into self-relevant information. However, we think that it is unlikely that the idio-IATs consistently correlated higher with validity criteria because they were used at the second position, as IATs at later positions have generally been found to yield less pronounced, less reliable, and less valid IAT effects (Nosek et al., 2007). Thus, by employing gen-IAT first and idio-IAT second, we put the hypothesis of higher validity of the idio-IAT to a particularly stringent test. In addition, the reverse order (idio-IAT first, then gen-IAT) would have been likely to invoke carry-over effects of idiographic self- and other-concepts due to a vivid pre-activation of these concepts during the idio-IAT (LeBel, 2010). Nevertheless, future research should use a between-participants approach to investigate the differential validities of generic and idiographic IATs to rule out any remaining ambiguities of the within-participants approach employed in the present research.

Finally, we used idiographic other-stimuli in our study, which may have increased error variance due to another non-standardised element on top of idiographic self-stimuli. However, given that the self-concept overlaps with ‘significant others’, such a procedure need not be detrimental; idiographic other-stimuli may actually help the assessment of the relevant implicit self-evaluations (Mashek, Aron, & Boncimino, 2003; DeHart, Pelham, Fiedorowicz, Carvallo, & Gabriel, 2011). Nevertheless, the role of the other-category in self-concept IATs is an important and under-researched issue. Future research may explore whether validity can be improved even further by using other-stimuli not pertaining to a close or well-known friend, but to a fictitious person or a prominent person known by all participants that can be kept constant across participants. Such a pragmatic mixture of idiographic and standardised elements might further reduce nuisance variance. This an open empirical question reaching beyond our initial theoretical question.

### Implications and Outlook

Our findings have implications for research on the self, and for self-concept IATs in particular. Part of the variance observed in I–E relations may be attributable to volatile self-representations. Generic representations seem to run a higher risk of merely assessing participants’ sorting speed, rather than self-related associations proper. Previous research may have underestimated I–E consistency as well as IAT-validity by relying on generic self- and other-representations (cf. Greenwald, Poehlman, Uhlmann, & Banaji, 2009). Some researchers have concluded from low I–E correlations that the IAT may not be a useful measure (e.g. Bosson et al., 2000); others have concluded that the dissociation indicates distinct mental representations (e.g. Rudolph, Schröder-Abé, Schütz, Gregg, & Sedikides, 2008; Krizan & Suls, 2009). Again, others have suggested that explicit measures might not be suitable validation criteria for implicit measures (e.g. Perugini et al., 2007). Such conclusions may have been premature. The importance of distinguishing conceptual from measurement level

cannot be stressed enough, as all theorizing involves assumptions about the conceptual level, whereas we can merely investigate the measurement level. Therefore, previous null-findings on the I–E consistency in self-concept domains invite a check with idiographic IATs. A cross-check may likewise remove inconsistencies with regard to the incremental value of self-concept IATs for predicting behaviour. Whether based on single studies or meta-analyses—estimates of I–E consistency and predictive validity are biased downward if associations are less than optimally assessed.

The quality of a measure partly depends on the mental representations elicited by the stimuli. They fill with life the concepts that people will associate during the task. Researchers need to weigh the potential for incremental utility of an idiographic over a generic approach against the additional effort spent on collecting idiographic stimuli. The idiographic approach requires flexible software tools that can use individually determined stimuli as items in the measurement procedure. From a diagnostic perspective the implications are clear; choosing the idiographic approach may require somewhat more effort, but it offers the possibility to explain extra variance that would remain unexplained by a generic approach. The CCM predicts that the present findings will extend to implicit measures other than IATs, but this prediction has not been tested yet. Therefore, at this stage, we conclude that idiographic outperform generic representations at least in self-concept IATs.

## ACKNOWLEDGEMENTS

The authors are grateful to Lisa Bossmann, Johannes Dabisch, Lisa Göbelbecker, Nadine Haag, Tillman Ihrig, Meike Meister, Anna Mokhart, Marius Prohl, Arnhild Proß, and Nathalie Seefried for their help in collecting the data of the Experiments 1–3. Wilhelm Hofmann's support in identifying the idiographic studies is clearly acknowledged as are Klaus Fiedler's and Jonathan Jong's helpful comments on earlier drafts of this manuscript. We further thank Kristin Steslow and Tristan Philip for proof-reading the manuscript. Financial support by the research pool of the University of Heidelberg granted to the first author (D.100200/08013) is gratefully acknowledged.

## REFERENCES

- Aiken, L. S., & West, S. G. (1991). *Multiple Regression: Testing and interpreting interactions*. Thousand Oaks: Sage.
- Allport, G. W. (1937). *Personality: A psychological interpretation*. New York: Holt.
- Antonovsky, A. (1979). *Health, stress and coping: New perspectives on mental and physical well-being*. San Francisco: Jossey-Bass.
- Antonovsky, A. (1993). The structure and properties of the Sense of Coherence Scale. *Social Science and Medicine*, 36, 725–733.
- Antonovsky, A. (1997). *Salutogenese. Zur Entmystifizierung der Gesundheit* (Dt. erweiterte Herausgabe von A. Franke) [Salutogenesis. De-mystifying health [extended German edition by A. Franke]]. Tübingen: DGVT.
- Asendorpf, J. B., Banse, R., & Mücke, D. (2002). Double dissociation between implicit and explicit personality self-concept: The case of shy behavior. *Journal of Personality and Social Psychology*, 83, 380–393.
- Baccus, J. R., Baldwin, M. W., & Packer, D. J. (2004). Increasing implicit self-esteem through classical conditioning. *Psychological Science*, 15, 498–502.
- Banse, R., & Fischer, I. (2002). *Implicit and explicit aggressiveness and the prediction of aggressive behavior*. Poster presented at the 11th European Conference on Personality of the European Society for Personality Psychology, Jena, Germany.
- Barden, J., Maddux, W. W., Petty, R. E., & Brewer, M. B. (2004). Contextual moderation of racial bias: The impact of social roles on controlled and automatically activated attitudes. *Journal of Personality and Social Psychology*, 87, 5–22.
- Blair, I. V. (2002). The malleability of automatic stereotypes and prejudice. *Personality and Social Psychology Review*, 6, 242–261.
- Bluemke, M., & Friese, M. (2006). Do features of stimuli influence IAT effects? *Journal of Experimental Social Psychology*, 42, 163–176.
- Bluemke, M., & Steinmayr, R. (2008). *Validation of Caprara's Irritability Scale and Gladue's Aggression Inventory in German*. Unpublished raw data.
- Bluemke, M., & Zumbach, J. (2007). Implicit and explicit Measures for analyzing the aggression of computer gamers. In G. Steffgen, & M. Gollwitzer (Eds.), *Emotions and aggressive behavior* (pp. 38–57). Göttingen: Hogrefe & Huber Publishers.
- Bluemke, M., Friedrich, M., & Zumbach, J. (2010). The influence of violent and nonviolent computer games on implicit measures of aggressiveness. *Aggressive Behavior*, 36, 1–13.
- Bosson, J. K., Swann, W. B., Jr., & Pennebaker, J. W. (2000). Stalking the perfect measure of implicit self-esteem: The blind men and the elephant revisited? *Journal of Personality and Social Psychology*, 79, 631–643.
- Buss, A. H., & Perry, M. P. (1992). The aggression questionnaire. *Journal of Personality and Social Psychology*, 63, 452–459.
- Caprara, G. V., Cinanni, V., D'Imperio, G., Passerini, S., Renzi, P., & Travaglia, G. (1985). Indicators of impulsive aggression: Present status of research on irritability and Emotional Susceptibility Scales. *Personality and Individual Differences*, 6, 665–674.
- von Collani, G., & Herzberg, P. Y. (2003). Eine revidierte Fassung der deutschsprachigen Skala zum Selbstwertgefühl von Rosenberg [A revised version of the German adaptation of Rosenberg's self-esteem scale]. *Zeitschrift für Differentielle und Diagnostische Psychologie*, 24, 3–7.
- von Collani, G., & Werner, R. (2005). Self-related and motivational constructs as determinants of aggression. An analysis and validation of a German version of the Buss-Perry Aggression Questionnaire. *Personality and Individual Differences*, 38, 1631–1643.
- Collins, A. M., & Loftus, E. F. (1975). A spreading activation theory of semantic processing. *Psychological Review*, 82, 407–428.
- Conner, T., & Barrett, L. F. (2005). Implicit self-attitudes predict spontaneous affect in everyday life. *Emotion*, 5, 476–488.
- DeHart, T., Pelham, B., Fiedorowicz, L., Carvallo, M., & Gabriel, S. (2011). Including others in the implicit self: Implicit evaluation of significant others. *Self and Identity*, 10, 127–135.
- Denzler, M., Förster, J., & Liberman, N. (2009). How goal-fulfillment decreases aggression. *Journal of Experimental Social Psychology*, 45, 90–100.
- DeWall, C. N., Baumeister, R. F., Stillman, T. F., & Gailliot, M. T. (2007). Violence restrained: Effects of self-regulatory capacity and its depletion on aggressive behavior. *Journal of Experimental Social Psychology*, 43, 62–76.
- Egloff, B., & Schmukle, S. C. (2002). Predictive validity of an Implicit Association Test for assessing anxiety. *Journal of Personality and Social Psychology*, 83, 1441–1455.
- Eriksson, M., & Lindström, B. (2006). Antonovsky's sense of coherence scale and the relation with health: a systematic review. *Journal of Epidemiology and Community Health*, 60, 376–381.
- Franck, E., De Raedt, R., Dereu, M., & Van den Abbeele, D. (2007). Implicit and explicit self-esteem in currently depressed individuals with and without suicidal ideation. *Journal of Behavior Therapy and Experimental Psychiatry*, 38, 75–85.

- Friese, M., & Hofmann, W. (2009). Control me, or I will control you: Impulses, trait self-control, and the guidance of behavior. *Journal of Research in Personality, 43*, 795–805.
- Friese, M., Hofmann, W., & Schmitt, M. (2008). When and why do implicit measures predict behavior? Empirical evidence for the moderating role of opportunity, motivation, and process reliance. *European Review of Social Psychology, 19*, 285–338.
- 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.
- Gebauer, J. E., Riketta, M., Broemer, P., & Maio, G. R. (2008). “How much do you like your name?” An implicit measure of global self-esteem. *Journal of Experimental Social Psychology, 44*, 1346–1354.
- Gladue, B. (1991). Qualitative and quantitative sex differences in self-reported aggressive behavioral characteristics. *Psychological Reports, 68*, 675–684.
- Glen, I. S., & Banse, R. (2004). Probing the malleability of implicit and explicit self-esteem: An interview approach. *Cahiers de Psychologie Cognitive/Current Psychology of Cognition, 22*, 133–158.
- Gollwitzer, M., Banse, R., Eisenbach, K., & Naumann, A. (2007). Effectiveness of the Vienna social competence training on explicit and implicit aggression: Evidence from aggressiveness-IAT. *European Journal of Psychological Assessment, 23*, 150–156.
- Govan, C. L., & Williams, K. D. (2004). Reversing or eliminating IAT effects by changing the affective valence of the stimulus items. *Journal of Experimental Social Psychology, 40*, 357–365.
- Greenwald, A. G., Banaji, M. R., Rudman, L. A., Farnham, S. D., Nosek, B. A., & Mellott, D. S. (2002). A unified theory of implicit attitudes, stereotypes, self-esteem, and self-concept. *Psychological Review, 109*, 3–25.
- 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.
- Greenwald, A. G., Poehlman, T. A., Uhlmann, E. L., & Banaji, M. R. (2009). Understanding and using the Implicit Association Test: III. Meta-analysis of predictive validity. *Journal of Personality and Social Psychology, 97*, 17–41.
- Greenwald, A. G., & Pratkanis, A. R. (1984). The self. In R. S. Wyer, & T. K. Srull (Eds.), *Handbook of social cognition* (pp. 129–178). Hillsdale, NJ: Erlbaum.
- Grover, V. P., Keel, P. K., & Mitchell, J. P. (2003). Gender differences in implicit weight identity. *International Journal of Eating Disorders, 34*, 125–135.
- Herzberg, P. Y. (2003). Faktorstruktur, Gütekriterien und Konstruktvalidität der deutschen Übersetzung des Aggressionsfragebogens von Buss und Perry [Psychometric evaluation and validity of the German translation of the Aggression Questionnaire by Buss and Perry]. *Zeitschrift für Differentielle und Diagnostische Psychologie, 24*, 311–323.
- Hofmann, W., Friese, M., & Strack, F. (2009). Impulse and self-control from a dual-systems perspective. *Perspectives on Psychological Science, 4*, 162–176.
- 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.
- Hofmann, W., Gschwendner, T., Nosek, B. A., & Schmitt, M. (2005). What moderates implicit–explicit consistency? *European Review of Social Psychology, 16*, 335–390.
- Hofmann, W., Gschwendner, T., & Schmitt, M. (2009). The road to the implicit self not taken: Discrepancies between self- and observer-inferences about implicit dispositions from nonverbal behavioral cues. *European Journal of Personality, 23*, 343–366.
- Hoorens, V., & Todorova, E. (1988). The name letter effect: Attachment to self or primacy of own name writing? *European Journal of Social Psychology, 18*, 365–368.
- Johnson, M. (2004). Approaching the salutogenesis of sense of coherence: The role of ‘active’ self-esteem and coping. *British Journal of Health Psychology, 9*, 419–432.
- Jordan, C. H., Spencer, S. J., Zanna, M. P., Hoshino-Browne, E., & Correll, J. (2003). Secure and defensive high self-esteem. *Journal of Personality and Social Psychology, 85*, 969–978.
- Jordan, C. H., Whitfield, M., & Zeigler-Hill, V. (2007). Intuition and the correspondence between implicit and explicit self-esteem. *Journal of Personality and Social Psychology, 93*, 1067–1079.
- Karpinski, A. (2004). Measuring self-esteem using the Implicit Association Test: The role of the other. *Personality and Social Psychology Bulletin, 30*, 22–34.
- Kitayama, S., & Karasawa, M. (1997). Implicit self-esteem in Japan: Name letters and birthday numbers. *Personality and Social Psychology Bulletin, 23*, 736–742.
- Koole, S. L., Dijksterhuis, A., & van Knippenberg, A. (2001). What’s in a name: Implicit self-esteem. *Journal of Personality and Social Psychology, 80*, 614–627.
- Krippendorff, K. (1970). Estimating the reliability, systematic error, and random error of interval data. *Educational and Psychological Measurement, 30*, 61–70.
- Krizan, Z. (2008). What is implicit about implicit self-esteem? *Journal of Research in Personality, 42*, 1635–1640.
- Krizan, Z., & Suls, J. (2008). Are implicit and explicit measures of self-esteem related? A meta-analysis for the Name-Letter Test. *Personality and Individual Differences, 44*, 521–531.
- Krizan, Z., & Suls, J. (2009). Implicit self-esteem in the context of trait models of personality. *Personality and Individual Differences, 46*, 659–663.
- Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (2005). *International affective picture system (IAPS): Affective ratings of pictures and instruction manual*. Technical Report A-6. University of Florida, Gainesville, FL.
- LeBel, E. P. (2010). Attitude accessibility as a moderator of implicit and explicit self-esteem correspondence. *Self and Identity, 9*, 195–208.
- Markus, H. (1977). Self-schemata and processing information about the self. *Journal of Personality and Social Psychology, 35*, 63–78.
- Markus, H., Hamill, R., & Sentis, K. P. (1987). Thinking fat: Self-schemas for body weight and the processing of weight relevant information. *Journal of Applied Social Psychology, 17*, 50–71.
- Markus, H., & Wurf, E. (1987). The dynamic self-concept: A social psychological perspective. *Annual Review of Psychology, 38*, 299–337.
- Mashek, D., Aron, A., & Boncimino, M. (2003). Including other in the self as overlap of cognitive elements: Evidence from source memory confusions. *Personality and Social Psychology Bulletin, 29*, 382–392.
- McConnell, A. R. (2011). The multiple self-aspects framework: Self-concept representation and its implications. *Personality and Social Psychology Review, 15*, 3–27.
- Miles, J. N. V., & Shevlin, M. E. (2001). *Applying regression and correlation: a guide for students and researchers*. London: Sage.
- Molenaar, P. C. M. (2004). A manifesto on psychology as idiographic science: Bringing the person back into scientific psychology, this time forever. *Measurement: Interdisciplinary Research and Perspectives, 2*, 201–218.
- Münsterberg, H. (1994). Psychology and history. *Psychological Review, 101*, 230–236. (Original work published 1899)
- Mussweiler, T., & Förster, J. (2000). The sex-aggression link: A perception-behavior dissociation. *Journal of Personality and Social Psychology, 79*, 507–520.

- Mussweiler, T., & Strack, F. (2000). The "relative self": Informational and judgmental consequences of comparative self-evaluations. *Journal of Personality and Social Psychology, 79*, 25–38.
- Nosek, B. A. (2007). Implicit-explicit relations. *Current Directions in Psychological Science, 16*, 65–69.
- Nosek, B. A., Greenwald, A. G., & Banaji, M. R. (2007). The Implicit Association Test at age 7: A methodological and conceptual review. In J. A. Bargh (Ed.), *Social psychology and the unconscious: The automaticity of higher mental processes* (pp. 265–292). New York, NY: Psychology Press.
- Nuttin, J. M. (1985). Narcissism beyond Gestalt and awareness: The name letter effect. *European Journal of Social Psychology, 15*, 353–361.
- Nyholm, M., Gullberg, B., Merlo, J., Lundqvist-Persson, C., Råstam, L., & Lindblad, U. (2007). The validity of obesity based self-reported weight and height: Implications for population studies. *Obesity, 15*, 197–208.
- Olson, M. A., Fazio, R. H., & Hermann, A. D. (2007). Reporting tendencies underlie discrepancies between implicit and explicit measures of self-esteem. *Psychological Science, 18*, 287–291.
- Perugini, M. (2005). Predictive models of implicit and explicit attitudes. *British Journal of Social Psychology, 44*, 29–45.
- Perugini, M., O’Gorman, R., & Prestwich, A. (2007). An ontological test of the IAT. Self-activation can increase predictive validity. *Experimental Psychology, 54*, 134–147.
- Perugini, M., Richetin, J., & Zogmaister, C. (2010). Prediction of behavior. In B. Gawronski, & B. K. Payne (Eds.), *Handbook of Implicit Social Cognition: Measurement, Theory, and Applications* (Chap. 14). New York: Guilford Press.
- Rand, C. S. W., Resnick, J. L., & Seldman, R. S. (1997). Assessment of socially acceptable body sizes by university students. *Obesity Research, 5*, 415–429.
- Ranganath, K. A., Smith, C. T., & Nosek, B. A. (2008). Distinguishing automatic and controlled components of attitudes from direct and indirect measurement methods. *Journal of Experimental Social Psychology, 44*, 386–396.
- Richetin, J., Richardson, D. S., & Mason, G. D. (2010). Predictive validity of IAT aggressiveness in the context of provocation. *Social Psychology, 41*, 27–34.
- Robins, R. W., Hendin, H. M., & Trzesniewski, K. H. (2001). Measuring global self-esteem: Construct validation of a single-item measure and the Rosenberg self-esteem scale. *Personality and Social Psychology Bulletin, 27*, 151–161.
- Rogers, C. R. (1959). A theory of therapy, personality and interpersonal relationships as developed in the client-centered framework. In S. Koch (Ed.), *Psychology: A study of a science. Vol. III. Formulations of the person and the social context* (pp. 184–256). New York: McGraw Hill.
- Rosch, E. (1978). Principles of categorization. In E. Rosch, & B. B. Lloyd (Eds.), *Cognition and categorization* (pp. 27–48). Hillsdale, NJ: Lawrence Erlbaum.
- Rosch, E., Mervis, C. B., Gray, W. D., Johnson, D. M., & Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive Psychology, 8*, 382–439.
- Rosenberg, M. (1979). *Conceiving the self*. New York: Basic Books.
- Rudolph, A., Schröder-Abé, M., Schütz, A., Gregg, A. P., & Sedikides, C. (2008). Through a glass, less darkly? Reassessing convergent and discriminant validity in measures of implicit self-esteem. *European Journal of Psychological Assessment, 24*, 273–281.
- Rydell, R., & Gawronski, B. (2009). I like you, I like you not: Understanding the formation of context-dependent automatic attitudes. *Cognition and Emotion, 23*, 1118–1152.
- Schütz, A., & Sellin, I. (2006). *Die Multidimensionale Selbstwertenskala (MSWS)* [The multidimensional self-esteem scale (MSSES)]. Göttingen: Hogrefe.
- Scott, W. A., Kline, J. A., Faguy-Coté, E., & Peterson, C. (1980). Centrality of cognitive attributes. *Journal of Research in Personality, 14*, 12–26.
- Sedikides, C., & Skowronski, J. J. (1993). The self in impression formation: Trait centrality and social perception. *Journal of Experimental Social Psychology, 29*, 347–357.
- Steffens, M. C., Kirschbaum, M., & Glados, P. (2008). Avoiding stimulus confounds in Implicit Association Tests by using the concepts as stimuli. *British Journal of Social Psychology, 47*, 217–243.
- Steiger, J. H. (1980). Tests for comparing elements of a correlation matrix. *Psychological Bulletin, 87*, 245–251.
- Stieger, S., Voracek, M., & Formann, A. K. (2011). How to administer the Initial Preference Task. *European Journal of Personality, 25*: n/a. doi: 10.1002/per.823
- Strack, F., & Deutsch, R. (2004). Reflective and impulsive determinants of social behavior. *Personality and Social Psychology Review, 8*, 220–247.
- Stucke, T., & Baumeister, R. F. (2006). Ego depletion and aggressive behavior: Is the inhibition of aggression a limited resource? *European Journal of Social Psychology, 36*, 1–13.
- Stunkard, A. J., Sorensen, T. I., & Schulsinger, F. (1983). Use of the Danish adoption Register for the study of obesity and thinness. In S. S. Kety, L. P. Rowland, R. L. Sidman, & S. W. Matthysse (Eds.), *The Genetics of Neurological and Psychiatric Disorders* (pp. 115–120). New York: Raven Press.
- Symons, C. S., & Johnson, B. T. (1997). The self-reference effect in memory: A meta-analysis. *Psychological Bulletin, 121*, 371–394.
- Tuerlinckx, F. (2004). The idiographic approach: Where do we come from and where do we go. *Measurement: Interdisciplinary Research and Perspectives, 2*, 240–243.
- Uhlmann, E., & Swanson, J. (2004). Exposure to violent video games increases automatic aggressiveness. *Journal of Adolescence, 27*, 41–52.
- Wanous, J. P., & Reichers, A. E. (1996). Estimating the reliability of a single-item measure. *Psychological Reports, 78*, 631–634.
- Wanous, J. P., Reichers, A. E., & Hudy, M. J. (1997). Overall job satisfaction: How good are single-item measures? *Journal of Applied Psychology, 82*, 247–252.
- Webb, E., Campbell, D., Schwartz, R., & Sechrest, L. (1966). *Unobtrusive measures: Nonreactive research in the social sciences*. Chicago: Rand McNally.
- Wentura, D., Kulfanek, M., & Greve, W. (2005). Masked affective priming by name letters: Evidence for a correspondence of explicit and implicit self-esteem. *Journal of Experimental Social Psychology, 41*, 654–663.
- Whitfield, M., & Jordan, C. H. (2009). Mutual influence of implicit and explicit attitudes. *Journal of Experimental Social Psychology, 45*, 748–759.
- Williamson, D. A., Gleaves, D. H., Watkins, P. C., & Schlundt, D. G. (1993). Validation of self-ideal body size discrepancy as a measure of body dissatisfaction. *Journal of Psychopathology and Behavior Assessment, 15*, 57–68.
- Windelband, W. (1998). History and natural science. *Theory and Psychology, 8*, 5–22. (Original work published 1894)
- Woodman, T., & Steer, R. (2011). Body self-discrepancies and women’s social physique anxiety: The moderating role of the feared body. *British Journal of Psychology, 102*, 147–160.
- Yamaguchi, S., Greenwald, A. G., Banaji, M. R., Murakami, F., Chen, D., Shiomura, K., . . . Krendl, A. (2007). Apparent universality of positive implicit self-esteem. *Psychological Science, 18*, 498–500.

## APPENDIX

GERMAN STIMULI AND ENGLISH COUNTERPARTS EMPLOYED IN IMPLICIT ASSOCIATION TESTS  
(EXPERIMENTS 1–3)

| <u>Experiment 1: Implicit self-esteem: Attribute stimuli (adjectives)</u> |                 |                 |                           |                           |
|---|-----------------|-----------------|---------------------------|---------------------------|
| <u>Attributes from MSES-scale</u>   | <u>Positive</u> | <u>Negative</u> | <u>Positive (English)</u> | <u>Negative (English)</u> |
| MSES: Emotional self-esteem   | Gut             | Schlecht        | Good                      | Bad                       |
|   | Wertvoll        | Wertlos         | Valuable                  | Worthless                 |
|   | Zufrieden       | Unzufrieden     | Content                   | Dissatisfied              |
| MSES: Social contact  | Offen           | Gehemmt         | Open-minded               | Inhibited                 |
|   | Gesellig        | Schüchtern      | Sociable                  | Shy                       |
|   | Kommunikativ    | Schweigsam      | Talkative                 | Quiet                     |
| MSES: Social criticism  | Beliebt         | Unbeliebt       | Well-liked                | Unpopular                 |
|   | Unabhängig      | Abhängig        | Independent               | Dependent                 |
|   | Verträglich     | Stur            | Agreeable                 | Stubborn                  |
| MSES: Performance   | Erfolgreich     | Erfolglos       | Successful                | Ineffective               |
|   | Fähig           | Unfähig         | Competent                 | Inapt                     |
|   | Belastbar       | Überfordert     | Resilient                 | Overstrained              |
| MSES: Physical attractiveness   | Schön           | Hässlich        | Beautiful                 | Ugly                      |
|   | Schlank         | Dick            | Slim                      | Fat                       |
|   | Attraktiv       | Dürr            | Attractive                | Meager                    |
| MSES: Athleticism   | Sportlich       | Unsportlich     | Sportive                  | Unathletic                |
|   | Athletisch      | Träge           | Athletic                  | Inactive                  |
|   | Fit             | Untrainiert     | [Physically] fit          | Unfit                     |

| <u>Experiment 2: Implicit weight identity: Attribute stimuli (adjectives)</u> |              |               |                       |                        |
|---|--------------|---------------|-----------------------|------------------------|
| <u>Attribute stimuli</u>  | <u>Thin</u>  | <u>Thick</u>  | <u>Thin (English)</u> | <u>Thick (English)</u> |
|   | Dürr         | Fett          | Skinny                | Fat                    |
|   | Mager        | Üppig         | Slender               | Busty                  |
|   | Hager        | Breit         | Scraggy               | Robust                 |
|   | Leicht       | Schwer        | Light                 | Heavy                  |
|   | Schmal       | Mollig        | Small                 | Plump                  |
|   | Schlank      | Füllig        | Slim                  | Full-figured           |
|   | Knochtig     | Beleibt       | Bony                  | Stout                  |
|   | Zierlich     | Pummelig      | Petite                | Chubby                 |
|   | Schmächtig   | Korpulent     | Tender                | Corpulent              |
|   | Ausgehungert | Übergewichtig | Famished              | Obese                  |

| <u>Experiment 3: Implicit aggressiveness: Attribute stimuli (verbs)</u> |                   |                 |                             |                           |
|---|-------------------|-----------------|-----------------------------|---------------------------|
| <u>BPAQ-scale</u>   | <u>Aggressive</u> | <u>Peaceful</u> | <u>Aggressive (English)</u> | <u>Peaceful (English)</u> |
| BPAQ: Physical aggression   | Schlagen          | Streicheln      | Hit                         | Caress (so.)              |
|   | Prügeln           | Kuscheln        | Club                        | Cuddle                    |
| BPAQ: Verbal aggression   | Anschreien        | Lachen          | Scream at (so.)             | Laugh                     |
|   | Lästern           | Reden           | Gossip                      | Talk                      |
| BPAQ: Hostility   | Drohen            | Lieben          | Threaten                    | Love                      |
|   | Quälen            | Vertrauen       | Tantalise (so.)             | Trust                     |
| BPAQ: Anger   | Streiten          | Beruhigen       | Quarrel                     | Settle (sth.)             |
|   | Zürnen            | Abregen         | Angry with (so.)            | Calm down                 |