Ego depletion moderates the influence of automatic and controlled precursors of reactive aggression: A double dissociation

(this preprint has been accepted for publication in *Social Psychology* and does not represent the final published version)

Alexander F. Schmidt[ab]

Philipp S. Zimmermann[a]

Rainer Banse[a]

Roland Imhoff[c]

[a]University of Bonn [b]University of Luxembourg [c]University of Cologne

Alexander F. Schmidt, Department of Psychology, Social and Legal Psychology, University of Bonn, Bonn, Germany and Institute for Health and Behavior, Health Promotion and Aggression Prevention, University of Luxembourg, Walferdange, Luxembourg. Philip Zimmermann and Rainer Banse, Department of Psychology, Social and Legal Psychology, University of Bonn, Bonn, Germany. Roland Imhoff, Department of Psychology, Social Cognition, University of Cologne, Cologne, Germany.

The authors gratefully acknowledge the helpful comments of Sarah Teige-Mocigemba and Matthias Blümke on earlier drafts of this manuscript and thank Amanda C. Jones for her assistance in manuscript preparation.

Correspondence concerning this article should be addressed to Alexander F. Schmidt, University of Luxembourg, INSIDE Research Unit, Route de Diekirch, L-7201 Walferdange, Luxembourg; E-mail: alexander.schmidt@uni.lu

Abstract

From a dual-systems perspective, it has been proposed that predictive validity of whether individuals act out or stifle their reactive aggression will be maximized if (a) automatic and (b) controlled precursors of aggression are assessed and (c) situational boundaries in favor of acting out or restraining oneself are specified. In the present research we experimentally manipulated participants' self-regulatory resources in an ego depletion paradigm and subsequently measured reactive aggression in the Taylor Aggression Paradigm. Assessing automatic and controlled precursors of reactive aggression via an Implicit Association Test of Aggressiveness (Agg-IAT) and a self-report reactive aggressiveness questionnaire, respectively, we demonstrated a theoretically expected double dissociation: Reactive aggression of ego depleted individuals was associated with the indirect measure (Agg-IAT) whereas non-depleted participants' reactive aggression correlated with the direct measure (self-report). The results corroborate the usefulness of both direct and indirect measures of aggressiveness and point to boundary conditions of these measures' criterion validity.

Keywords: Reactive aggression, implicit measures, Implicit Association Test (IAT), ego depletion, self-control

Ego depletion moderates the influence of automatic and controlled precursors of reactive aggression: A double dissociation

Aggression is among the most pressing social problems. Humans are equipped with a substantial personal potential for aggressive behavior and daily life offers a multitude of relevant situational catalysts. The root causes of aggression have been extensively researched over the last decades and interpersonal provocation has been specifically regarded as "the most important single cause of human aggression" (Anderson & Bushman, 2002, p. 37). Although a third party's interference with achieving personal goals may result in a rather immediate impulse to aggress against the other person, individuals are typically able to inhibit their aggressive impulses as social norms impede immediate reactive aggression (Ramirez, 1993). In the present research we sought to extend the individual difference perspective on aggressiveness to a dual-systems framework of reactive aggression (Hofmann, Friese, & Strack, 2009; Richetin & Richardson, 2008) as experimental research on more complex models of the interplay of dispositional aggressiveness and situational boundary conditions is yet lacking.

The *Reflective-Impulsive Model* (RIM; Strack & Deutsch, 2004) is concerned with two complementing information processing systems that interact in the determination of behavioral outcomes. According to the RIM, both the impulsive and the reflective system operate in parallel and activate behavioral schemata independently (i.e., to act aggressively or non-aggressively). The *impulsive system* influences behavior through spreading activation (i.e., automatic precursors) whereas in the *reflective system* behavior is a consequence of deliberate decision processes (i.e., controlled precursors). Thus, whether individuals act out or stifle their aggressiveness depends not only on individual differences in controlled (e.g., Bettencourt, Tally, Benjamin, & Valentine, 2006) but also automatic (e.g., Richetin, Richardson, & Mason, 2010) precursors of reactive aggression. Both systems interact either in synergistic or antagonistic ways. This means that several

behavioral options can be activated by both information processing systems concurrently. Which behavioral schema will be finally executed depends on the strength of activation that has been triggered by either the reflective or the impulsive system for each behavioral option (Strack & Deutsch, 2004). As long as both automatic and controlled precursors of reactive aggression synergistically activate the same behavioral options (i.e., to act out or stifle aggression), the prediction of behavioral outcomes is straightforward. However, in cases of antagonistic directedness of behavioral options, when one system inhibits whereas the other disinhibits reactive aggression, behavior execution will depend not only on automatic and controlled precursors but also on their interaction with situational boundary conditions that impact either the strengthening of automatic or controlled precursors (Hofmann et al., 2009; Perugini, Richetin, & Zogmaister, 2010; Richetin & Richardson, 2008).

One candidate for such a moderating situational factor that has previously been associated with an increase in aggressive behavior is the amount of self-control efforts that are maintained in a certain situation (DeWall, Baumeister, Stillman, & Gailliot, 2007; Stucke & Baumeister, 2006; Vohs, Glass, Maddox, & Markman, 2011) or as a disposition (Hofmann, Gschwendner, Friese, Wiers, & Schmitt, 2008). A prominent example of a situational factor theorized to reduce self-control effort is ego depletion (ED), the phenomenon that self-control decreases after initial exertion similar to how a muscle's strength deteriorates after making use of it (Baumeister, Vohs, & Tice, 2007, but see Inzlicht & Schmeichel, 2012, for a more recent motivational account). Specifically, ED manipulations have shown theoretically consistent effects on (a) self-reported hypothetical reactive aggression in vignette scenarios (DeWall et al., 2007; Experiment 4), (b) negative evaluations of a provoking experimenter (Stucke & Baumeister, 2006; Experiments 2, 3) or an ostensibly provoking other participant (DeWall et al., 2007, Experiment 3), and (c) administering aversive stimuli such as doses of hot sauce (DeWall

et al., 2007; Experiment 1) or blasts of white noise (DeWall, 2007; Experiment 2; Vohs et al., 2011; Experiment 1) to a confederate. Especially the latter two experiments can be regarded as convincing support for the acute behavioral impact of ED on aggressive behavior.

However, as argued above, individuals may differ in the extent to which automatic and controlled precursors command them to lash out aggressively. A lack of acutely available self-control resources or motivation to control oneself might selectively impair the impact of the reflective system on the activation of behavioral schemata. Concurrently, this specific lack should strengthen the impact of the impulsive system on the activation of behavioral schemata. On the other hand, in case of a situation where self-regulatory resources are available or a strong motivation to control oneself exists, the behavioral impact of the reflective system should be strengthened and the contribution of the impulsive system should be decreased.

Measuring automatic and controlled precursors of aggression. It is commonly argued that indirect measures (e.g., De Houwer, Teige-Mocigemba, Spruyt, & Moors, 2009; Nosek, Hawkins, & Frazier, 2011) represent the best-fitting measurement approach to tap into such automatic precursors processed in the impulsive system. Most of these measures are operationalized by the latencies of immediate responses towards relevant classes of stimuli (e.g., simple detection, classification, sorting tasks), thus minimizing potential influences from the reflective system of behavioral activation (Hofmann et al., 2009; Richetin & Richardson, 2008). In case of aggressiveness automatic precursors could be represented, for example, by an implicit association between aggressive behavior and the self. Compared to other prominent implicit measures, the Implicit Association Test (IAT; Greenwald, McGhee, & Schwartz, 1998; for individual differences in aggressiveness see Banse & Fischer, 2002; Banse, Messer, & Fischer, in press; Grumm,

Hein, & Fingerle, 2011; Richetin et al., 2010) represents an optimal indirect measure as it has shown superior psychometric properties (Bar-Anan & Nosek, 2014).

On the other hand, direct self-report measures that rely on verbal responses to specific item lists generated through effortful decisions are usually seen as tapping into more controlled processes (Nosek et al., 2011). According to this notion, self-reported standards and inclinations relating to the expression of aggressive behavior can be regarded as controlled precursors of reactive aggression and are a valid way to tap into these (Bettencourt et al., 2006).

Present Study

In the present research we were interested in examining how automatic and controlled precursors' predictive validity would be moderated by experimentally manipulated situational boundary conditions. First, we wanted to replicate the causal impact of ED on reactive aggression that has been reported in prior studies (DeWall et al., 2007; Stucke & Baumeister, 2006; Vohs et al., 2011). From a methodological perspective, the only two experimental ED studies using behavioral measures of aggression have been based on single trial assessments of either spontaneous (i.e., unprovoked) aggression (DeWall, 2007; Experiment 2) or reactive aggression (Vohs et al., 2011; Experiment 1). Only in an ancillary analysis, Vohs et al. (2011) reported an ED effect on aggregated reactive aggression levels. Therefore, we sought to increase the reliability of the dependent variable by utilizing a variant of the Taylor Aggression Paradigm (TAP, Taylor, 1967) with an increased number of trials.

Second, by expanding the individual difference perspective to a dual-systems framework, we were interested in the moderating role of ED for the criterion validity of automatic and controlled precursors of reactive aggression as theoretically postulated by Richetin and Richardson (2008). Crucially, we did *not* predict a crossover interaction pattern (which would imply less aggressive behavior for participants with high levels of

automatic aggressiveness when self-control resources are sufficient), but focused our hypotheses specifically on the pattern of simple slopes in the moderated hierarchical regression analyses. We hypothesized a double dissociation pattern (Perugini et al., 2010) of, on the one hand, automatic precursors being associated with reactive aggression in the high-ED condition but not in the low-ED condition. On the other hand, controlled precursors were expected to be associated with reactive aggression in the low- but not in the high-ED condition. A similar double dissociation pattern of reactive aggression has been reported by Hofmann et al. (2008; Study 3) but limited to *self-reported* behavior (i.e., negative feedback ratings after provocation) and a correlational design leaving potential alternative explanations of confounding third variables open. As of yet, methodologically more conclusive experimental demonstrations of this double dissociation only exist for the domain of food and alcohol consumption (Friese, Hofmann, & Wänke, 2008) but are missing for aggressive behavior.

Method

Participants

Participants were 61 psychology students volunteering to take part in the experiment for partial course credit and the possibility of winning 10 or 20 Euros as compensation (see Procedure section for a detailed description). Three participants were identified as multivariate outliers utilizing influence plots based on Cook's distances indicating the joint influence of studentized residuals and leverage on the central regression analysis (i.e., participants with the three highest Cook's coefficients were excluded from the sample)[1]. Thus, the remaining sample (N = 58; mean age M = 24.9 years; SD = 5.5) consisted of 8 men and 50 women. Age was independent from sex; $t < 1$. To control for possibly confounding sex differences in aggressiveness (Archer, 2004), group allocation for each sex was randomized separately.

Measures

Controlled precursors of aggression. Controlled precursors of reactive aggression were measured with the 13-item Reactive Aggressiveness (RA) subscale of a standard German aggressiveness questionnaire (*Fragebogen zur Erfassung von Aggressivitätsfaktoren* [Inventory for the assessment of aggressiveness factors]; FAF; Hampel & Selg, 1998), which is based on Buss and Durkee's (1957) Hostility Inventory. Responses were given in a forced-choice (true, false) answer-format. The RA subscale has been shown to differentiate between ice-hockey (increased aggressiveness) and volleyball players and to be positively related to the same laboratory-induced reactive aggression paradigm that was used in the present experiment (Banse et al., in press). Because the original 13-item RA subscale had low internal consistency in this sample ($\alpha = .52$) we removed five items with the lowest part whole corrected item-total correlations (r_{it}s < .20) from it via α-maximization in order to increase its reliability. Subsequently, the modified RA scale reached at least acceptable internal consistency ($\alpha = .60$). Construct validity has not been threatened by this elimination as evidenced by the content of the remaining items (see Appendix for the item sets of the original and modified scales) and the correlation of the original and the revised scale ($r = .87$, $p < .001$)[2].

Automatic precursors of aggression. Automatic precursors were assessed via a modified Aggressive Behavior IAT (Agg-IAT; Banse et al., in press) that has been shown to correlate with self-reported and peer-observed aggressiveness as well as with the same modified TAP that was utilized as a measure of reactive aggression in this experiment (Banse et al., in press). The Agg-IAT is based on a double-discrimination task of items belonging to the target dimension *me vs. others* and the attribute dimension *aggressive vs. peaceful.* Its set-up follows the classic five-block structure as described by Greenwald et al. (1998). In the first two blocks participants were asked to differentiate stimuli on either the *aggressive vs. peaceful* or the *me vs. other* dimension. In the critical third block, both discrimination tasks were combined. Stimuli related to *me/peaceful* and *others/aggressive* were assigned to the

same key. In the fourth block, key assignment of the attribute dimension was reversed. In the critical fifth block, both tasks were again combined assigning the same response key to stimuli related to *me/aggressive* and *others/peaceful*. The complete Agg-IAT protocol is presented in Table 1.

The items of the *me vs. others* dimension consisted of German words related to the self (*me*) and occupational labels (*others*). The *others* items were conceptualized with concrete occupational labels instead of the more commonly used descriptive labels of "them" or "others" because it is not clear who these "others" are and participants might use particularly aggressive, non-aggressive, or similarly aggressive comparison exemplars. This, however, is unlikely to happen for specific professions that were pretested to represent professions occupied by people with stereotypically moderate levels of aggressiveness (Banse et al., in press). As we wanted to assess female and male participants and the German language offers gendered occupation labels, half of the *others* items indicated male and the other half female occupational labels. For the *aggressive vs. peaceful* dimension ten words that are related to peaceful or aggressive concepts and interaction behaviors were used (Table 1).

The answer category labels were presented at the left and right side on the top of the screen and the stimuli appeared further below horizontally centered in the middle of the screen. Response errors were marked with a red "X" until the correct response key was pressed. Participants were asked to respond as quickly and accurately as possible. In order to avoid confounding person effects with procedural effects of the IAT, the order of blocks was kept constant for all participants. Error trials were discarded from the analysis. Scores for the Agg-IAT were calculated as the untransformed mean response latency difference in both critical blocks divided by the pooled *SD* of all response latencies in these blocks (D_2-measure; Greenwald, Nosek, & Banaji, 2003). Higher IAT scores indicate a relatively closer association between *me* and *aggressive*/*others* and *peaceful* as compared to *me* and *peaceful*/*others* and *aggressive*. They can be interpreted as an implicit measure of an

aggressiveness self-concept. Internal consistency of the Agg-IAT as based on the D_2-scores from the first and the second half of trials of the critical blocks was $\alpha = .68$.

Dependent variable. The focal dependent variable was measured by a modified version of the TAP (Taylor, 1967). The TAP constitutes a well-established behavioral laboratory measure of reactive aggression (Giancola & Parrott, 2008) that is related to self-reported physical aggression (e.g., Giancola & Zeichner, 1995). In the TAP, participants allegedly competed in 50 trials against an ostensible second player in a simple reaction time task (i.e., pressing a key as fast as possible after the presentation of a target on screen) with a predetermined order of a fixed set of interstimulus-intervals varying between 1.950 and 6.000ms). In our modified version of the TAP participants were led to believe that whoever reacted more slowly received a blast of white noise via headphones (Bond & Lader, 1986). Prior to each trial, participants could choose volume levels (on a scale from 1 to 8) for the noise presented to their opponent in case the opponent lost. The constant predetermined order for all participants consisted of two escalation phases during each of which noise levels steadily increased. Both escalation phases were interspersed by seven lose-trials during which the minimum noise level was presented (for a detailed description see Banse et al., in press). Participants lost in half of the trials. Maximum noise levels were restricted to 90dB (minimum 68dB) preventing hearing damage during exposure to short impulse intervals (2.000ms; Acton, 1967). Selected noise volume levels for all trials except from the first (i.e., unprovoked trial) were averaged as an indicator of reactive aggressiveness (Giancola & Parrott, 2008) with high internal consistency ($\alpha = .99$).

Procedure

Participants were recruited via bulletin boards in the psychology department. They were asked to arrive five minutes prior to the starting time of the experiment at a special meeting point outside the laboratory. Upon arrival at the meeting point they met the

experimenter and a confederate posing as the ostensible other participant. Confederates were always of the same sex as the participants in order to avoid differential provocation effects due to mixed gender participant-opponent combinations (Bettencourt & Miller, 1996). Participants were told that the study dealt with achievement behavior under stress conditions to disguise the true nature of the experiment. In order to further motivate the participants, they were informed that the participant with the fastest total reaction time would be awarded 20 Euros and the three participants with the three following ranks would be awarded 10 Euros each (cash prizes were actually awarded via a raffle). Also, they were instructed that the first part of the study (the ED manipulation described below) was to be taken by every participant for himself whereas the second part consisted of a competitive reaction time game that would be played against the other participant (confederate) over a network connection between two computers in adjacent rooms in the laboratory. Subsequently, after providing their informed consent the participant and confederate were seated separately in two adjacent rooms and the participant completed the ED manipulation. The procedure has been approved by the departmental ethics review board.

Experimental manipulation. To manipulate the degree of ED, we used a modified Stroop task, which has been validated for this purpose (Imhoff, Schmidt, & Gerstenberg, 2014; Study 1). All participants were asked to press correspondingly colored keys to indicate the color of the text in which presented color word stimuli were written. In the low-ED condition, the text color always corresponded to the color words (i.e., there was no interference between semantic meaning and color perception). In the high-ED condition, the color words never matched the text color, thus the automatic response to process the semantic meaning of the word had to be inhibited. Additionally, if the color word was presented in blue (25% of the trials), participants were asked to press the key corresponding to the meaning of the word, thus, preventing them from strategically ignoring the meaning of the words. After

completing 12 practice trials, participants received feedback regarding their accuracy. If participants answered more than 25% of the items incorrectly, they had to repeat the practice phase until they scored below the error criterion. Subsequently, participants completed 180 test trials in both conditions.

After finishing the ED manipulation and the TAP, participants were asked to complete the Agg-IAT followed by the RA subscale of the FAF questionnaire. This order was chosen to prevent participants from suspecting the aggression-related intent of the TAP. Afterwards, participants were thanked and asked to leave an email address for the debriefing and to be notified after the whole study was completed in case they won monetary compensation.

Results

Age, participant sex, RA, and error rates of the Agg-IAT were independent from the experimental manipulation ($|rs| < .13$, $ps > .30$, Table 2). As hypothesized, individuals showed higher TAP reactive aggression levels in the high-ED condition as compared to the low-ED condition ($r = .36$, $p = .006$)[3]. Unexpectedly, the high-ED group showed significantly lower Agg-IAT scores ($r = -.47$, $p < .001$).

To test the predicted moderation patterns, we conducted a hierarchical multiple regression analysis (Cohen, Cohen, Aiken, & West, 2003; Table 3). In order to control for possible artifacts due to the small sample size we bootstrapped (1,000 samples, bias-corrected 95%-confidence intervals) all regression analyses reported in the following. Regressing TAP mean volume levels after provocation on the effect-coded ED manipulation (-1 for low-ED, +1 for high-ED conditions) as well as Agg-IAT score and self-reported RA (both *z*-standardized), the ED manipulation and the Agg-IAT significantly predicted higher TAP reactive aggression scores (Table 3; Step 1). In order to test the moderation hypothesis, all three two-way interaction terms between the three variables ED, RA, and Agg-IAT were entered into the regression in the second step. The predicted double dissociation was confirmed as the interaction of ED with automatic (Agg-IAT) was significant ($p = .004$) and

the interaction of controlled aggression precursors (RA) with ED was marginally significant (p = .08; Table 3; Step 2), whereas neither the interaction of RA and the Agg-IAT nor the higher order three-way interaction of ED, RA, and the Agg-IAT in the third step reached significance (all ps > .45; Table 3)[4].

As the critical test of the predicted double dissociation pattern, corresponding interaction plots (Figure 1) revealed that the Agg-IAT showed stronger effects on TAP reactive aggression in the high-ED as compared to the low-ED condition. Bootstrapped (1,000 samples, bias-corrected 95% confidence intervals) simple slope analyses (± 1 *SD*) corroborated our hypothesis that the Agg-IAT predicted TAP reactive aggression for highly ego depleted individuals, b = .91, SE = .28, p = .005, but not for participants in the low-ED condition, b = -.06, SE = .15, p = .63. Conversely, RA exclusively exerted more pronounced aggression effects in the low-ED as compared to the high-ED condition. This was supported by corresponding simple slope analyses revealing that the self-report RA subscale predicted reactive aggression for less ego depleted participants, b = .65, SE = .29, p = .02, but not for highly ego depleted individuals, b = -.05, SE = .24, p = .85 (Figure 1)[5].

Discussion

One of our research aims was to elucidate the role of self-control resources as an important boundary condition impacting the regulation of reactive aggression. We replicated previously reported ED main effects on aggressive behavior (DeWall et al., 2007; Stucke & Baumeister, 2006, Vohs et al., 2011) showing that individuals deprived of their self-regulatory resources administered louder blasts of white noise to their ostensible opponents as compared to less-depleted control participants.

Double dissociation of reactive aggression

As a new contribution and in line with the dual-systems framework (Hofmann et al., 2009; Perugini et al., 2010; Richetin & Richardson, 2008) our findings corroborated the hypothesized double dissociation that a direct self-report measure of dispositional reactive

aggressiveness (RA) was associated with reactive aggression only when self-regulatory resources were not impeded, whereas for an indirect measure of aggressiveness (Agg-IAT) this was the case only under the condition of impeded self-control. Notably, adding the moderators to the regression model explained additional two-thirds of the variance in reactive aggression as compared to the simple main effects alone (Table 3) corroborating that more sophisticated models that take self-regulatory boundary conditions into consideration indeed increase incremental validity.

The reported double dissociation is concordant with recent motivational (Inzlicht & Schmeichel, 2012) and neurocognitive findings of ED not only increasing brain activity in areas coding reward value and desirability but also decreasing functional connectivity between reward-sensitive areas and inhibitive structures associated with cognitive self-control (Wagner, Altman, Boswell, Kelley, & Heatherton, 2013). These processes are a possible explanation for the idea that ED strengthens the relative impact of the impulsive as compared to the reflective system and might be particularly true for reactive aggression induced by the TAP that functions as a prominent primer of automatic aggressive behavioral scripts (Berkowitz, 2008).

Limitations and strengths

It may be seen as a methodological limitation that individual differences in aggressiveness were measured subsequent to the experimental manipulation and the dependent variable because this procedural order might have systematically influenced the assessment of automatic and controlled precursors of aggression. However, measuring precursors prior to all the other procedural steps would have rendered the research hypotheses unnecessarily salient and the genuine purpose of the TAP as an experimental aggression measure would have been given away too easily. Nevertheless, it is an unexpected finding that the high-ED (vs. low-ED) group appeared less aggressive at the automatic level as assessed via the Agg-IAT. We have randomized group allocation, and so systematic a priori

differences should be ruled out. Yet, there exists no strong rationale as to why an ED manipulation should selectively impact on automatic precursors. It might be hypothesized that the Agg-IAT group differences indicate systematic changes in interference inclination as a function of the ED manipulation. However, such a confound due to weakened test motivation or executive control seems unlikely as error rates between groups did not differ (r = -.08, p = .553, Table 2). Moreover, this would not necessarily prevent meaningful interindividual differences in IAT scores or the validity of interindividual IAT scores within each ED group despite a mean shift across all participants if groups are collapsed. Independent of this hypothetical explanation, it is more crucial to discuss whether this unexpected effect threatens the validity of our findings. We would argue that it does not: The fact that participants showed particularly less aggressive self-concepts at the automatic level in the condition where we expected automatic precursors to be the central predictor of increased aggression should have worked against our hypothesized pattern (i.e., a floor effect). Thus, the finding of our predicted pattern of effects can indeed be interpreted as a rather conservative test of our double dissociation hypothesis.

It is a liability to this study that the results reported here rest on a small sample size. Over recent years, an ongoing debate on sample sizes and replicability in social psychology has led many researchers to the insight that there is a need for better powered studies. This is a highly desirable move of the field as a whole but leaves the question unanswered what should happen with data that were already collected. We strongly believe that it would not do justice to the field as well as the participants to merely hide these studies in the file drawer. Instead we have employed analyses that reduce the unduly influence of a few outliers within a small sample. Specifically, we have controlled for multivariate outliers utilizing robust non-parametric bootstrapping algorithms to reduce distortion by single influential cases who are largely different from the rest of the sample. Moreover, we believe that the fairly complex double dissociation pattern is strongly embedded in a well-outlined *a priori* theoretical

framework (Richetin & Richardson, 2008; Hofmann et al., 2008) that has been demonstrated in other behavioral domains (Friese, Hofmann, & Wänke, 2008; Hofmann et al., 2009). This framework allowed for highly specific hypotheses safeguarding against capitalizing on pure chance effects due to the removal of single outliers and might bolster the confidence one can have in the results from this single study experiment.

On the other hand, it is a particular strength of this experiment that we have used a laboratory behavioral measure of reactive aggression instead of much more subtle and remote forms of self-reported aggression in prior research or behavioral measures based on psychometrically less favorable single trial measurements (DeWall et al., 2007, Experiment 2; Vohs et al., 2011, Experiment 1).

Conclusions and outlook

The popularity of indirect measures has been attributed to the fact that these approaches benefit from being (a) inherently less transparent than self-report measures (due to the indirect character of the measurement procedure) and (b) able to tap into automatic precursors (because of the factors contributing to the implicitness of the measurement outcomes; De Houwer et al. 2009). From this reasoning it is often concluded that indirect measures reflect the “true value” or the “true self” in the sense of individuals’ genuine motives that are usually concealed from self-reports. However, it remains unclear what the “true self” is supposed to be (Gawronski, 2009). On the one hand, it might be assumed that the “true self” is revealed under circumstances of failing deliberate control. On the other hand, it might be hypothesized that the “true self” can be inferred from what a person deliberately does and explicitly chooses in a controlled mode. This conundrum can be solved by incorporating a dual-systems perspective that claims differential predictive validities for direct and indirect measures tapping into controlled and automatic behavioral precursors, respectively (Perugini et al., 2010). From this perspective, both forms of the self are equally “true” but differentially relevant under distinct situational conditions. Therefore, for diagnostic purposes the pressing

question remains, under what boundary conditions will implicit measures be incrementally valid predictors of specific behaviors above and beyond explicit measures? This question has been elucidated with the double dissociation pattern in this experiment: Although provocation and depletion of self-regulatory resources make aggression more likely in general, individual differences do indeed matter as not everyone will aggress in situations of limited self-control.

Moreover, this conception may inform current cognitive-behavioral intervention practices (e.g., Hollin & Bloxsom, 2007), which strongly focus on behavioral regulation via the reflective system (i.e., cognitive deliberation) neglecting the self-regulatory effects of automatic precursors under the detrimental effects of acute stress and affective arousal (Van Gelder, 2013) that have to be faced in everyday situations outside of intervention settings. Future research on treating aggression problems might, thus, be informed by recent advances in the field of automatic cognitive bias modification (Wiers, Gladwin, Hofmann, Salemink, & Ridderinkhof, 2013) in order to develop more successful intervention paradigms including therapeutic modules that might help buffer the impact of or even modify automatic (pro)-aggression precursors in situations of depleted self-control resources – a combination of risk factors that led to the highest levels of reactive aggression in this experiment.

References

Acton, W. I. (1967). A review of hearing damage risk criteria. *Annals of Occupational Hygiene, 10,* 143-153. doi: 10.1093/annhyg/10.2.143

Anderson, C. A., & Bushman, B. J. (2002). Human aggression. *Annual Review of Psychology, 53,* 27-51. doi: 10.1146/annurev.psych.53.100901.135231

Archer, J. (2004). Sex differences in real-world settings: A meta-analytic review. *Review of General Psychology, 8,* 291-322. doi: 10.1037/1089-2680.8.4.291

Banse, R., Messer, M., & Fischer, I. (in press). Indirect and direct measures of aggressiveness, social desirability, and the prediction of aggressive behavior. *Aggressive Behavior.*

Banse, R. & Fischer, I. (2002, July). *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.

Bar-Anan, Y., & Nosek, B. A. (2014). A comparative investigation of seven indirect measures of social cognition. *Behavior Research Methods, 46,* 668-688. doi: 10.3758/s13428-013-0410-6

Baumeister, R. F., Vohs, K. D., & Tice, D. M. (2007). The strength model of self-control. *Current Directions in Psychological Science, 16,* 351-355. doi: 10.1111/j.1467-8721.2007.00534.x

Buss, A. H., & Durkee, A. (1957). An inventory for assessing different kinds of hostility. *Journal of Consulting Psychology, 21,* 343-349. doi: 10.1037/h0046900

Berkowitz, L. (2008). On the consideration of automatic as well as controlled psychological processes in aggression. *Aggressive Behavior, 34,* 117-129. doi: 10.1002/ab.20244

Bettencourt, B. A., & Miller, N. (1996). Gender differences in aggression as a function of provocation: A meta-analysis. *Psychological Bulletin, 119,* 422-447. doi:10.1037/0033-2909.119.3.422

Bettencourt, B. A., Talley, A., Benjamin, A. J., Valentine, J. (2006). Personality and aggressive behavior under provoking and neutral conditions: A meta-analytic review. *Psychological Bulletin, 132,* 751-777. doi: 10.1037/0033-2909.132.5.751

Bond, A., & Lader, M. (1986). A method to elicit aggressive feelings and behaviour via provocation. *Biological Psychology, 22,* 69-79. doi: 10.1016/0301-0511(86)90021-9

Buss, A. H., & Durkee, A. (1957). An inventory for assessing different kinds of hostility. *Journal of Consulting Psychology, 21,* 343-349. doi: 10.1037/h0046900

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences*. Mahwah: Lawrence Erlbaum.

De Houwer, J., Teige-Mocigemba, S., Spruyt, A., & Moors, A. (2009). Implicit measures: A normative review. *Psychological Bulletin, 135,* 347-368. doi: 10.1037/a0014211

DeWall, C. N., Baumeister, R. F., Stillman, T. F., & Gailliot, M. T. (2007). Violence restrained: Effects of self-regulation and its depletion on aggression. *Journal of Experimental Social Psychology, 43,* 62-76. doi: 10.1016/j.jesp.2005.12.005

Friese, M., Hofmann, W., & Wänke, M. (2008). When impulses take over: Moderated predictive validity of implicit and explicit attitude measures in the prediction of food choice and consumption behavior. *British Journal of Social Psychology, 19,* 285-338. doi:10.1348/014466607X241540

Gawronski, B. (2009). Ten frequently asked questions about implicit measures and their frequently supposed, but not entirely correct answers. *Canadian Psychology, 50*, 141-150. doi:10.1037/a0013848

Giancola, P. R., & Parrott, D. J. (2008). Further evidence for the validity of the Taylor Aggression Paradigm. *Aggressive Behavior, 34,* 214-229. doi:10.1002/ab.20235

Giancola, P. R., & Zeichner, A. (1995). Construct validity of a competitive reaction time aggression paradigm. *Aggressive Behavior, 21,* 199-204. doi:10.1002/1098-2337(1995)21:3\199:AID-AB2480210303[3.0.CO;2Q

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. doi:10.1037//0022-3514.74.6.1464

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. doi:10.1037/h0087889

Grumm, M., Hein, S., & Fingerle, M. (2011). Predicting aggressive behavior in children with the help of measures of implicit and explicit aggression. *International Journal of Behavioral Development, 35,* 352-357. doi: 10.1177/0165025411405955

Hampel, R., & Selg, S. (1998). *FAF – Fragebogen zur Erfasung von Aggressivitätsfaktoren* [FAF – Inventory for the assessment of aggressiveness factors]. Göttingen: Hogrefe.

Hofmann, W., Friese, M., & Strack, F. (2009). Impulse and self-control from a dual systems perspective. *Perspectives on Psychological Science, 4,* 162-176. doi: 10.1111/j.1745-6924.2009.01116.x

Hofmann, W., Gschwendner, T., Friese, M., Wiers, R. W., & Schmitt, M. (2008). Working memory capacity and self-regulatory behavior: Toward an individual differences perspective on behavior determination by automatic versus controlled processes. *Journal Personality and Social Psychology, 95,* 962-977. doi: 10.1037/a0012705

Hollin, C. R., & Bloxsom, C. A. J. (2007). Treatments for angry aggression. In T.A. Gannon, T. Ward, A. R. Beech & D. Fisher (Eds.), *Aggressive offenders' cognition: Theory, research and practice* (pp. 215-229). Chichester, West Sussex: Wiley. doi: 10.1002/9780470746295.ch11

Imhoff, R., Schmidt, A. F.**,** Gerstenberg, F. X. R. (2014). Exploring the interplay of trait self-control and ego depletion: Empirical evidence for ironic effects. *European Journal of Personality, 28,* 413-442. doi:10.1002/per.1899

Inzlicht, M. & Schmeichel, B. J. (2012). What is ego depletion? Toward a mechanistic revision of the resource model of self-control. *Perspectives on Psychological Science, 7,* 450-463. doi: 10.1177/1745691612454134

Nosek, B. A., Hawkins, C. B., & Frazier, R. S. (2011). Implicit social cognition: From measures to mechanisms. *Trends in Cognitive Sciences, 15,* 152–159. doi:10.1016/j.tics.2011.01.005

Perugini, M., Richetin, J., & Zogmaister, C. (2010). Prediction of behavior. In B. Gawronski, & B. K. Payne (Eds.), *Handbook of social cognition – Measurement, theory, and applications* (p. 255- 277). New York, NY: Guilford.

Ramirez, J. M. (1993). Acceptability of aggression in four Spanish regions and a comparison with other European countries. *Aggressive Behavior, 19,* 185-197. doi: 10.1002/1098-2337(1993)19:3<185::AID-AB2480190304>3.0.CO;2-T

Richetin, J., & Richardson, D. S. (2008). Automatic processes and individual differences in aggressive behavior. *Aggression and Violent Behavior, 13,* 423-430. doi: 10.1016/j.avb.2008.06.005

Richetin, J., Richardson, D. S., & Mason, G. D. (2010). Predictive validity of IAT aggressiveness in the context of provocation. *Social Psychology, 41,* 27-34. doi 10.1027/1864-9335/a000005

Strack, F., & Deutsch, R. (2004). Reflective and impulsive determinants of social behavior. *Personality and Social Psychology Review, 8,* 220–247. doi: 10.1207/s15327957pspr0803_1

Stucke, T. S., & 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. doi: 10.1002/ejsp.285

Taylor, S. P. (1967). Aggressive behavior and physiological arousal as a function of provocation and the tendency to inhibit aggression. *Journal of Personality, 35,* 297-310. doi: 10.1111/j.1467-6494.1967.tb01430.x

Van Gelder, J.-L. (2013). Beyond rational choice: The hot/cool perspective of criminal decision making. *Psychology, Crime & Law, 19,* 745-763. doi: 10.1080/1068316X.2012.660153

Vohs, K. D., Glass, B. D., Maddox, W. T., & Markman A. B. (2011). Ego depletion is not just fatigue: Evidence from a total sleep deprivation experiment. *Social Psychology and Personality Science, 2,* 166-173. doi: 10.1177/1948550610386123

Wagner, D.D., Altman, M., Boswell, R. G., Kelley, W. M., & Heatherton, T. F. (2013). Self-regulatory depletion enhances neural responses to rewards and impairs top-down control. *Psychological Science.* Advance Online Publication. doi: 10.1177/0956797613492985

Wiers, R. W., Gladwin, T. E., Hofmann, W., Salemink, E., Ridderinkhof, K. R. (2013). Cognitive bias modification and cognitive control training in addiction and related psychopathology: Mechanisms, clinical perspectives, and ways forward. *Clinical Psychological Science, 1,* 192-212. doi: 10.1177/2167702612466547

Footnotes

[1] Control analyses on the full sample with outliers included could not replicate the central moderation pattern described below mirroring the profound influence of the three multivariate outliers as identified by Cook's distances.

[2] Control analyses showed that the results from the crucial hierarchical regression analysis described below could not be replicated with the full 13-item RA scale. However, controlling for the unreliability of the eliminated subset of items by utilizing factor scores of the whole 13-item RA subscale, the central results remained virtually unaltered: Running the same hierarchical multiple regression analysis with factor scores from an exploratory factor analysis (principal components, one factor solution, 21% of variance explained) of the whole 13-item RA subscale replicated the moderation pattern as reported below ($\Delta R^2_{Step1} = .25$, $p = .002$; $\Delta R^2_{Step2} = .15^{**}$, $p = .009$; $\Delta R^2_{Step3} = .00$, $p = .993$).

[3] A 2 (ED vs. Non-ED) x 2 (Escalation vs. Non-Escalation Phases in the TAP) mixed-model ANOVA revealed a marginally significant main effect of the ED manipulation ($F = 3.23$, $p = .08$, $\varepsilon = .05$) that ran into the expected direction as ego depleted participants showed more aggression in the TAP. Moreover, Escalation also showed a significant main effect on participants' aggression ($F = 5.41$, $p = .02$, $\varepsilon = .08$) in the expected direction (more reactive aggression after escalation). However, no ED x Escalation interaction emerged ($F < 1$), thus corroborating that the escalation in the TAP effected both groups in a similar way.

[4] Controlling for potential effects of participant sex in the central regression revealed that it was unrelated to TAP volume levels ($p = .44$) and its inclusion left the crucial moderations unaltered.

[5] We also explored whether the direct and indirect aggressiveness measures moderated the escalation effect (see Footnote 3) utilizing a difference measure of both escalation conditions. No moderation could be revealed. This corroborates that there is no differential

escalation effect of the ED manipulation but rather an effect on the impact of automatic vs. controlled precursors on reactive aggression as a function of situational boundary conditions reflecting self-control limitations.

Table 1

Overview of the Aggressive Behavior IAT (Agg-IAT) Protocol (German Items in Brackets).

Trials	Block	Left Key	Right Key	Target Items		Attribute Items	
				Me (Ich)	Others (Andere)	Peaceful (friedlich)	Aggressive (aggressiv)
20	1	aggressive	peaceful	Me (mir)	architect (Architektin)	talk (reden)	hunt (jagen)
20	2	others	Me	My (mein)	accountant (Buchhalter)	conciliation (Versöhnung)	revenge (Rache)
4 + 80	3	others or	me or	Me (mich)	educator (Erzieherin)	conversation (Gespräch)	punch (Faustschlag)
		aggressive	peaceful	I (ich)	farmer (Landwirt)	exchange (Austausch)	fight (Kampf)
20	4	peaceful	aggressive	Self (Selbst)	cook (Köchin)	compromise (Kompromiss)	hit (schlagen)
4 + 80	5	others or	me or		gatekeeper (Pförtner)	agreement (Verständigung)	avenge (rächen)
		peaceful	aggressive		cabinet maker (Tischlerin)	settlement (Einigung)	retaliate (zurückschlagen)
					filling station attendant (Tankwart)	counsel (Beratung)	threat (Drohung)
					dentist (Zahnärztin)	agree (einigen)	attack (Angriff)
					carpenter (Zimmermann)	give in (nachgeben)	beat (hauen)

Table 2

Descriptive Statistics, Reliabilities (Cronbach's α; in Brackets), and Zero-Order Correlations of Variables in the Experiment.

	Min	*Max*	*M*	*SD*	1.	2.	3.	4.	Ego Depletion Manipulation[a]
Sex[a]	-	-	-	-	-.26	-.02	-.01	.09	.10[b]
Age	19	50	24.93	5.47	.05	-.09	-.20	-.12	-.13
1. FAF Reactive Aggressiveness	.00	.75	.13	.17	(.62)	.09	.11	.14	.01
2. Aggressive Behavior IAT	-1.06	0.30	-0.44	0.33		(.69)	.03	.08	**-.47*** **
3. Aggressive Behavior IAT Errors	.01	.24	.07	.04			-	-.08	-.08
4. TAP Reactive Aggressiveness	1.00	5.73	2.10	1.36				(.99)	**.36** **

Note. $N = 58$; [a] higher values depict female sex/high ego depletion condition; [b] r_ϕ (Fisher's exact test)

** $p < .01$, *** $p < .001$

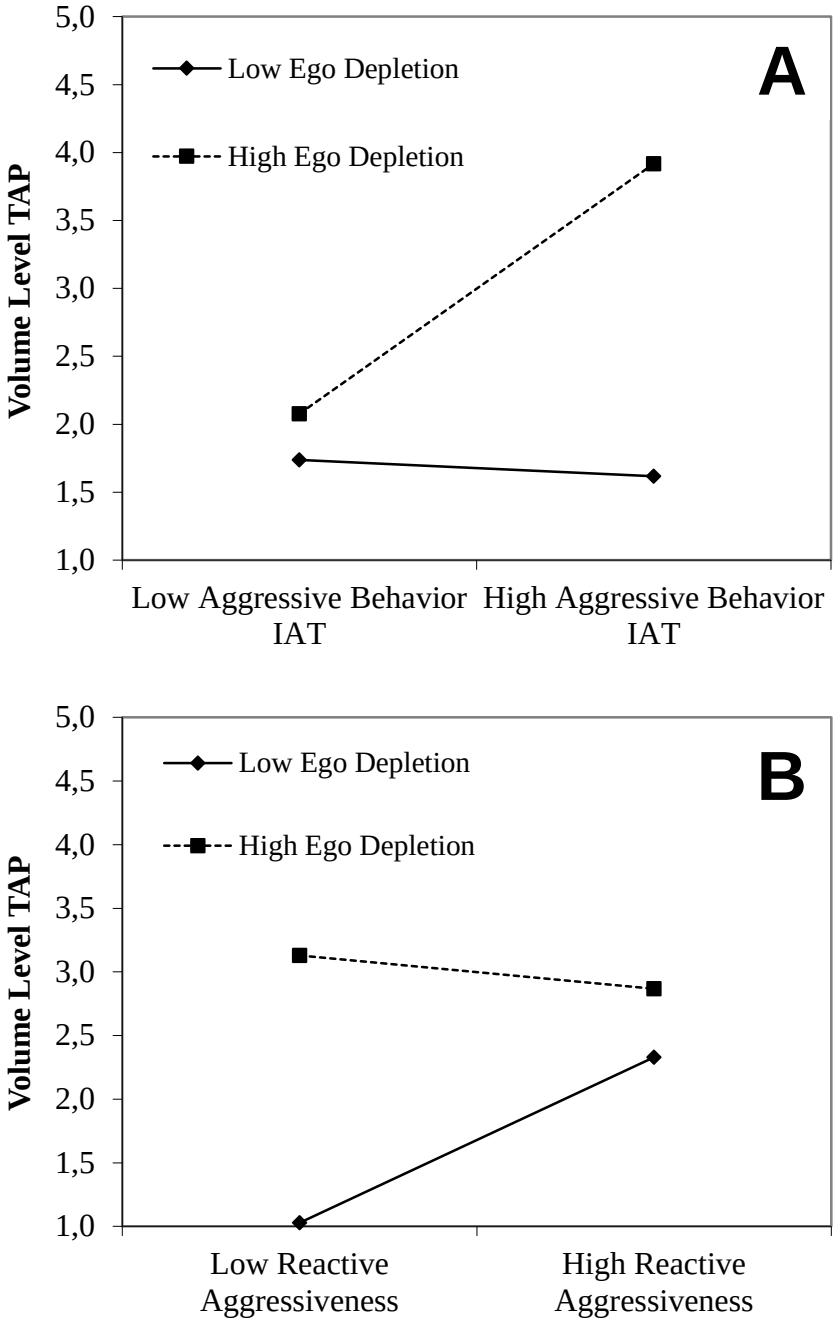
Table 3

Summary of Bootstrapped Hierarchical Regression Analysis for Controlled and Automatic Precursors of Reactive Aggression in the Modified Taylor Aggression Paradigm.

Predictor	ΔR^2	ß	*SE* ß	*p*
Step 1	.22			.004
Ego Depletion (ED)		.67	.20	.002
Aggr. Behav. IAT (Agg-IAT)		.42	.18	.022
React. Aggressiveness (RA)		.15	.17	.345
Step 2	.15			.014
ED		.66	.19	.001
IAT		.43	.16	.015
RA		.26	.16	.066
ED x Agg-IAT		.49	.16	.004
ED x RA		-.39	.24	.080
Agg-IAT x RA		-.13	.19	.463
Step 3	.00			.560
ED		.65	.20	.002
Agg-IAT		.39	.18	.026
RA		.16	.26	.438
ED x Agg-IAT		.49	.18	.005
ED x RA		-.37	.26	.088
Agg-IAT x RA		-.16	.26	.449
ED x Agg-IAT x RA		-.13	.25	.499

Note. $N = 58$.

Figure 1. *Reactive Aggression (Taylor Aggression Paradigm; TAP) as a Function of Ego Depletion and Automatic Precursors (Panel A) versus Controlled Precursors (Panel B) of Aggression.*



Appendix

Items remaining in the α-maximized FAF Reactive Aggressiveness Scale (English translations from the original German scale):

1. It is better to break someone's nose than to be a coward.
2. Someone who seriously insults me will earn a slap in the face.
3. A horse that does not pull well needs to be whipped.
4. Sometimes I imagine how much people who did not treat me well should actually suffer.
5. My motto is: Never trust a stranger!
6. If I have to resort to physical violence in order to defend my rights, I am willing to do so.
7. If a fly is annoying me, I will not be satisfied only until I have caught it.
8. If a friend of mine gets accosted, we will retaliate together.

Items removed from the original FAF Reactive Aggressiveness Scale (English translations from the original German scale):

1. If someone is unjust to me, I wish him/her a fierce punishment.
2. I wish to fiercely punish someone who has malignantly abandoned me.
3. I am careful with people who are friendlier than I would have expected.
4. A dog that doesn't obey needs to be hit.
5. If someone is nasty to my friend, I will join in when revenge is taken.