

Fachbereich Erziehungswissenschaft und Psychologie
der Freien Universität Berlin

**Dealing with interpersonal conflict:
Psychological and neuroimaging investigations of forgiveness**

Dissertation

zur Erlangung des akademischen Grades

Doktor der Philosophie (Dr. phil.)

Doctor of Philosophy (Ph.D.)

Vorgelegt von:

Ramzi Fatfouta, B. Sc.

Berlin, 2015

Vorblatt

Erstgutachter:

Prof. Dr. Hauke R. Heekeren

Zweitgutachterin:

Prof. Dr. Michela Schröder-Abé

Datum der Disputation: 08.09.2015

Eidesstattliche Erklärung

Hiermit erkläre ich an Eides statt,

- dass ich die vorliegende Arbeit selbstständig und ohne unerlaubte Hilfe verfasst habe,
- dass ich mich nicht bereits anderwärts um einen Doktorgrad beworben habe und keinen Doktorgrad in dem Promotionsfach Psychologie besitze und
- dass ich die zugrunde liegende Promotionsordnung vom 02.12.2008 kenne.

Berlin, den 30.06.2015

Ramzi Fatfouta

Table of contents

Acknowledgments.....	V
Summary	VI
Zusammenfassung	VIII
List of original publications	X
List of abbreviations.....	XI
List of figures	XII
1. Introduction	1
1.1. Forgiveness	2
1.2. Measurement issues in forgiveness research.....	2
1.3. Personality correlates of forgiveness: The case of narcissistic traits	4
1.4. Clinical correlates of forgiveness: The case of obsessive-compulsive symptoms.....	6
1.5. Neural correlates of forgiveness.....	8
2. Research aims	10
2.1. Establish new methods for studying forgiveness	10
2.2. Characterize the mechanisms behind narcissists' lack of forgiveness.....	10
2.3. Examine the link between OC symptoms and revenge.....	11
2.4. Identify a neural mechanism of forgiveness	11
3. Methodology.....	12
3.1. General methodology	12
3.1.1. Forgiveness Implicit Association Test.....	12
3.1.2. Transgression-recall paradigm.....	13
3.1.3. Questionnaires	13
3.2. Specific methodology.....	14
3.2.1. Real-life interaction and ultimatum game task	14
3.2.2. Functional magnetic resonance imaging.....	16
4. Summary of empirical studies.....	18
4.1. Study 1: Forgiving, fast and slow: Validity of the implicit association test for predicting differential response-latencies in a transgression-recall paradigm	18
4.2. Study 2: How forgiveness affects processing time: Mediation by rumination about the transgression.....	19
4.3. Study 3: Narcissism and lack of interpersonal forgiveness: The mediating role of state anger, state rumination, and state empathy	21
4.4. Study 4: Can the shrew be tamed? Narcissism, lack of forgiveness, and the moderating role of conciliatory gestures.....	23
4.5. Study 5: Associations between obsessive-compulsive symptoms, revenge, and the perception of interpersonal transgressions.....	24
4.6. Study 6: Forgiveness involves reduced connectivity between medial prefrontal and dorsal anterior cingulate cortex.....	25
5. Discussion	31
5.1. Specific discussion	31
5.1.1. New methods for studying forgiveness (Studies 1 and 2)	31
5.1.2. Mechanisms behind narcissists' lack of forgiveness (Studies 3 and 4).....	33
5.1.3. OC symptoms, revenge, and interpersonal transgressions (Study 5)	34
5.1.4. A neural mechanism of forgiveness (Study 6).....	36
5.2. General discussion	38
5.2.1. A neurocognitive model of forgiveness	38
5.2.2. Implications for future research.....	40
6. Conclusion	42
7. References.....	43
8. Appendix	56

Acknowledgments

I wish to thank my supervisor Hauke Heekeren for providing me enough freedom to pursue my own academic interests. I am deeply indebted to his encouragement, guidance, and excellent scientific feedback — thank you! I am also very grateful to Angela Merkl for introducing me to the topic of forgiveness, supporting my passion for research, and mentoring me during all stages of my dissertation studies. I would like to extend my gratitude to Michela Schröder-Abé for the innumerable discussions, e-mails, and phone calls, which helped me to survive a couple of arduous review processes. I thank my colleague Tanja Gerlach for sparking my interest in narcissism, discussing scientific and nonscientific topics, and sharing many laughs. I further thank all members of my dissertation committee for assisting me in finalizing my dissertation. I wish to express my appreciation to my colleagues at the Heekeren lab. Special thanks go to Yulia Oganian, Julia Rodríguez Buritica, Stefan Schulreich, and Dar Meshi. Also, I would like to thank the Cluster “Languages of Emotion” for funding two years of my Ph.D. (and again, thank you Hauke for organizing funding for the remaining year). Importantly, I owe many thanks to my family, especially Jolanta and Mohamed, who unconditionally supported me during the ups and downs of life as a doctoral student. Without them I would not be where I am today. Thanks to Firdaous and her husband Jan for informational, emotional, and moral support. Finally, no words can fully describe my gratitude for Ivo. Thank you with all my heart.

Summary

Forgiveness is central to social life because how we deal with relational conflicts influences intra- and interpersonal functioning. Despite recent progresses in forgiveness research, many problems remain unresolved. One set of problems is methodological and involves the overreliance on forgivers' self-reports. Another set of problems is conceptual and involves the neglect of important correlates of forgiveness, such as personality, psychopathology, and brain function. The six studies constituting this dissertation represent an attempt to address these issues and, via this, to advance a multifaceted understanding of forgiveness.

Study 1 developed and examined the psychometric properties of an indirect (i.e., non-self-report) measure of forgiveness, the forgiveness Implicit Association Test (IAT). As an objective behavioral criterion of forgiving, Study 1 assessed the response time (RT) individuals need to rate their current thoughts and feelings toward a real-life transgressor (i.e., their state forgiveness). Study 2 extended Study 1 by examining the processes underlying the generation of individuals' forgiveness-response decisions. Studies 3 and 4 examined how the personality trait of narcissism relates to forgiveness (in particular, lack thereof). Specifically, the mediating role of socio-cognitive variables (Study 3) and the moderating role of conciliatory gestures (e.g., apology) were probed (Study 4). Study 5 explored how intrusive thoughts and actions (i.e., obsessive-compulsive [OC] symptoms) relate to forgiveness' conceptual foil, namely, revenge. Finally, Study 6 combined behavioral and neuroimaging techniques to identify a neural mechanism of forgiveness: the interplay between the medial prefrontal and dorsal anterior cingulate cortex.

In Study 1, I show that individuals high in implicit (IAT-measured) and explicit (self-reported) forgiveness are faster to rate their state forgiveness and that the forgiveness IAT has incremental validity regarding the prediction of RT. Study 1 suggests that forgiveness involves both implicit and explicit aspects, further accentuating the value of RTs for capturing forgiveness-related information processing. In Study 2, I demonstrate that the link between forgiveness and RT is mediated by differential rumination about the transgression; individuals who were more forgiving also tended to ruminate less about their transgression, which in turn, promoted quicker responses. Thus, the ability to inhibit transgression-related thoughts seems to speed information processing when rating one's state forgiveness. Regarding personality correlates of forgiveness, in Study 3 I provide evidence that states associated with lack of forgiveness — anger, rumination, and diminished empathy —

differentially mediate the link between two distinct dimensions of narcissism (i.e., admiration and rivalry). Whereas admiration was negatively related to lack of forgiveness via higher empathy, rivalry was positively related to lack of forgiveness via more anger and rumination and less empathy. By contrast, in Study 4 I show that the positive rivalry/lack of forgiveness link can be altered (i.e., is moderated) by a verbal expression of regret together with a costly apology. Whereas motivations to retaliate against the transgressor remained unchanged, motivations to avoid the transgressor could in fact be reduced. Together, Studies 3 and 4 suggest narcissism to be a multidimensional construct and underscore the distinct forgiveness-related consequences of its facets. In Study 5, I show that OC symptoms in general, and obsessing in particular, positively relate to a number of revenge-related phenomena (i.e., vengeful attitudes, dispositions, and motivations) and heightened perceptions of interpersonal transgressions in daily life. Similar to Studies 3 and 4, individuals high in OC symptoms seem to be characterized by pronounced unforgiveness. Finally, in Study 6 I show that forgiveness involves reduced connectivity between two brain regions: the dorsal anterior cingulate cortex (DACC), which is involved in conflict detection, and the medial prefrontal cortex (MPFC), which is involved in mentalizing. MPFC/DACC connectivity was inversely related to implicit but not explicit forgiveness, again highlighting the utility of the forgiveness IAT even for the assessment of spontaneous brain activity. In sum, all six studies suggest that forgiveness is a complex phenomenon that is influenced by a multitude of situational as well as relational factors and can only be fully understood if we use a multi-method approach.

Keywords: forgiveness, interpersonal relationships, implicit social cognition, narcissism, obsessions, functional magnetic resonance imaging (fMRI)

Zusammenfassung

Vergebung ist zentral für das Sozialleben, denn wie wir mit zwischenmenschlichen Konflikten umgehen hat sowohl intra- als auch interpersonale Konsequenzen. Trotz jüngster Fortschritte in der Vergebungsforschung bleiben viele Probleme ungelöst. Probleme methodischer Art betreffen die nahezu ausschließliche Verwendung von Selbstberichtsfragebögen für die Erfassung von Vergebung. Probleme konzeptueller Art betreffen die Vernachlässigung wichtiger Korrelate von Vergebung wie Persönlichkeit, Psychopathologie und Hirnfunktionen. Die sechs empirischen Beiträge, die den Kern dieser Dissertation bilden, stellen einen Versuch dar, diesen Problemen nachzugehen. Ziel ist es, ein differenzierteres Verständnis von Vergebung zu gewinnen.

Studie 1 entwickelte und evaluierte die psychometrischen Eigenschaften eines indirekten (d.h. nicht auf Selbstbericht beruhenden) Vergebungsmaßes, nämlich den Impliziten Assoziationstest (IAT) für Vergebung. Als objektives Verhaltenskriterium für Vergebung diente die Reaktionszeit (RZ), mit der Personen ihre momentanen Gedanken und Gefühle im Hinblick auf eine real erlebte Verletzung einschätzen (d.h. ihre „Zustandsvergebung“). Studie 2 erweiterte Studie 1 um die Betrachtung von Prozessen, die der Generierung dieser RZ-Werte zugrunde liegen. Hinsichtlich der Persönlichkeitskorrelate von Vergebung untersuchten Studien 3 und 4 wie dispositionaler Narzissmus mit Vergebung (bzw. deren Mangel) zusammenhängt. Insbesondere wurden die vermittelnde Rolle sozio-kognitiver Variablen (Studie 3) sowie die moderierende Wirkung versöhnlicher Gesten (z.B. das Vorhandensein einer Entschuldigung) getestet (Studie 4). Studie 5 fokussierte darauf, wie Zwangsgedanken und damit einhergehende Zwangshandlungen (d.h. obsessiv-kompulsive Symptome) mit Rache zusammenhängen. Studie 6 kombinierte behaviorale und funktionelle Bildgebungsverfahren, um einen neuronalen Mechanismus von Vergebung zu identifizieren: das Zusammenspiel zwischen medial präfrontalem und dorsal anteriorem Kortex.

Studie 1 zeigt, dass hoch vergebende Personen — sowohl implizit (IAT) als auch explizit (selbstberichtet) gemessen — schneller waren, wenn es um die Einschätzung ihrer Gedanken und Gefühle gegenüber dem Transgressor geht und dass der Vergebungs-IAT inkrementell valide für die RZ-Vorhersage ist. Studie 1 deutet darauf hin, dass Vergebung sowohl implizite als auch explizite Aspekte beinhaltet. Daneben scheinen RZen einen Mehrwert für die Erfassung vergebungsbezogener Informationen zu besitzen. Studie 2 demonstriert, dass der Zusammenhang zwischen Vergebung und RZ über interindividuelle Unterschiede in der Ruminationsneigung über die Transgression vermittelt wird: Personen,

die vergebender waren, ruminieren weniger über die Transgression, was wiederum schnelleres Antwortverhalten begünstigt. Demnach scheint die Fähigkeit, transgressionsbezogene Gedanken zu inhibieren, die Informationsverarbeitung beim Einschätzen der „Zustandsvergebung“ zu beschleunigen. Studie 3 weist auf, dass der Zusammenhang zwischen zwei distinkten Narzissmusfacetten (d.h. Bewunderung und Rivalität) und dem Mangel an Vergebung differenziell über die sozio-kognitiven Variablen Zustandsärger, Zustandsrumination und Zustandsempathie vermittelt werden. Die Beziehung zwischen Bewunderung und Mangel an Vergebung war negativ und wurde über mehr Empathie für den Transgressor vermittelt; die Beziehung zwischen Rivalität und Mangel an Vergebung war positiv und wurde über mehr Ärger, mehr Rumination und weniger Empathie vermittelt. Studie 4 hingegen verdeutlicht, dass die positive Beziehung zwischen Rivalität und Mangel an Vergebung verändert (d.h. moderiert) werden kann, sofern eine verbale Entschuldigung mit Kosten für den Transgressor dargeboten wird. Allerdings wird nur die Vermeidungsmotivation gegenüber dem Transgressor reduziert, wohingegen die Rachemotivation nahezu unverändert bleibt. Zusammen deuten Studien 3 und 4 darauf hin, dass Narzissmus ein mehrdimensionales Konstrukt darstellt, dessen Facetten distinkte Konsequenzen für Vergebung besitzen. Studie 5 liefert einen Beleg dafür, dass obsessiv-kompulsive Symptome im Allgemeinen und Obsessionen im Besonderen positiv mit verschiedenen rachebezogenen Phänomenen (z.B. Einstellungen, Dispositionen und Motivationen) sowie einer erhöhten Wahrnehmung von interpersonalen Verletzungen zusammenhängen. Ähnlich zu Studien 3 und 4 scheinen Personen, die an obsessiv-kompulsiven Symptomen leiden, durch einen Mangel an Vergebung charakterisierbar zu sein. Schließlich zeigt Studie 6, dass Vergebung mit einer reduzierten funktionellen Konnektivität zwischen dem dorsalen anterioren cingulären Kortex (DACC) und dem medialen präfrontalen Kortex (MPFC) einhergeht. Die MPFC/DACC Konnektivität hing negativ mit impliziter aber nicht expliziter Vergebung zusammen, was erneut den inkrementellen Nutzen des Vergabungs-IATs sogar für die Erfassung spontaner Hirnaktivität verdeutlicht. Zusammengenommen deuten diese Befunde daraufhin, dass Vergebung ein komplexes Phänomen darstellt, das von einer Vielzahl situativer und relationaler Faktoren beeinflusst wird und nur unter Einsatz mehrerer Messmethoden besser verstanden werden kann.

Schlagwörter: Vergebung, zwischenmenschliche Beziehungen, implizite soziale Kognition, Narzissmus, Obsessionen, funktionelle Magnetresonanztomographie (fMRT)

List of original publications

This dissertation is based on the following original research articles:

Study 1

Fatfouta, R., Schröder-Abé, M., & Merkl, A. (2014). Forgiving, fast and slow: Validity of the implicit association test for predicting differential response latencies in a transgression-recall paradigm. *Frontiers in Psychology, 5*, 1-14. doi: 10.3389/fpsyg.2014.00728

The original article is online available at: <http://dx.doi.org/10.3389/fpsyg.2014.00728>

Study 2

Fatfouta, R. (2015). How forgiveness affects processing time: Mediation by rumination about the transgression. *Personality & Individual Differences, 82*, 90-95. doi: 10.1016/j.paid.2015.03.016

The original article is online available at: <http://dx.doi.org/10.1016/j.paid.2015.03.016>

Study 3

Fatfouta, R., Gerlach, T. M., Schröder-Abé*, M., & Merkl*, A. (2015). Narcissism and lack of interpersonal forgiveness: The mediating role of state anger, state rumination, and state empathy. *Personality and Individual Differences, 75*, 36-40. doi: 10.1016/j.paid.2014.10.051 *equal contribution

The original article is online available at: <http://dx.doi.org/10.1016/j.paid.2014.10.051>

Study 4

Fatfouta*, R., & Gerlach*, T. M. (under review). Can the shrew be tamed? Narcissism, lack of forgiveness, and the moderating role of conciliatory gestures. *European Journal of Personality*. *equal contribution

Study 5

Fatfouta, R., & Merkl, A. (2014). Associations between obsessive-compulsive symptoms, revenge, and the perception of interpersonal transgressions. *Psychiatry Research, 219*(2), 316-321. doi: 10.1016/j.psychres.2014.05.038

The original article is online available at: <http://dx.doi.org/10.1016/j.psychres.2014.05.038>

Study 6

Fatfouta, R., Meshi, D., Merkl, A., & Heekeren, H. R. (under review). Forgiveness involves reduced connectivity between medial prefrontal and dorsal anterior cingulate cortex. *Cerebral Cortex*.

List of abbreviations

AI	Anterior Insula
ACC	Anterior Cingulate Cortex
BOLD	Blood Oxygen Level Dependent
CI	Confidence Interval
DACC	Dorsal Anterior Cingulate Cortex
DLPFC	Dorsolateral Prefrontal Cortex
EPI	Echo Planar Imaging
fMRI	Functional Magnetic Resonance Imaging
FWHM	Full Width at Half Maximum
GLM	General Linear Model
IAT	Implicit Association Test
MNI	Montreal Neurological Institute
MPFC	Medial Prefrontal Cortex
MP-RAGE	Magnetization-Prepared Rapid Acquisition Gradient Echo
MRI	Magnetic Resonance Imaging
NARC	Narcissistic Admiration and Rivalry Concept
NARQ	Narcissistic Admiration and Rivalry Questionnaire
NPI	Narcissistic Personality Inventory
OC	Obsessive Compulsive
OCD	Obsessive Compulsive Disorder
OCI-R	Obsessive-Compulsive Inventory-Revised
PPI	Psychophysiological Interaction
RT	Response Time
TE	Echo Time
TR	Repetition Time
TRIM	Transgression-Related Interpersonal Motivations
TTF	Tendency to Forgive
TMS	Transcranial Magnetic Stimulation

List of figures

Figure 1. Example trial during block 3 of the forgiveness IAT.....	12
Figure 2. Timeline of a single trial in the Ultimatum Game.....	15
Figure 3. Scatterplots of the forgiveness/rumination-response time relationship.....	20
Figure 4. Mediation models illustrating the direct and indirect effects of forgiveness on response time.....	20
Figure 5. Mediation models illustrating the direct and indirect effects of narcissism on lack of forgiveness.....	22
Figure 6. Effect of apology on narcissists' lack of forgiveness.....	24
Figure 7. Behavioral results in the UG task.....	27
Figure 8. Affective responses to unfairness.....	27
Figure 9. Fairness, attractiveness, and trustworthiness ratings.....	28
Figure 10. BOLD signal change related to unfairness.....	28
Figure 11. Overlap of BOLD signal change related to unfairness.....	29
Figure 12. Brain activation associated with unfairness by the partner.....	29
Figure 13. Reduced connectivity between MPFC and DACC predicts forgiveness.....	30
Figure 14. A neurocognitive model of forgiveness.....	39

1. Introduction

The weak can never forgive. Forgiveness is the attribute of the strong.

— Mahatma Gandhi

Conflict is a ubiquitous feature of social life. Breaches of trust, broken promises, and failures of cooperation are just some of the challenges humans can face in interpersonal relationships. Many of these relationships, however, do not forfeit their value despite the occurrence of conflict. This capacity to forgo revenge and, instead, to take a conciliatory stance toward a transgressing party is known as forgiveness (McCullough, Kurzban, & Tabak, 2011, 2013). The concept of forgiveness plays a central role in everyday life and is deeply rooted in philosophical and religious traditions (for reviews, see McCullough & Worthington, 1999; Rye et al., 2000). Over the past decades, forgiveness has also become an increasingly important topic of empirical research. Psychologists, for example, aim to describe, explain, and predict individual differences in forgiveness. In pursuing this aim, they try to answer one key question: Who forgives whom and under what circumstances? In addition, psychological researchers have become interested in extending methods for assessing forgiveness beyond self-report. Neuroscientists, by contrast, aim to uncover the neural correlates of forgiveness. The key question they try to answer is how specific brain areas influence our decision to forgive (or not to forgive). The overarching goal of this dissertation was to contribute to answering these questions by adopting a multi-method approach to the study of forgiveness.

In this dissertation, I will present the results of six empirical studies that investigated the psychological and neural correlates of forgiveness. The thesis is structured as follows: First, I will briefly summarize the relevant theoretical and empirical background necessary to place the dissertation studies in a broader context. Specifically, I will focus on issues in the measurement of forgiveness, personality and clinical correlates of forgiveness (in particular, lack thereof) and, finally, the neural correlates of forgiveness. Second, I will explain the main research aims of this thesis. Third, I will provide a succinct summary of the materials and methods used. Fourth, I will sum up each of the six studies, which together form the main part of this thesis. Finally, I will discuss theoretical and practical implications of these studies. The dissertation concludes with a neurocognitive model of forgiveness integrating the intra- and interpersonal consequences of forgiveness into a coherent framework.

1.1. Forgiveness

Forgiveness can be conceptualized either as a state or a trait. Whereas state forgiveness pertains to individuals' thoughts and feelings toward a specific transgressor (McCullough et al., 1998; McCullough, Worthington, & Rachal, 1997), trait forgiveness characterizes individuals' general (i.e., cross-situational and cross-temporal) tendency to forgive others (Allemand, Sassin-Meng, Huber, & Schmitt, 2008; Brown, 2003). In this dissertation, I will focus on both — forgiveness as a transgression-specific response and as a disposition.

1.2. Measurement issues in forgiveness research

To measure forgiveness, researchers have usually adopted scenario-based (e.g., Berry, Worthington, Parrott, O'Connor, & Wade, 2001) or questionnaire-based (e.g., Brown, 2003) approaches. In the former case, individuals are required to imagine a hypothetical transgression and then to indicate their likelihood of forgiving a transgressor. In the latter, individuals are required to rate statements such as “I tend to get over it quickly when someone hurts my feelings” (trait forgiveness; Brown, 2003) or “I'm going to get even” (state forgiveness; McCullough et al., 1997). Common to both of these approaches is their reliance on respondents' verbal report. They are therefore referred to as direct measures (De Houwer, 2006).

Direct measures, albeit valid and reliable, may face researchers with two problems (Greenwald et al., 2002). First, individuals may be unwilling to disclose information about their forgiveness due to social-desirability concerns. For example, those with low-to-moderate forgiveness may self-enhance by presenting themselves in a favorable light (e.g., Paulhus & John, 1998). Second, individuals may be unaware of certain aspects of their forgiveness due to limited introspective abilities. For example, they may state or believe they have forgiven a transgressor, yet implicitly continue holding a grudge — a phenomenon also known as hollow forgiveness (Fincham, 2010; Fincham & Beach, 2001). To address these problems and, hence, to complement existing self-report methods, indirect (i.e., non-self-report) measures are needed (for a comment, see Fatfouta, Jacobs, & Merkl, 2013).

The need for indirect forgiveness measures can be better understood if we look at how humans process information about themselves and their social environment. Dual-process models (e.g., Back, Schmukle, & Egloff, 2009; Strack & Deutsch, 2004) posit two systems of information processing: reflective (or controlled) and impulsive (or automatic). In the reflective system, behavior results from slow decision processes and leads to propositional

representations of the self (explicit personality self-concept; e.g., “I am a forgiving person”). In the impulsive system, behavior results from fast spread-of-activation processes and leads to associative representations of the self (implicit personality self-concept; e.g., me-forgiving). Whereas individual differences in the explicit personality self-concept can be assessed directly, variations in the implicit personality self-concept can be assessed indirectly (Back et al., 2009). Importantly, indirect measures have been shown to yield incremental validity over direct measures, thereby providing a more nuanced picture of an individual’s underlying personality (Back et al., 2009; Bornstein, 2002; Rudolph, Schröder-Abé, Riketta, & Schütz, 2010).

At first sight, it may be difficult to think of forgiveness as an “automatic” reaction to a transgression. Indeed, the traditional view on forgiveness is that it requires a deliberate decision (e.g., Fincham, 2000). This decision involves overcoming one’s negative (i.e., revenge and avoidance) motivations toward the transgressor (McCullough et al., 1998) and, instead, restoring a more positive (i.e., benevolent) stance toward him/her (McCullough, Root, & Cohen, 2006). Self-help books such as “Forgiveness is a Choice” (Enright, 2001) or “Choosing Forgiveness: Your Journey to Freedom” (DeMoss, 2008) illustrate this line of thought. In everyday life, however, individuals do not always reflect on their tendency to forgive others (e.g., “Do I forgive him/her?”). Specifically, in close relationships forgiveness may be evoked automatically without individuals’ awareness (for a review, see Karremans & Van Lange, 2008). The rationale here is that over time people learn to associate forgiveness with positive relational outcomes (e.g., commitment), thus reinforcing automatic/impulsive responding in a constructive (vs. destructive) manner (Karremans & Aarts, 2007). Hence, if a person repeatedly engages in forgiving behavior, he/she may develop associative links between the self and forgiveness (for a similar theoretical perspective, see Schnabel & Asendorpf, 2010). So far, however, this implicit self-concept of forgiveness (or implicit forgiveness) has not yet been examined.

Thus, consistent with dual-process process models of social behavior (Back et al., 2009; Strack & Deutsch, 2004), it is assumed that forgiveness likely involves both explicit *and* implicit processes. To shed light on these implicit processes, indirect measures may prove useful. Developing such an indirect forgiveness measure is crucial and can assist in circumventing the problems inherent in self-report (Fatfouta et al., 2013). Nevertheless, there has been no study that has addressed this issue. Study 1 sought to close this gap. Study 2 aimed to examine the processes underlying forgiveness-response decisions.

1.3. Personality correlates of forgiveness: The case of narcissistic traits

Regardless of its conceptualization, forgiveness has been associated with beneficial consequences for both personal and social functioning (for a review, see Worthington, Witvliet, Pietrini, & Miller, 2007). At the individual level, high forgiveness is associated with higher empathy (Brown, 2003), higher emotional stability (McCullough & Hoyt, 2002), and greater satisfaction with life (Brown & Phillips, 2005). Likewise, individuals high in forgiveness tend to be less angry (Berry, Worthington, O'Connor, Parrott, & Wade, 2005), less stressed (Harris et al., 2006), and less likely to ruminate about negative interpersonal events (Ysseldyk, Matheson, & Anisman, 2007). At the relational level, forgiveness is similarly associated with positive outcomes, such as higher relationship satisfaction (Allemand, Amberg, Zimprich, & Fincham, 2007), greater commitment (Finkel, Rusbult, Kumashiro, & Hannon, 2002), and increased closeness to the partner (Tsang, McCullough, & Fincham, 2006).

Despite these benefits, some individuals have a more difficult time granting forgiveness than others. Specifically, narcissism has been suggested to be “the antithesis of the forgiving personality” (Emmons, 2000, p. 164). Narcissism can be defined either as a form of psychopathology (i.e., narcissistic personality disorder) or, as in the present thesis, a continuous personality trait (for a conceptual introduction, see Pincus & Lukowitsky, 2010). As a personality trait, narcissism is marked by a grandiose self-view, lack of empathy, feelings of entitlement, and interpersonal exploitation (Rhodewalt & Peterson, 2009). Narcissism has also been related to revenge fantasies (Raskin & Novacek, 1991) and aggressive responses following ego-threats (Bushman & Baumeister, 1998; Twenge & Campbell, 2003).

However, studies examining the narcissism/lack of forgiveness link have yielded mixed results, with some supportive (Eaton, Struthers, & Santelli, 2006a; Exline, Baumeister, Bushman, Campbell, & Finkel, 2004) and others not (Brown, 2004). This may be attributed to several factors, including aggregation of adaptive and maladaptive narcissism dimensions into one composite score (Ackerman et al., 2011), as measured by the narcissistic personality inventory (NPI; Raskin & Hall, 1979). To address this problem and, hence, to obtain a more nuanced understanding of narcissism, researchers have called for a distinction of these dimensions (Ackerman et al., 2011; Clarke, Karlov, & Neale, 2014).

The Narcissistic Admiration and Rivalry Concept (NARC; Back et al., 2013) allows for such a distinction. According to the NARC, narcissism can be conceptualized in two higher-order dimensions: admiration and rivalry. In essence, admiration denotes an agentic

tendency toward self-enhancement via self-promotion, whereas rivalry denotes an antagonistic tendency toward self-protection via self-defense. Individuals high in admiration show self-assured and expressive behavior, are popular among their peers, and have a problem-focused approach toward transgressions. Individuals high in rivalry, by contrast, show aggressive and combative behavior, are disliked by their peers, and demonstrate more hostile reactions (Back et al., 2013; Dufner et al., 2014). So far, however, the facet-specific contributions of both narcissism facets to lack of forgiveness remain largely unexplored.

To better understand narcissists' lack of forgiveness, it is important to examine the factors underlying the narcissism/lack of forgiveness link. Three socio-cognitive factors are deemed relevant here: anger toward the transgressor (state anger), rumination about the transgression (state rumination), and empathy toward the transgressor (state empathy). These variables are not only central to accounting for individuals' lack of forgiveness (for meta-analytic evidence, see Fehr, Gelfand, & Nag, 2010) but they are also related to narcissism. For example, narcissism has been linked to greater levels of anger (Rhodewalt & Morf, 1998), especially following interpersonal rejection (Twenge & Campbell, 2003). Moreover, narcissists are more likely to ruminate about past angry experiences, which in turn fuels their destructive behavior (Krizan & Johar, 2014). Finally, narcissists often fail to appreciate the feelings of others, that is, they are relatively low in empathy (Watson, Grisham, Trotter, & Biderman, 1984; Watson & Morris, 1991). Based on these previous findings, it is hypothesized that these socio-cognitive variables (i.e., state anger, state rumination, and state empathy) may be key factors underlying the lack of forgiveness in rivalry and admiration.

Aside from examining the factors mediating the narcissism/lack of forgiveness link, it is also important to investigate moderators of this link. So far, the conditions under which narcissists' lack of forgiveness can be mitigated remain unknown. However, prior work has pointed to the value of conciliatory gestures (e.g., making amends or apologizing) in facilitating forgiveness (Riek & Mania, 2012). Apologies can convey concern and remorse for the individuals' hurt (Fehr et al., 2010) and, hence, might decrease aggressive behavior toward the transgressor (McCullough, Pedersen, Tabak, & Carter, 2014; Ohbuchi, Kameda, & Agarie, 1989). Moreover, apologies can signal the transgressor's willingness to refrain from similar transgressions in the future, thereby paving the way for the re-establishment of positive relations (McCullough et al., 2013). Past research has suggested that apologies could reduce ego-threats (Eaton, Struthers, & Santelli, 2006b), indicating that apology might mitigate narcissists' defensive behavior. Thus, when the transgressor apologizes for his/her wrongdoing, this should reduce narcissists' lack of forgiveness.

In sum, narcissism likely encompasses two distinct trait dimensions — admiration and rivalry (Back et al., 2013). Together, these dimensions may provide a more comprehensive view of narcissism’s interplay with socio-cognitive variables as well as conciliatory gestures in predicting forgiveness (or lack thereof) in response to transgressions. Yet, no studies to date have provided evidence in this direction. Studies 3 and 4 sought to illuminate this issue.

1.4. Clinical correlates of forgiveness: The case of obsessive-compulsive symptoms

Although a large body of evidence has implicated failure to forgive/vengeance as a positive predictor of psychological distress (see Enright & Fitzgibbons, 2000), several important questions remain unanswered. One unanswered question concerns the association between vengeance and obsessive-compulsive disorder (OCD). This is striking given the proposed similarity between fantasies of revenge (e.g., recurrent urges to get even) and intrusive thoughts/obsessions (Horowitz, 2007), a hallmark feature of OCD (Leckman et al., 2010). To date, there is only indirect evidence that the concept of “revenge” may play a role in OCD. Still, an involvement of vengeance phenomena in OCD can be inferred from two lines of research: first, psychoanalytic accounts of OCD, and second, research on anger in OCD.

First, according to psychoanalytic theory, OCD can be conceptualized as a compromise between two conflicting tendencies — the ego and superego versus unacceptable (e.g., aggressive) impulses originating from the id. Obsessive-compulsive (OC) symptoms, such as obsessing are taken to reflect either an outbreak of these intolerable id-impulses or a punishment of a strict superego. Hence, whereas non-OCD individuals have evolved an adaptive way of dealing with these impulses, individuals with OCD are assumed to suppress them (Fenichel, 1945; Kempke & Luyten, 2007). This view, albeit controversial, has only recently gained interest from psychiatric researchers. For example, Moritz et al. (2009) have shown that individuals with OCD report more latent aggression (i.e., disguised hostility/aggression toward others) than healthy controls. This latent aggression, in turn, can promote ambivalence toward others, fuel interpersonal conflicts, and reveal itself in constant late arrivals to therapy sessions (Moritz, Kempke, Luyten, Randjbar, & Jelinek, 2011; Moritz et al., 2009).

Second, a limited number of studies have investigated the role of anger and aggression in OCD. Results from clinical (Whiteside & Abramowitz, 2005) and non-clinical (Whiteside & Abramowitz, 2004) samples suggest that OC symptoms are related to greater anger experience but not anger expression. Similarly, more recent studies have shown that

individuals with OCD tend to suppress their aggressive tendencies rather than acting them out overtly (Moscovitch, McCabe, Antony, Rocca, & Swinson, 2008; Radomsky, Ashbaugh, & Gelfand, 2007). The association between OC symptoms and aggression — particularly obsessing and aggressive fantasies — was also demonstrated in another study by Cogan et al. (2004).

Anger may be conceived as a central affective component of revenge (Emmons, 1992). However, anger and revenge represent distinct, albeit related, phenomena (Wilkowski, Hartung, Crowe, & Chai, 2012). Whereas anger pertains to a subjective emotional experience following provocation (Spielberger, 1988; Wilkowski, Robinson, & Troop-Gordon, 2010), revenge pertains to beliefs about the desirability or morality of vengeful acts for goal attainment (e.g., restoring the moral balance; McCullough, Bellah, Kilpatrick, & Johnson, 2001). Importantly, and similar to anger, revenge does not need to be acted out; instead feelings and fantasies of revenge can provide feelings of safety, satisfaction, and the experience of regained control and power. Additionally, they can help to overcome feelings of injustice and victimization (Gäbler & Maercker, 2011; Seebauer, Froß, Dubaschny, Schönberger, & Jacob, 2014).

Prior work has established a firm link between OCD and a variety of social difficulties, such as impairment of family relations and feeling victimized by others (Painuly, Grover, Mattoo, & Gupta, 2012; Piacentini, Bergman, Keller, & McCracken, 2003; Storch et al., 2006). Furthermore, it has been suggested that interpersonal conflicts and, by extension, transgressions, may arise from external criticism and/or from contravening the rules imposed by OCD individuals (Moritz, Niemyer, Hottenrott, Schilling, & Spitzer, 2012; Storch et al., 2012). For example, a family member may ridicule the individual with OCD (e.g., accusations of madness) or refrain from providing reassurance behavior (e.g., re-check whether the door is locked; Purdon, 2011). Against this backdrop, it seems plausible that individuals with OCD may wish to restore relational justice, that is, resort to revenge.

Taken together, recent studies have linked anger and (latent) aggression to OC symptoms (e.g., Moritz et al., 2011; Moritz et al., 2009; Whiteside & Abramowitz, 2004; Whiteside & Abramowitz, 2005). How the outcome of unresolved anger (i.e., revenge) relates to these symptoms, however, has remained unexplored. Further, it has remained unexplored whether and to what extent the frequency of interpersonal transgressions incurred is associated with OC symptom severity. Study 5 aimed to tackle these open questions.

1.5. Neural correlates of forgiveness

Aside from personality and clinical correlates of forgiveness, it is also important to examine its neural correlates. Crucially, an examination of the neural correlates can help identify core psychological processes underlying our ability to forgive (Newberg, d'Aquili, Newberg, & deMarici, 2000). So far, neuroscience research on forgiveness has relied on three types of paradigms: forgivability-judgment tasks (i.e., judging the forgivability of fictitious transgressions; Farrow et al., 2005; Farrow et al., 2001; Hayashi et al., 2010; Young & Saxe, 2009), mental-imagery tasks (i.e., imagining granting forgiveness toward hypothetical transgressors; Ricciardi et al., 2013), and game-theoretic tasks (i.e., refraining from punishment in an economic exchange; Brüne, Juckel, & Enzi, 2013; Strang, Utikal, Fischbacher, Weber, & Falk, 2014; Will, Crone, & Güroğlu, 2015). Across these tasks, the regions most consistently associated with forgiveness comprised anterior cingulate and medial prefrontal cortical areas — specifically, the dorsal anterior cingulate cortex (DACC) and the medial prefrontal cortex (MPFC) (e.g., Farrow et al., 2001; Hayashi et al., 2010; Will et al., 2015).

The DACC has been found to play a critical role in higher cognitive functions, such as error detection, conflict monitoring, and expectancy violation (Botvinick, Braver, Barch, Carter, & Cohen, 2001; Brown, 2013; Shenhav, Botvinick, & Cohen, 2013; Shenhav, Straccia, Cohen, & Botvinick, 2014; Sheth et al., 2012; Somerville, Heatherton, & Kelley, 2006). Moreover, the DACC has been implicated in affective processing, most notably the processing of social pain (Eisenberger, 2012a; Eisenberger, Lieberman, & Williams, 2003; Meyer, Williams, & Eisenberger, 2015) — that is, the unpleasant feeling that is associated with damage to one's social value or social connectedness (Eisenberger, 2012b). Recently, meta-analytic evidence has emerged supporting the contribution of the DACC to social pain across a variety of contexts, such as unreciprocated cooperation, negative social evaluation, or rejection (Rotge et al., 2015). Consistent with these findings, the DACC has been proposed to act as a “neural alarm system” consisting of two components (Eisenberger & Lieberman, 2004; Spunt, Lieberman, Cohen, & Eisenberger, 2012): First, a monitoring mechanism that detects discrepancies from desired (vs. actual) goal states (e.g., incurring an interpersonal transgression) and, second, a sounding mechanism that signals that the discrepancy needs to be addressed. Accordingly, the DACC is involved in both detection and appraisal of negative interpersonal events (also see Eisenberger, 2015; Kawamoto, Ura, & Nittono, 2015).

The MPFC has been implicated in various processes underlying high-level social cognition, such as self-reference, person perception, and mental-state reasoning/mentalizing

(for reviews, see Amodio & Frith, 2006; Frith & Frith, 2006; Van Overwalle, 2011). Mentalizing-related MPFC activity has reliably been observed in tasks in which individuals try to infer others' mental states (i.e., thoughts, feelings, and beliefs), specifically similar and/or socially close others (Heatherton et al., 2006; Mitchell, Macrae, & Banaji, 2006; Ochsner et al., 2005). Most recently, Welborn and Lieberman (2015) proposed that the MPFC might support mentalizing about "specific minds" and, hence, enable a person-specific theory of mind. In line with these previous findings, it has been suggested that the MPFC provides a signal about personally relevant social information (Krienen, Tu, & Buckner, 2010). It should also be noted that mentalizing is not confined to the MPFC, but likely encompasses other regions, such as the temporoparietal junction (Saxe & Wexler, 2005). Yet, recent meta-analytic evidence strongly implicates the MPFC in mentalizing-related processes (Denny, Kober, Wager, & Ochsner, 2012; Van Overwalle, 2009) — a finding consistent with lesion studies indicating that damage to the MPFC can impair mentalizing abilities (Rowe, Bullock, Polkey, & Morris, 2001; Shamay-Tsoory, Tomer, Berger, & Aharon-Peretz, 2003).

To shed more light on the MPFC/DACC involvement in forgiveness, it is important to consider its precursors: For forgiveness to take place, there has to be (1) an initial recognition that some sort of discrepancy has occurred (e.g., an interpersonal transgression), followed by (2) an attempt to resolve that discrepancy (Newberg et al., 2000). Stated differently, forgiveness not only requires the transgressed party to perceive the hurt incurred (Fincham, 2000; Fincham & Beach, 2001), but also to make sense of the reasons and motivations underlying the transgressor's behavior — briefly, the ability to "mentalize" about a specific person, the transgressor (e.g., Hayashi et al., 2010; Ricciardi et al., 2013). So far, however, neuroscience research on forgiveness is mute as to how the MPFC/DACC might interact in influencing our decision to forgive — specifically, in close interpersonal relationships, where forgiveness is deemed most relevant (Karremans & Aarts, 2007; Karremans et al., 2011). In addition, none of the aforementioned neuroscience studies tested for an association between individual differences (in forgiveness) and neural correlates of forgiveness.

In sum, most of the previous neuroscience research on forgiveness has been informed by paradigms involving fictitious transgressions, with hypothetical consequences for the transgressor and the transgressed party (Farrow et al., 2005; Farrow et al., 2001; Young & Saxe, 2009). In these paradigms, forgiveness has been shown to involve conflict-related (DACC) and mentalizing-related (MPFC) components. However, forgiveness is a complex psychological process (Worthington, 2006). It therefore seems unlikely that single brain areas in isolation can implement such a process (Newberg et al., 2000). Instead, a neural network

explaining the interplay between MPFC/DACC may be a better marker of forgiveness (for a discussion of network approaches in neuroscience research, see Aue, Lavelle, & Cacioppo, 2009). Consistent with the functional distinction between the DACC and the MPFC in terms of discrepancy detection on the one hand (DACC) and its regulation via mentalizing-related computations (MPFC) on the other (Etkin, Egner, & Kalisch, 2011), it is suggested that forgiveness depends on interactions between these two brain areas. As no study to date has attempted to examine such a neural mechanism, Study 6 aimed to close this gap in the literature.

2. Research aims

The general aim of this thesis was to advance our understanding of the psychological and neural correlates of forgiveness. The specific aims were fourfold:

2.1. Establish new methods for studying forgiveness

There is a mono-method bias (i.e., overreliance on a single measurement method) in forgiveness research, which calls for the establishment of alternative measures (Hoyt & McCullough, 2005; McCullough, Hoyt, & Rachal, 2000). Study 1 developed and initially validated an indirect forgiveness measure — the forgiveness Implicit Association Test (forgiveness IAT; Fatfouta, Schröder-Abé, & Merkl, 2014). Specifically, Study 1 examined the differential contribution of implicit and explicit forgiveness to the prediction of an aspect of forgiveness-related behavior — the response time (RT) with which individuals generated responses to state forgiveness ratings following a transgression recall (faster responses were taken to reflect greater forgiveness; see section Methodology below). Thus, the aim of Study 1 was to test the forgiveness IAT's incremental validity beyond a corresponding self-report measure. Study 2 (Fatfouta, 2015) built on Study 1 by scrutinizing the processes underlying forgiveness-response decisions. In particular, Study 2 probed whether individuals high in forgiveness are faster to rate their state forgiveness, because their reduced ruminative tendencies may enable better inhibition of transgression-related information in memory. Study 2 thus aimed to provide initial evidence for the mediating role of rumination about the transgression in the forgiveness/RT link.

2.2. Characterize the mechanisms behind narcissists' lack of forgiveness

Previous research has usually conceived of narcissism as a unitary construct, yet more recent approaches advocate a multidimensional approach. Study 3 (Fatfouta, Gerlach, Schröder-Abé, & Merkl, 2015) analyzed the facet-specific contribution of two distinct

narcissism dimensions (i.e., admiration and rivalry) to lack of forgiveness (i.e., high revenge and avoidance motivation toward the transgressor). Moreover, Study 3 investigated how three socio-cognitive variables (i.e., state anger, state rumination, and state empathy) that can be prompted by a transgression mediate the narcissism/lack of forgiveness link. In doing so, Study 3 aimed to provide evidence for the notion that these socio-cognitive variables are key factors underlying narcissists' lack of forgiveness. Study 4 (Fatfouta & Gerlach, under review), by contrast, inspected conditions under which narcissists — especially those high in rivalry — can be appeased. Stated differently, Study 4 aimed to test the moderating role of conciliatory gestures (i.e., apology/amends) in the narcissism/lack of forgiveness link.

2.3. Examine the link between OC symptoms and revenge

Whereas anger has only recently been linked to OC symptoms, it remains unclear how these symptoms relate to the outcome of unresolved anger, that is, revenge. Study 5 (Fatfouta & Merkl, 2014) sought to close this gap by examining how OC symptomatology would relate to various aspects of revenge (i.e., vengeful attitudes, motivations, and dispositions). Based on previous studies in psychiatric research, Study 5 also tried to answer the question of whether individuals high in OC symptoms would perceive interpersonal transgressions more frequently in their daily interactions.

2.4. Identify a neural mechanism of forgiveness

Previous neuroscience research has suggested that forgiveness involves aspects of conflict detection (represented by the DACC) and mentalizing (involving the MPFC). A neurobiological mechanism explaining the interplay of MPFC/DACC in instigating forgiveness is currently lacking, especially in the context of socially close relationships. How does the brain implement the decision to forgive a loved one as compared to a stranger when being treated unfairly? Study 6 (Fatfouta, Meshi, Merkl, & Heekeren, under review) combined behavioral and neuroimaging techniques in an interactive social exchange to address this question. Specifically, Study 6 examined which brain areas are involved when individuals experience transgressions by a close other (i.e., one's partner) and by a non-close other (i.e., an unknown person). Moreover, Study 6 analyzed the relationship between task-related changes in brain activity and individual differences in forgiveness, as assessed by the aforementioned forgiveness IAT and a corresponding self-report.

3. Methodology

In this chapter, I will briefly describe the methodology of the six empirical studies constituting this thesis (Fatfouta, 2015; Fatfouta & Gerlach, under review; Fatfouta et al., 2015; Fatfouta & Merkl, 2014; Fatfouta et al., under review; Fatfouta et al., 2014). Given the methodological diversity among studies, I will focus on the most pertinent materials. Specifically, subsection 3.1. (General methodology) describes methods that have been used in at least two research studies and lists the most important questionnaires; subsection 3.2. (Specific methodology) describes methods that have been used in study six. For a complete description of the respective methodological details, see the Method sections of each study (Appendix).

3.1. General methodology

3.1.1. Forgiveness Implicit Association Test

In the forgiveness IAT used in Studies 1 and 6, participants were required to classify stimulus words as quickly and accurately as possible into one of two categories. These categories varied across five blocks: In blocks 1 and 2, participants practiced classifying stimuli from the target category (me-others) or attribute category (forgiving-vengeful). In block 3, both discrimination tasks were combined (i.e., me-forgiving share one key; others-vengeful share the other key). In block 4, the labels of the attribute category were reversed (vengeful-forgiving). In block 5, participants completed the reversed combined discrimination task (i.e., me-vengeful share one key; others-forgiving share the other key). Figure 1 details an example trial of the forgiveness IAT.

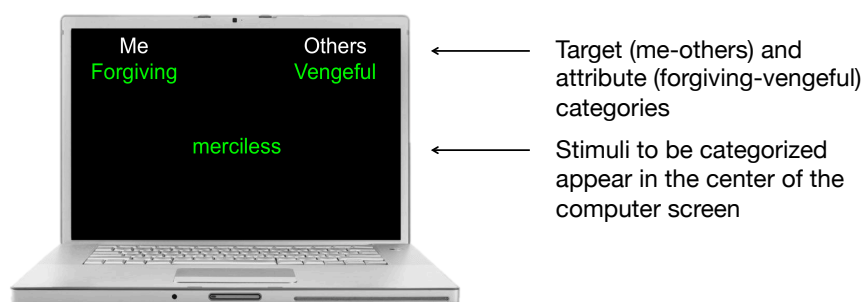


Figure 1. Example trial during block 3 of the forgiveness IAT. Participants responded to each stimulus word by pressing either a left or right key. After a correct key press, the next trial began with a new stimulus word.

We kept the order of critical blocks (i.e., blocks 3 and 5) constant to optimize the reliability of individual differences (Banse, 2001). However, a constant block order can inflate IAT scores (so-called compatibility-order effect; Greenwald, Nosek, & Banaji, 2003). We therefore doubled the number of trials in the reversed attribute-discrimination task (i.e., block 4) — a procedure known to mitigate the compatibility-order effect (Nosek, Greenwald, & Banaji, 2005). We computed IAT scores using the D_1 algorithm proposed by Greenwald et al. (2003). Higher scores reflect a stronger automatic association between me-forgiving relative to me-vengeful (i.e., higher implicit forgiveness).

3.1.2. Transgression-recall paradigm

In the transgression-recall paradigm used in Studies 1 through 5, participants were required to (1) relive an interpersonal transgression and, when ready, to (2) rate their current thoughts and feelings toward the person who had hurt them (i.e., their state forgiveness). In particular, participants completed a computer-based questionnaire containing multiple transgression-related items (e.g., "I am finding it difficult to act warmly toward him/her"; McCullough et al., 2006, p. 897). In Studies 1 and 2, we complemented the recall paradigm by measuring the time with which participants generated responses to state forgiveness ratings. Consistent with research on attitude accessibility (Fazio, 1990a, 1990b), shorter RTs were conceptualized as an indication of greater forgiveness-cue accessibility. In these studies, each item appeared separately on the computer screen. Participants then had to indicate their responses via a click on the respective response category. The next item appeared either by clicking on a "next" button (Study 1) or immediately following response selection (Study 2). RT was unobtrusively recorded as the time between item presentation and that click. Importantly, participants had no prior knowledge of the RT measurement. We analyzed RT data using standard procedures (e.g., accounting for positive skewness; Fazio, 1990b; Ratcliff, 1993) and, most importantly, controlling for individual differences in general (i.e., baseline) speed (Lischetzke, Angelova, & Eid, 2011; Lischetzke, Cuccodoro, Gauger, Todeschini, & Eid, 2005). Stated differently, we made sure that differential RTs in the transgression-recall paradigm were not confounded by construct-irrelevant variance due to reading speed and/or motor behavior.

3.1.3. Questionnaires

This section provides an overview of the most important questionnaires used in the six studies. For details on these questionnaires, see the Method sections of each study (Appendix).

The Transgression-Related Interpersonal Motivations Inventory (TRIM; McCullough et al., 1998; McCullough et al., 2006) was used to assess individuals' motivations toward their transgressors following the transgression recall (i.e., their state forgiveness; see 3.1.2.). The TRIM consists of three subscales: revenge, avoidance, and benevolence. Higher levels of revenge and avoidance indicate a lack of forgiveness toward the transgressor, whereas higher levels of benevolence indicate more conciliation and goodwill. We used the TRIM in Studies 1 through 5.

The Tendency to Forgive Scale (TTF; Brown, 2003) was used to assess individuals' general (i.e., cross-situational) tendency to forgive transgressions (i.e., their trait forgiveness). Higher TTF scores indicate greater trait forgiveness. We used the TTF in Studies 1, 2, 5, and 6. In Study 5, we focused on forgiveness' conceptual foil, namely revenge. We therefore reverse-coded two of the four TTF items in Study 5 (for details, see the Method section of Study 5, Appendix).

The Narcissistic Admiration and Rivalry Questionnaire (NARQ; Back et al., 2013) was used to assess individuals' narcissism levels. The NARQ consists of two subscales: admiration and rivalry. Higher levels of admiration reflect a tendency toward assertive self-promotion, whereas higher levels of rivalry reflect a tendency to toward antagonistic self-protection. We used the NARQ in Studies 3 and 4. In particular, we calculated uniqueness scores by partialling each of the two narcissism facets (i.e., admiration and rivalry) out of the other one (for a similar approach, see Penke & Asendorpf, 2008).

The Obsessive-Compulsive Inventory-Revised (OCI-R; Foa et al., 2002) was used to assess individuals' degree to which they have been bothered by OC-related symptoms in the past month. The OCI-R consists of six subscales: washing, checking, ordering, obsessing, hoarding, and neutralizing. Higher OCI-R scores indicate higher OC symptom severity. We used the OCI-R in all four substudies of Study 5.

3.2. Specific methodology

3.2.1. Real-life interaction and ultimatum game task

Study 6 combined a real-life interaction with an experimental design, in which romantically involved couples played a social-exchange game, the Ultimatum Game (UG; Güth, Schmittberger, & Schwarze, 1982). In the UG, one player (the proposer) offers how to split a sum of money (e.g., 10€) with another player (the responder). If the responder accepts the offer, the sum is divided according to the proposal. However, if the responder rejects the offer, neither player receives anything. Responders typically reject offers < 20-30% of the

proposer’s endowment and, hence, retaliate for low offers (Camerer, 2003). Indeed, such rejection behavior has been shown to reflect retaliatory motives (Crockett et al., 2013; Fehr & Gächter, 2000). However, responders can also accept those offers and, hence, refrain from retaliation (i.e., forgive).

In our UG task, participants believed they had been randomly assigned to be the “proposer” and their partner had been assigned to be the “responder”. In reality, however, all participants were responders. As responders, they decided whether to accept or reject offers from their partner or an unknown person (gender-matched to the partner). Participants also played with a non-human (i.e., computer) proposer (not relevant to the present research question). Participants believed that the offers were real (i.e., actually being made during the experimental session). In reality, however, all offers were predetermined to establish experimental control. Offers ranged from fair (offer 4€ or 5€ out of 10€) to unfair (offer 3€, 2€, or 1€ out of 10€). During the UG, we scanned one member of each couple with functional magnetic resonance imaging (fMRI; see subsection below). The respective partner performed a behavioral version of the task in an adjacent testing room. Study 6 therefore allowed us to investigate how relationship closeness modulates forgiveness and its neural mechanism following transgressions (in our case, unfairness in the UG).

Participants completed 180 trials divided across three conditions: PARTNER (75 trials), UNKNOWN (75 trials), and COMPUTER (30 trials). Each trial began with a fixation cross, followed by a picture of the proposer, and the offer. While the offer was presented, participants had to indicate their decision (accept or reject) via button press. Finally, the outcome for the current round was presented. Figure 2 details the experimental design of a single trial in the UG.

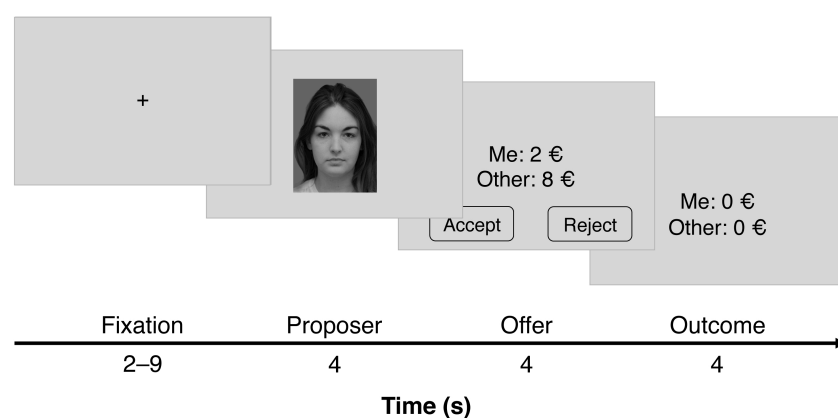


Figure 2. Timeline of a single trial in the Ultimatum Game. In each trial, participants first viewed a fixation cross, then a photograph of the proposer, the offer, and then the outcome. While the offer was on screen, participants responded whether they accepted or rejected it by button press.

For behavioral analyses, we examined individuals' acceptance rates (% of accepted offers) using a 2 (fair vs. unfair) by 2 (PARTNER vs. UNKNOWN) repeated-measures analysis of variance, with fairness level (fair vs. unfair) and relationship closeness (PARTNER vs. UNKNOWN) as within-subject factors.

3.2.2. Functional magnetic resonance imaging

As mentioned above, Study 6 used fMRI to identify a neural mechanism of forgiveness. Briefly, fMRI relies on task-related changes in the magnetic properties of hemoglobin and, more specifically, changes in blood oxygen level dependent (BOLD) signal. The BOLD signal thus provides an indirect measure of neural activity (for a review, see Logothetis, 2008). One key advantage of fMRI over other neuroimaging methods (e.g., positron emission tomography) is its noninvasiveness. Furthermore, and important to the study of forgiveness, fMRI allows for the investigation of phenomena that people are unwilling or incapable to disclose when asked to do so (Aue et al., 2009).

For MRI data acquisition, we used a 3 Tesla system (Siemens Magnetom Trio; Erlangen, Germany) located at the Dahlem Institute for Neuroimaging of Emotion (D.I.N.E., <http://www.loe.fu-berlin.de/dine/>), Freie Universität Berlin, Germany. We acquired anatomical images using a T1-weighted magnetization-prepared rapid acquisition with gradient echo (MP-RAGE) sequence (256 x 256 matrix, 176 sagittal slices, slice thickness = 1 mm). We acquired functional images using T2*-weighted Echo Planar Imaging (EPI; time of repetition [TR] = 2,000 ms; echo time [TE] = 30 ms; 64x64 matrix; flip angle = 70°; field of view = 192 mm; interslice gap = 0.6 mm). We collected a total of 37 axial slices (3 x 3 x 3 mm voxels) parallel to the anterior and posterior commissure per volume (545 volumes/run; three runs per participant).

For fMRI data preprocessing and analysis, we used FMRIB's Software Library, version 5.0.6 (FSL; Smith et al., 2004). We preprocessed the fMRI data using standard procedures, including non-brain tissue removal (Smith, 2002), motion correction (Jenkinson, Bannister, Brady, & Smith, 2002), spatial smoothing (6-mm full-width at half-maximum [FWHM] Gaussian kernel), high-pass temporal filtering (cutoff = 80 s), registration to each individual's structural image using boundary-based registration (Greve & Fischl, 2009), and normalization into standard space (Montreal Neurological Institute [MNI]) using linear registration with 12 degrees of freedom (Jenkinson et al., 2002; Jenkinson & Smith, 2001). We then estimated a general linear model (GLM) of the BOLD signal using event-related regressors for each condition (for a list of all regressors, see Method section of Study 6,

Appendix). Individual level contrasts were computed and then entered into a higher-level mixed-effects analysis using voxel-wise one-sample *t* tests.

First, we examined brain regions showing greater activity for unfair versus fair exchanges from human proposers (contrast [unfair PARTNER + unfair UNKNOWN > fair PARTNER + fair UNKNOWN]). Second, we investigated changes in brain activity in the context of unfair treatment by one's partner relative to unfair treatment by an unknown person [unfair PARTNER > unfair UNKNOWN]. Third, to determine task-dependent changes in functional connectivity, we performed a psychophysiological interaction (PPI) analysis (Friston et al., 1997; O'Reilly, Woolrich, Behrens, Smith, & Johansen-Berg, 2012). PPI analysis allows for identifying voxels in the brain whose activity depend on the interaction between two factors (O'Reilly et al., 2012): a psychological factor (i.e., the task of interest) and a physiological factor (i.e., the time course of a given seed region). Here, the task of interest was the contrast [unfair PARTNER > unfair UNKNOWN] and the seed region was the MPFC based on a priori hypotheses regarding its role in forgiveness (e.g., Young & Saxe, 2009). Finally, we were also interested in the relationship between brain activity and forgiveness in the UG as well as individual differences in forgiveness, assessed outside the MRI scanner. To assess forgiveness in the UG, we calculated the absolute difference between participants' acceptance of unfair partner exchanges and unfair unknown exchanges (Δ forgiveness index = $|\% \text{ of accepted unfair exchanges}_{\text{PARTNER}} - \% \text{ of accepted unfair exchanges}_{\text{UNKNOWN}}|$). Values greater (smaller) than zero indicate that an individual accepts more (less) unfair offers from his/her partner than from the unknown person. Specifically, the forgiveness index (Δ) reflects the tendency to forgive one's partner (vs. the unknown person). To assess individual differences in forgiveness, we used implicit (forgiveness IAT) and explicit (TTF scale) measures. We then regressed the extracted parameter estimates from the PPI analysis onto these variables. For more details on the specific fMRI hypotheses and the PPI analysis, see Method section of Study 6 (Appendix).

4. Summary of empirical studies

In this chapter, I will briefly summarize the six empirical studies constituting this thesis. For the complete research articles, see Appendix.

4.1. Study 1: Forgiving, fast and slow: Validity of the implicit association test for predicting differential response-latencies in a transgression-recall paradigm

Background and objective: Forgiveness has traditionally been assessed via self-report, yet recent research highlights the value of indirect measures. Indirect measures tap into more automatic (i.e., non-deliberative) aspects of information processing and have been shown to provide increments in predictive validity. However, no studies to date have examined such an indirect measure of forgiveness. To fill this gap, we developed an Implicit Association Test (IAT) of forgiveness and examined its reliability, test-retest stability, and incremental validity over a direct (i.e., self-report) measure of forgiveness. As a forgiveness-related behavioral criterion, we used the time with which individuals rated their thoughts and feelings toward a person who had hurt them (i.e., their state forgiveness).

Method: Participants ($N = 104$) first completed the forgiveness IAT and then a computer-based questionnaire assessing their baseline speed of responding, their state forgiveness (while RT was unobtrusively recorded), and their perceived severity of the transgression. Additionally, participants completed measures of explicit forgiveness and social desirability. To examine temporal stability, the last thirty participants were re-contacted after one month to complete the forgiveness IAT in a second testing session.

Results: We obtained three main results: First, implicit (IAT-measured) and explicit (self-reported) forgiveness were not significantly correlated. Second, implicit and explicit forgiveness predicted shorter RT of state forgiveness ratings. That is, those high in explicit as well as implicit forgiveness were faster to describe and rate their thoughts and feelings toward the person who had hurt them. Third, implicit forgiveness demonstrated high split-half reliability, moderate stability over time and, even more important, evidence of incremental validity over explicit forgiveness. In addition, implicit (but not explicit) forgiveness was unrelated to social desirability.

Conclusion: These results suggest that RT of state forgiveness ratings is informative when examining dynamic aspects of information processing related to the transgression recall, further corroborating the notion of implicit forgiveness. The newly developed forgiveness IAT may advance forgiveness research beyond what is known from self-report.

4.2. Study 2: How forgiveness affects processing time: Mediation by rumination about the transgression

Background and objective: In the previous study (Fatfouta et al., 2014), we have demonstrated that individuals high (vs. low) in forgiveness are faster to describe and rate their state forgiveness. So far, however, the mechanism through which forgiveness is related to RT remains unknown. To address this open question, a mediation model is proposed wherein the relationship between forgiveness and RT would be mediated by rumination about the transgression. Specifically, individuals high in forgiveness should ruminate less about their transgression, which in turn, should promote faster information processing (i.e., shorter latencies).

Method: The procedure was similar to that used by Fatfouta et al. (2014) except that this study was conducted online¹. Participants ($N = 767$) completed a questionnaire assessing their trait forgiveness, their baseline speed of responding, their rumination about a recalled transgression, and their state forgiveness (while RT was unobtrusively recorded).

Results: Results from a bootstrapping mediation analysis (5,000 samples, bias-corrected confidence interval [CI] method; Hayes, 2013) replicated our initial findings (Fatfouta et al., 2014) — trait forgiveness predicted shorter RT of state forgiveness ratings (Fig. 3A). That is, those high in forgiveness were faster to rate their thoughts and feelings toward their transgressor. Furthermore, rumination about the transgression predicted longer RT of state forgiveness ratings (Fig. 3B). That is, those who ruminated more about the transgression were slower to rate their thoughts and feelings toward the person who had hurt them. Finally, and as hypothesized, rumination about the transgression mediated the relationship between trait forgiveness and RT (Fig. 4).

Conclusion: These findings extend those of Fatfouta and colleagues (2014) and provide important information regarding rumination about the transgression as a mediator of the forgiveness-RT link. The mediation analysis indicates that RT of state forgiveness ratings may be better characterized as reflecting a person's reduced ruminative tendency rather than an enhanced accessibility of forgiveness-related cognitions in memory.

¹At the time where Study 2 was conducted, the IAT module was not implemented in the online-survey tool. Therefore, the forgiveness IAT could not be applied in Study 2.

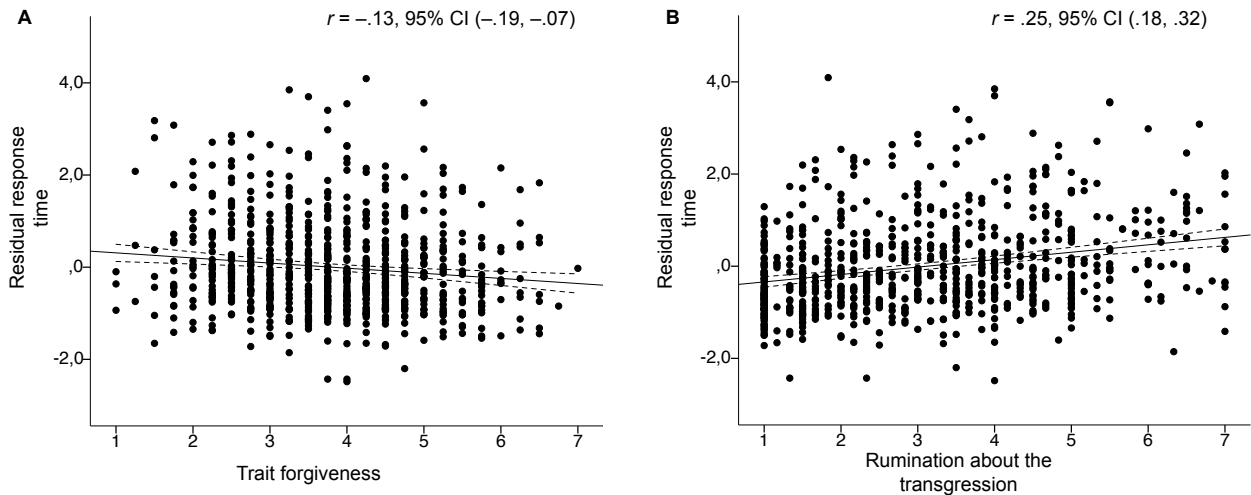


Figure 3. Scatterplots of the forgiveness/rumination-response time relationship. The panels show the relationship between trait forgiveness (A) as well as rumination about the transgression (B) and residual (i.e., baseline-corrected) response time. Dashed lines indicate the 95% confidence intervals (CIs) for the regression line.

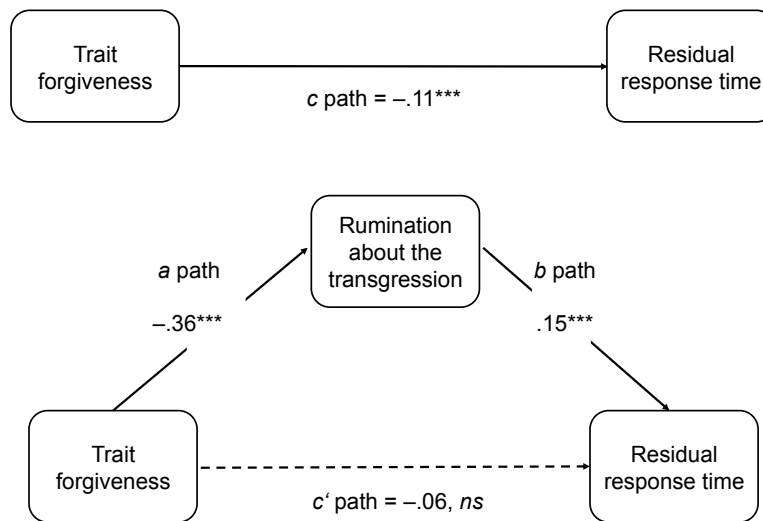


Figure 4. Mediation model illustrating the direct and indirect effects of forgiveness on response time. Path c is the total effect of trait forgiveness on response time (sum of direct and indirect effects: $c = c' + ab$); c' is the direct effect of trait forgiveness on response time; a is the direct effect of trait forgiveness on rumination; b is the direct effect of rumination on response time. When rumination was included in the model the direct effect of trait forgiveness on response time (dashed line) was no longer significant. Coefficients represent unstandardized regression coefficients.

*** $p < .001$, two-tailed, ns = not significant.

4.3. Study 3: Narcissism and lack of interpersonal forgiveness: The mediating role of state anger, state rumination, and state empathy

Background and objective: As summarized in the introduction (subsection 1.3.), narcissism plays a key role in the psychology of forgiveness. Previous research, however, has usually relied on a one-dimensional conceptualization of narcissism. This leaves open the question of how different narcissism dimensions relate to lack of forgiveness (i.e., revenge and avoidance). Furthermore, the mechanisms underlying the narcissism/lack of forgiveness link are largely unexplored. To fill these gaps, we investigated how two facets of narcissism — admiration and rivalry — are related to lack of forgiveness. In addition, we examined how different socio-cognitive variables, namely state anger, state rumination, and state empathy, mediate these associations.

Method: Participants ($N = 1,040$) were recruited online to complete the NARQ (see subsection 3.1.3.) and were then instructed to recall a transgression in which someone had hurt them. Subsequently, participants rated their cognitions and affective state, including anger toward the transgressor, ruminative thoughts about the transgression, and empathic feelings for the transgressor. Finally, participants rated their lack of forgiveness.

Results: Using bootstrapping mediation analysis (5,000 samples, bias-corrected CI method; Hayes, 2013), we obtained the following main results: First, narcissistic admiration was unrelated to lack of forgiveness. However, empathy for the transgressor acted as a suppressor variable in the mediation models. In particular, those high in admiration experienced greater empathy for their transgressor, which in turn predicted less revenge and avoidance (Fig. 5A). Second, narcissistic rivalry was positively related to lack of forgiveness and this relationship was mediated by greater anger and rumination and less empathy (Fig. 5B).

Conclusion: These results suggest that the mechanisms underlying narcissism and lack of forgiveness are better understood by disentangling the two dimensions of narcissism (i.e., admiration and rivalry) rather than by conflating them into a single construct. Additionally, findings highlight the importance of considering narcissists' cognitive and affective responses that may be prompted by a transgression incident.

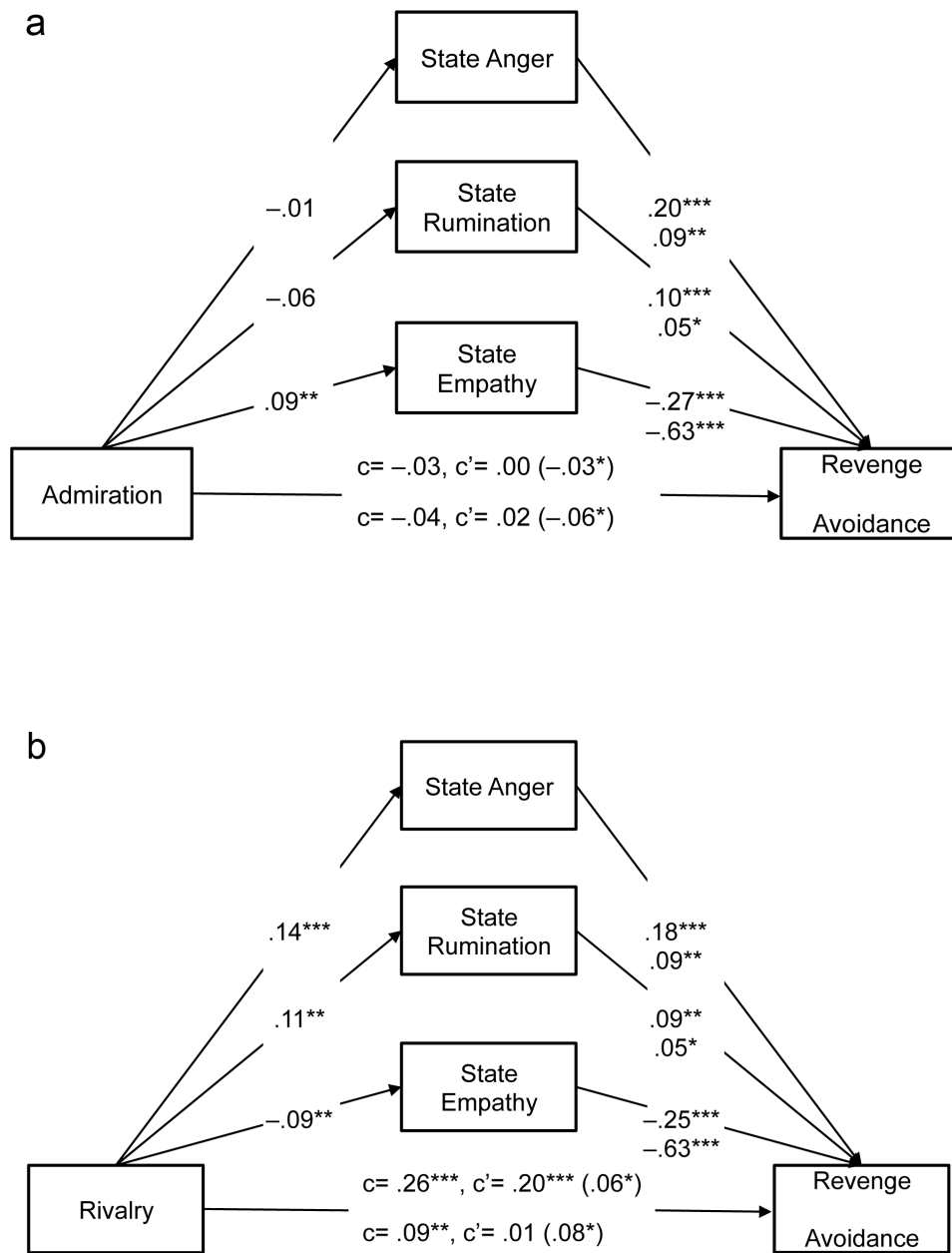


Figure 5. Mediation models illustrating the direct and indirect effects of narcissism on lack of forgiveness. Depicted are the unique effects of admiration (panel a) and rivalry (panel b) on forgiveness facets (revenge and avoidance) via state anger, state rumination, and state empathy. Coefficients represent standardized regression coefficients while controlling for gender and age (not shown in figure); *c* is the total effect of each narcissism facet on revenge/avoidance, and *c'* is the direct effect after correcting for proposed mediators; total indirect effects are in parentheses.
N = 1,040. **p* < .05, ***p* < .01, ****p* < .001 (two-tailed).

4.4. Study 4: Can the shrew be tamed? Narcissism, lack of forgiveness, and the moderating role of conciliatory gestures

Background and objective: Narcissists, specifically those high in rivalry, have shown to be less likely to forgive interpersonal transgressions (Back et al., 2013; Fatfouta et al., 2015). However, little is known about the factors that might change the narcissism/lack of forgiveness link. Previous studies hypothesized conciliatory gestures (i.e., apologies/offers of compensation) to mitigate individuals' lack of forgiveness, yet no studies to date have investigated whether apology moderates the relationship between narcissism and lack of forgiveness. We addressed this gap by conducting one correlational and two experimental studies.

Method: Across all three studies, participants (total $N = 2,127$) were recruited online to complete the NARQ (see subsection 3.1.3.). Within each study, apology was measured or manipulated as follows: In Study 1 ($N = 767$), we measured perceptions of apology following a real-life transgression via self-report. In Study 2 ($N = 930$), we manipulated the presence (vs. absence) of an apology following a hypothetical transgression via an experimental design. In Study 3 ($N = 430$), we extended the manipulation to encompass specific conciliatory gestures (verbal expression of regret, verbal expression of regret and offer of compensation, verbal and nonverbal expression of regret, or verbal and nonverbal expression of regret and offer of compensation).

Results: Across all three studies, narcissistic rivalry (but less so admiration) was positively related to lack of forgiveness. Apology, by contrast, was negatively related to lack of forgiveness (over and above the effects rivalry and admiration, respectively). Furthermore, the relationship between narcissistic rivalry and revenge appeared to be robust to attempts by the transgressor to apologize or to compensate. However, and importantly, the relationship between rivalry and avoidance was attenuated when the transgressor offered a costly apology to the victim in addition to a verbal expression of regret (Fig. 6).

Conclusion: These results provide the first evidence that narcissists' tendency to avoid the transgressor can be partly reduced, provided that one is willing to incur a personal cost. Moreover, these results again accentuate the value of a two-dimensional conceptualization of narcissism, as it allows for a differentiated view on narcissism's interplay with conciliatory gestures following interpersonal hurt.

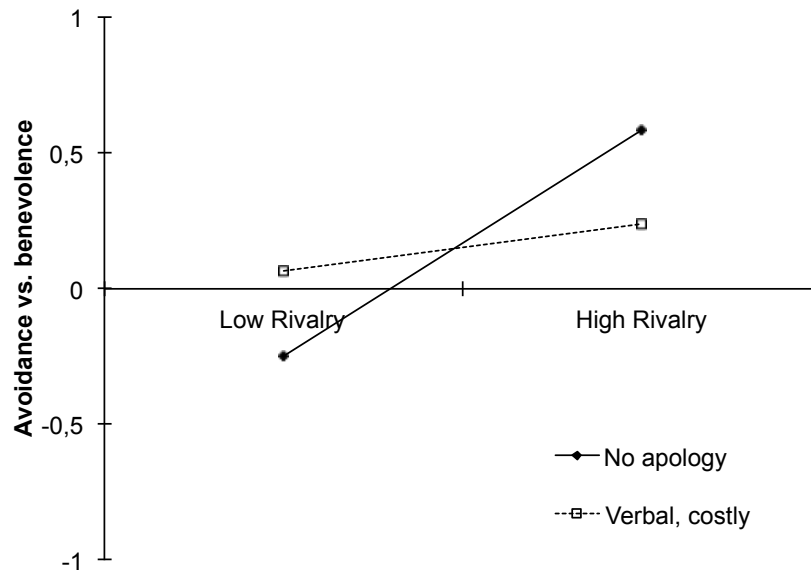


Figure 6. Effect of apology on narcissists' lack of forgiveness. Depicted are predicted values for avoidance versus benevolence as a function of narcissistic rivalry (± 1 *SD* of the mean) and apology.

4.5. Study 5: Associations between obsessive-compulsive symptoms, revenge, and the perception of interpersonal transgressions

Background and objective: Heightened anger and aggression have previously been found in individuals with OCD. How the outcome of unresolved anger (in particular, revenge) relates to OC symptoms remains, however, unknown. It also remains unknown how OC symptoms relate to the perception of interpersonal transgressions. Do individuals high in OC symptoms perceive transgressions more frequently in their everyday social life? We addressed these gaps by conducting four studies. Specifically, we systematically examined the associations between OC symptoms and three aspects of revenge (attitudes, dispositions, and motivations; Studies 1-3, respectively) as well as transgression frequency (Study 4).

Method: Across all four studies, participants (total $N = 679$) were recruited online to complete the OCI-R (see subsection 3.1.3.). Within Studies 1-3, revenge was measured as follows: In Study 1 ($N = 153$), we measured revenge-seeking attitudes. In Study 2 ($N = 178$), we measured dispositional revenge. In Study 3 ($N = 173$), we measured motivations to retaliate against a real-life transgressor. In Study 4 ($N = 175$), we measured participants' frequency of interpersonal transgression occurrence during the past 14 days.

Results: Individuals high in OC symptoms reported more positive attitudes toward revenge (Study 1), higher levels trait revenge (Study 2), and heightened revenge motivation

toward a real-life transgressor (Study 3). Importantly, the most robust association appeared between obsessions (i.e., intrusive and repetitive thoughts) and revenge. Moreover, individuals high in OC symptoms perceived interpersonal transgressions more frequently in their daily lives (Study 4).

Conclusion: These results suggest that revenge and interpersonal hurt may advance our understanding of OC symptoms. When confirmed in clinical studies, targeting revenge phenomena in psychotherapy may prove beneficial.

4.6. Study 6: Forgiveness involves reduced connectivity between medial prefrontal and dorsal anterior cingulate cortex

Background and objective: Neuroimaging studies suggest that forgiveness involves mentalizing-related and conflict-related components. However, a comprehensive brain mechanism of forgiveness remains unexplored — specifically, in the context of close interpersonal relationships. To fill this gap, we used fMRI to investigate changes in brain activity when individuals face transgressions by a close other (i.e., their own partner) relative to a non-close other (i.e., an unknown person). In particular, we examined whether forgiveness depends on interactions between mentalizing-related regions (in particular, the MPFC) and regions modulated by conflicting response tendencies (in particular, the DACC). To further determine how participants' brain activity predicts individual differences in forgiveness, we used measures of implicit (forgiveness IAT) and explicit (TTF scale) forgiveness as outcome criteria.

Method: Twenty-three romantically involved couples (final $N = 46$) participated in this study. Forgiveness was operationalized as an individual's tendency to refrain from retaliation following unfair treatment in an iterated social-exchange game, the UG (for details, see Fig. 2, subsection 3.2.1.). All participants played the UG, but only one member of each couple ($n = 23$) was scanned with fMRI while playing the UG (for scanning parameters, see subsection 3.2.2.). After completing the UG, all participants rated (1) the perceived fairness, attractiveness, and trustworthiness of their partner (PARTNER), the unknown person (UNKNOWN), and a gender-matched person who had participated in the previous testing session (OTHER); (2) their affective responses to unfairness and, finally, (3) completed the previously developed forgiveness IAT and the TTF scale.

Results: Behavioral results revealed that participants were more likely to forgive their partner than the unknown person, as indexed by a significantly higher proportion of accepted

unfair exchanges (Fig. 7). A control analysis excluded the possibility that this effect was due to differences in how resources were actually being shared with one's partner (i.e., higher acceptance rates for partner exchanges were not driven by couples' sharing of money in the relationship). Emotion ratings revealed, among others, that participants experienced greater forgiveness and benevolence in response to unfair partner (vs. unknown) treatment (Fig. 8). Moreover, participants perceived their partner to be fairer, more attractive, and more trustworthy than the unknown person (Fig 9; black and white bars); however, the romantic partners and the unknown persons were comparable regarding these characteristics (Fig 9; white and hatched bars). Neuroimaging results revealed a significant BOLD signal change in bilateral anterior insula (AI), dorsal anterior cingulate cortex (DACC), and right dorsolateral prefrontal cortex (DLPFC) when being treated unfairly by a human proposer (Fig. 10, A and B). Comparing the neural correlates of unfairness separately for the partner and the unknown person revealed a regional overlap in right insular cortex, middle occipital gyrus, and DACC (Fig. 11). Importantly, brain regions showing greater activation for unfair treatment by one's partner than by an unknown person revealed only one significant cluster in the MPFC (Fig. 12, A and B). A reverse-inference analysis (Yarkoni, Poldrack, Nichols, Van Essen, & Wager, 2011) indicated that the observed task-related change in MPFC activity likely reflected mentalizing-related processes (Fig. 12C). To identify brain regions exhibiting a change in functional connectivity with the MPFC depending on whether the partner or the unknown person behaved unfairly, we performed a PPI analysis (O'Reilly et al., 2012), which revealed a significant cluster within the DACC (Fig. 13A). This MPFC/DACC functional connectivity, in turn, was negatively related to participants' forgiveness index, defined as the absolute difference between accepted unfair partner exchanges and accepted unfair unknown exchanges (Fig. 13B). That is, the lower the task-related change in MPFC/DACC connectivity, the higher the individual's acceptance of unfair partner (vs. unfair unknown) exchanges (i.e., the greater the individual's forgiveness). Importantly, this pattern was similar for the forgiveness IAT assessed outside the scanner: lower task-related change in MPFC/DACC connectivity inversely predicted participants' IAT scores (i.e., greater implicit forgiveness). A leave-one-subject out cross-validation excluded the possibility that this result was due to a selection bias. By contrast, MPFC/DACC connectivity did not predict participants' TTF scores (Fig. 13C).

Conclusion: These results provide the first evidence for a neural mechanism of forgiveness in the context of close relationships. Having demonstrated that relationship

closeness modulates prefrontal brain function during interactive decision-making, this study advances our understanding of the context-dependent nature of social behavior.

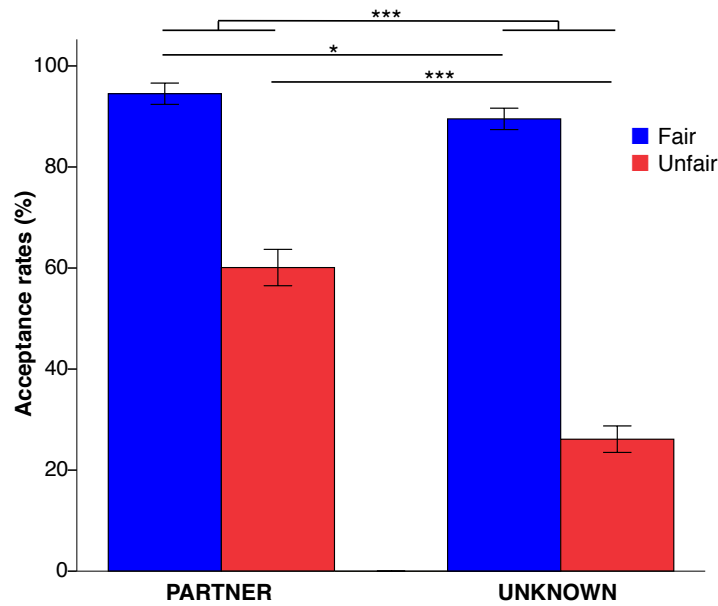


Figure 7. Behavioral results in the UG task. Mean acceptance rates (% of accepted offers) for fair and unfair offers from the partner and the unknown person. Error bars indicate one standard error of the mean, calculated within-subjects.

*** $P < 0.001$ (two-tailed), * $P < 0.05$ (two-tailed).

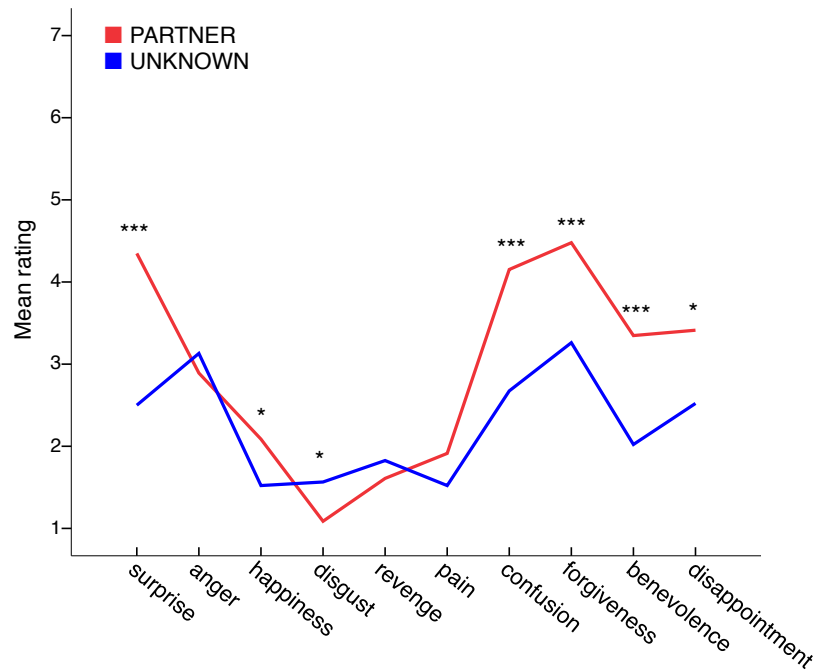


Figure 8. Affective responses to unfairness. Mean self-reported emotion ratings (1 = not at all, 7 = very much) in response to unfair offers in the UG, by proposer.

*** $P < 0.001$ (two-tailed), * $P < 0.05$ (two-tailed).

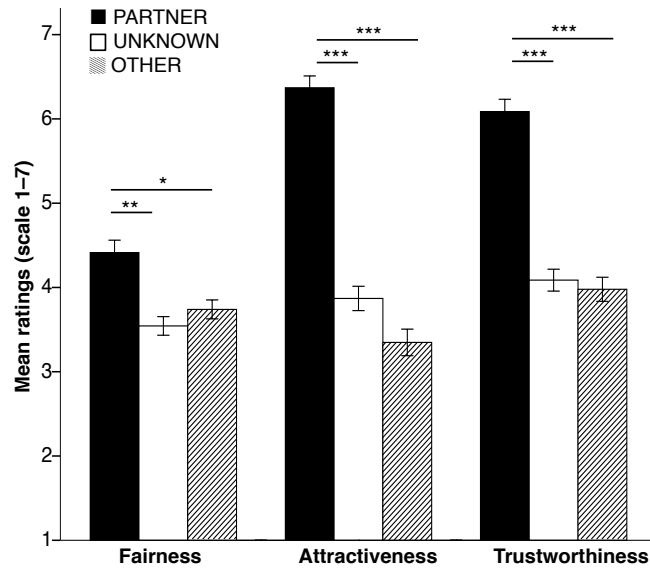


Figure 9. Fairness, attractiveness, and trustworthiness ratings. Mean rating scores for perceived fairness, trustworthiness, and attractiveness, by person. Persons included the romantic partner (PARTNER), the unknown person (UNKNOWN), and the person from the previous experimental session (OTHER). Error bars indicate one standard error of the mean, calculated within-subjects. *** $P < 0.001$ (two-tailed), ** $P < 0.005$ (two-tailed), * $P < 0.05$ (two-tailed).

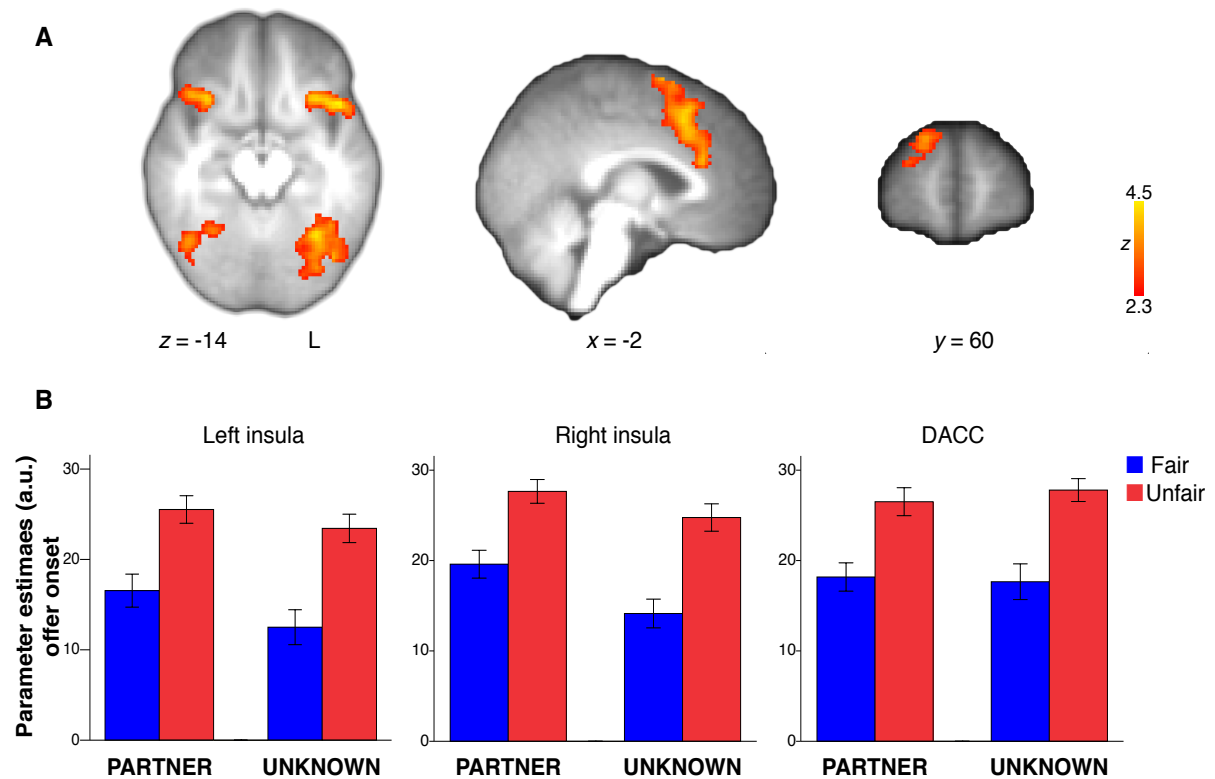


Figure 10. BOLD signal change related to unfairness. (A) Significant changes in BOLD signal for the contrast [unfair PARTNER + unfair UNKNOWN > fair PARTNER + fair UNKNOWN] were found in bilateral AI, DACC, right cerebellum, left middle occipital gyrus, and right DLPFC. For a complete list of results, see Table 2 in the Results section of Study 6. Results were thresholded at $z > 2.3$, $P < 0.05$, cluster corrected. L = left hemisphere. (B) Mean parameter estimates in arbitrary units (a.u.) within bilateral AI and DACC for fair and unfair offers from the partner and the unknown person. Error bars indicate one standard error of the mean, calculated within-subjects.

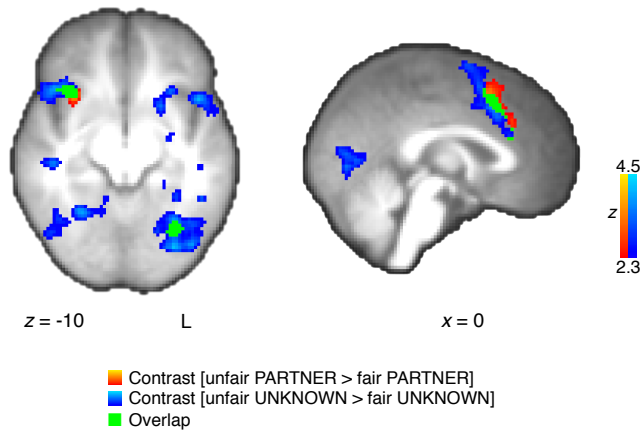


Figure 11. Overlap of BOLD signal change related to unfairness. Comparison of z statistic images of unfair offers from each proposer [unfair PARTNER > fair PARTNER in red-yellow; unfair UNKNOWN > fair UNKNOWN in blue-light blue]. Overlapping activations are displayed in green. The overlap included areas in right insular cortex, left middle occipital gyrus, and DACC. Results were thresholded at $z > 2.3$, $P < 0.05$, cluster corrected.

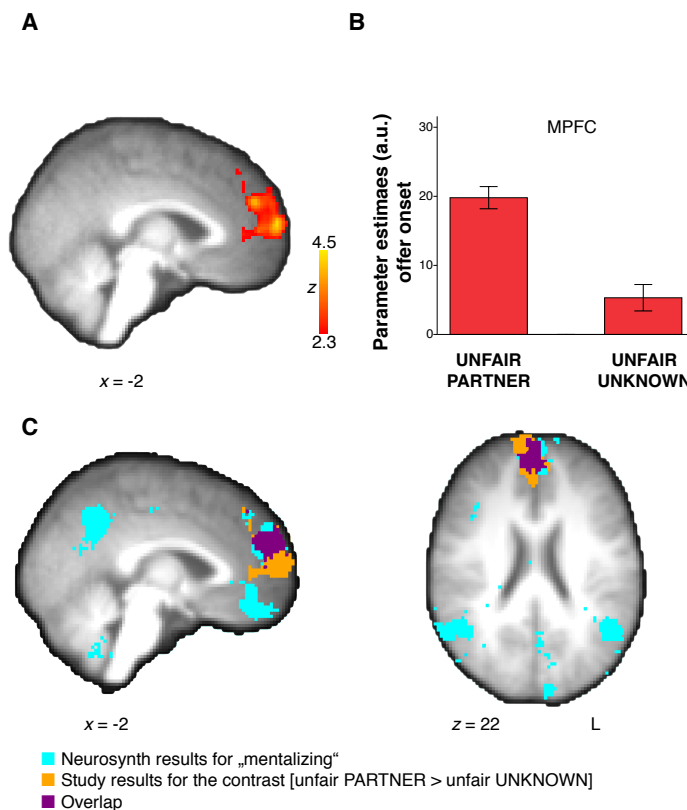


Figure 12. Brain activation associated with unfairness by the partner. (A) Significant changes in BOLD signal for the contrast [unfair PARTNER > unfair UNKNOWN] were found in MPFC. Results were thresholded at $z > 2.3$, $P < 0.05$, cluster corrected. (B) Mean parameter estimates in arbitrary units (a.u.) within MPFC for unfair offers from the partner and the unknown person. Error bars indicate one standard error of the mean, calculated within-subjects. (C) Comparison of Neurosynth meta-analysis search for the term “mentalizing” (light blue) to study results for the contrast [unfair PARTNER > unfair UNKNOWN] (orange, same as in A). Overlapping activations are displayed in purple (660 voxels, 1.98 cm³). Neurosynth results were corrected for multiple comparisons using a whole-brain false discovery rate (FDR) threshold of $P < 0.05$.

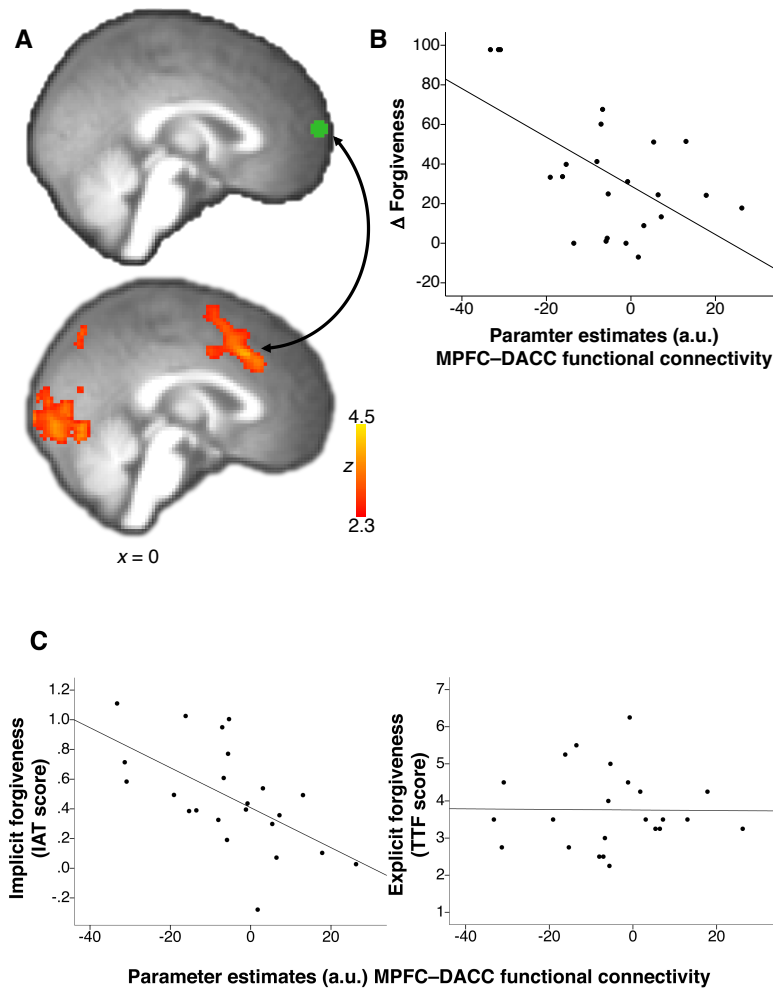


Figure 13. Reduced connectivity between MPFC and DACC predicts forgiveness. (A) PPI analysis identifying brain regions showing connectivity changes with the MPFC as a function of receiving unfair offers from the partner versus the unknown person. The upper panel depicts our functionally defined seed region (i.e., MPFC). The lower panel depicts brain regions whose connectivity with MPFC increases for unfair offers from the partner (vs. the unknown person). Results were thresholded at $z > 2.3$, $P < 0.05$, cluster corrected. (B) Regression line depicting the relationship between in MPFC/DACC functional connectivity and forgiveness index (Δ) across participants. The x-axis represents the mean parameter estimates in arbitrary units (a.u.) for task-related change in MPFC/DACC functional connectivity. The y-axis represents the forgiveness index, defined as the absolute difference between accepted unfair offers from the partner versus the unknown person [$\Delta \text{ forgiveness} = (\% \text{ of accepted unfair offers}_{\text{PARTNER}} - \% \text{ of accepted unfair offers}_{\text{UNKNOWN}})$]. (C) Regression lines depicting the relationship between MPFC/DACC functional connectivity and forgiveness levels across participants. The x-axis in both plots represents the mean parameter estimates in arbitrary units (a.u.) for task-related change in MPFC/DACC functional connectivity. The y-axis represents individual differences in forgiveness as measured by implicit (left panel; IAT score) and explicit (right panel; TTF score) measures.

5. Discussion

In this chapter, I will discuss how the results of the six empirical studies have contributed to address the research aims outlined above. Specifically, subsection 5.1. (Specific discussion) briefly evaluates study-specific results; subsection 5.2. (General discussion) integrates these results into a neurocognitive model of forgiveness. Finally, I will explain how this model may serve a framework for testing future predictions about the impact of forgiveness on the individual and relational level.

5.1. Specific discussion

5.1.1. New methods for studying forgiveness (Studies 1 and 2)

Study 1 developed and examined a non-self-report (i.e., indirect) measure of forgiveness, the forgiveness IAT. The forgiveness IAT demonstrated high internal consistency (i.e., split-half reliability was $> .80$) and moderate temporal stability after a time span of four weeks (i.e., test-retest reliability was $.50$). This discrepancy (satisfactory internal consistency vs. low-to-moderate stability) is typically observed in the IAT literature (Lane, Banaji, Nosek, & Greenwald, 2007; Nosek, Greenwald, & Banaji, 2007) and may imply that there are context-specific variations in forgiveness IAT scores (e.g., changes in the to-be-measured association strengths; Schnabel, Asendorpf, & Greenwald, 2008). This would be in line with evidence that IATs can capture both state-specific and trait-specific variance (Schmukle & Egloff, 2005). Indeed, recent studies suggest that forgiveness is particularly malleable, that is, sensitive to contextual changes (Karremans & Aarts, 2007; Karremans & Van Lange, 2010).

Regarding implicit-explicit correspondence (i.e., correlations between IAT and direct measures), the forgiveness IAT was virtually unrelated to a commonly used forgiveness self-report, the TTF scale. At first glance, this seems inconsistent with past research showing that direct and indirect measures of the same construct show low to moderate degrees of convergence (for meta-analytic evidence, see Hofmann, Gawronski, Gschwendner, Le, & Schmitt, 2005). Note, however, that the forgiveness IAT and the TTF scale differ in (at least) two factors: (a) content (i.e., the stimuli included in both instruments) and (b) processing mode (i.e., reflective vs. impulsive). Importantly, these factors have been shown to lower the implicit-explicit correspondence (see Hofmann, Gschwendner, Nosek, & Schmitt, 2005).

Regarding predictive (and incremental) validity, results were in line with the hypothesis that individuals should be faster to rate their forgiveness when they have largely forgiven someone compared to when they continue to harbor a grudge. Implicit (IAT-

measured) and explicit (self-reported) forgiveness predicted shorter RT of state forgiveness ratings. Importantly, the forgiveness IAT showed evidence of incremental validity and, hence, added to the prediction of behavior beyond the TTF scale. Thus, consistent with previous studies (Egloff & Schmukle, 2002; Schnabel, Banse, & Asendorpf, 2006), we demonstrated that the IAT and the direct measure explained separate aspects of the criterion variance (for an overview of predictive models of implicit and explicit measures, see Perugini, 2005). From a dual-process perspective, then, this would speak in favor of the distinction of and existence between automatic versus reflective information processing (Perugini, Richetin, & Zogmaister, 2010).

To note, Study 1 was mute as to the processes underlying the link between forgiveness and RT. Study 2 extended Study 1 by proposing and testing rumination about the transgression as a potential mediating mechanism of this link. Importantly, Study 2 conceptually replicated the finding that those high in forgiveness were faster to rate their state forgiveness (Fatfouta et al, 2014). This is noteworthy, given that Study 2 — unlike Study 1 — was conducted online and used a demographically diverse (i.e., heterogeneous) sample of participants. However, despite these procedural differences, the relation between individuals' forgiveness and their speed of rating their thoughts/feelings toward a transgressor seemed to be quite robust. Also, Study 2 provided evidence for a significant mediation effect (i.e., the direct effect of forgiveness on RT vanished when rumination was included into the model). Theoretically, this indicates that individuals with a general propensity to forgive also tend to ruminate less about their transgressions, which in turn promotes faster responding when asked to describe and rate one's state forgiveness. Lower rumination about the transgression may suggest a successful inhibition of unwanted (i.e., intrusive) memories associated with a particular transgression (Koster, De Lissnyder, Derakshan, & De Raedt, 2011). This would agree with studies showing that rumination is accompanied by deficits in inhibitory control over information in memory, specifically emotional information (Joormann, Levens, & Gotlib, 2011).

Together, results from Studies 1 and 2 add credence to the notion that forgiveness — like other social-cognitive processes — is likely to operate through a dual system: reflective (i.e., propositional self-representations: “I am forgiving”) and impulsive (i.e., associative self-representations: “me-forgiving”). Results further point to the utility of RT to state forgiveness items as promising tool for forgiveness research.

5.1.2. Mechanisms behind narcissists' lack of forgiveness (Studies 3 and 4)

Study 3 investigated (1) how two facets of narcissism — admiration and rivalry — relate to lack of forgiveness (i.e., high revenge and avoidance motivation toward the transgressor) as well as (2) how socio-cognitive factors (i.e., state anger, state rumination, and state empathy) mediate this relationship. Admiration showed no relations with revenge/avoidance, but this result could be explained by the suppressing effect of empathy. Specifically, individuals high in admiration (i.e., those who self-enhance to be admired by others) were capable of empathizing with their transgressor, which in turn mitigated their lack of forgiveness. By contrast, rivalry showed positive relations with revenge and this relation was mediated by anger, rumination, and empathy. Put differently, individuals high in rivalry (i.e., those who derogate others to avoid social failure) struggle to forgive their transgressor and this struggle can be explained by states associated with difficulty to forgive.

These particular results accentuate the value of conducting a more fine-grained analysis of narcissism (Ackerman et al., 2011; Clarke et al., 2014). Specifically, these findings add credence to the notion that narcissism can have both adaptive and maladaptive consequences: Adaptive narcissism (i.e., admiration) was positively related to empathy for the transgressor but it was negatively related to rumination about the hurt. Maladaptive narcissism (i.e., rivalry) showed inverse relations with these socio-cognitive variables; it was negatively related to empathic feelings but positively related to ruminative thoughts. This is consistent with the conceptualization of admiration as the bright side of narcissism and rivalry as its dark side (Back et al., 2013). Moreover, these findings cohere with evidence that especially the maladaptive dimension of narcissism entails adverse interpersonal consequences (Hepper, Hart, & Sedikides, 2014; Leckelt, Küfner, Nestler, & Back, in press). Hence, it is important to examine factors that can mitigate these consequences.

As a first step toward this challenge, Study 4 investigated how apology by the transgressor moderates narcissists' lack of forgiveness. Consistent with Study 3, rivalry (but less so admiration) was positively related to lack of forgiveness. Crucially, the link between rivalry and motivations to avoid the transgressor could be reduced when the transgressor offered a costly apology (in addition to a verbal expression of regret). However, the motivation to retaliate against the transgressor remained virtually unchanged. Consistent with previous studies showing that different narcissism dimensions can yield different interpersonal consequences (Back et al., 2013; Fatfouta et al., 2015), these results again indicate that the maladaptive (i.e., rivalry) rather than adaptive (i.e., admiration) dimension of narcissism is linked to lack of forgiveness.

Study 4 suggests that individuals high in rivalry may have an unmitigated revenge motivation toward their transgressor and this motivation persists irrespective of attempts to apologize. High revenge motivation can be detrimental to interpersonal functioning (McCullough et al., 1998; McCullough et al., 1997). However, revenge can also be conceived as a functional response to conflict (Fitness & Peterson, 2008; Gollwitzer, 2009). Specifically, revenge can ward off the reoccurrence of similar transgressions (McCullough et al., 2013) and communicate a “message” to the wrongdoer (“don't mess with me”; e.g., Funk, McGeer, & Gollwitzer, 2014; Gollwitzer, Meder, & Schmitt, 2011). Thus, it may be speculated that individuals high in rivalry pursue specific revenge-related goals, such as preventing others from self-exploitation or inducing them to acknowledge one deserves better treatment (McCullough et al., 2011, 2013).

Why are individuals high in rivalry less likely to avoid the transgressor following costly apology? We deem two answers plausible here: First, costly apology has been shown to be effective in communicating sincerity (Ohtsubo & Watanabe, 2009) and, hence, could signal that similar transgressions — and by extension potential ego-threats — are not likely to be repeated in the future (McCullough et al., 2014). Second, narcissism has formerly been associated with an approach motivation toward “functional” people — that is, people who provide interpersonal benefits (Jonason & Schmitt, 2012). Consistent with costly signaling theory (Miller, 2007), individuals high in rivalry may then approach the transgressor because his/her costly apology signals potential resource access and, hence, a desirable quality as a person. These resources may entail both financial benefits (e.g., invitation to a prestigious restaurant) and relational ones (e.g., commitment).

In sum, results from Studies 3 and 4 point to the value of distinguishing between two distinct dimensions of narcissism (i.e., admiration and rivalry) to better understand its interpersonal consequences following interpersonal transgressions. Results further point to the utility of considering key factors underlying narcissists’ lack of forgiveness (Study 3) as well as ways to curtail it (Study 4).

5.1.3. OC symptoms, revenge, and interpersonal transgressions (Study 5)

Study 5 aimed to better understand the hitherto neglected relationships between OC symptoms and (1) revenge as well as (2) the frequency of incurring interpersonal transgressions. Across the three studies, using three independent samples, results demonstrated that OC symptoms were associated with more positive attitudes towards revenge (Study 1), higher trait revenge (Study 2), and higher revenge motivation in response

to a real-life transgression (Study 3). Interestingly, the most consistent relationships occurred between revenge and the OCI-R obsessing subscale. This finding lends empirical support to the notion that individuals can be “obsessed with revenge” and also parallels previous research on the positive link between obsessing and anger/aggression (Moritz et al., 2012; Storch et al., 2012). Noteworthy, the revenge-obsession link concurs with findings from research on posttraumatic stress disorder, showing that feelings of revenge predict intrusions/re-experiencing symptoms (Orth, Montada, & Maercker, 2006). In this light, one might suggest that ruminative features of the transgression may induce anger, which under certain circumstances (e.g., negative appraisal of the transgression) can instigate feelings of revenge (for a similar argument, see Gäbler & Maercker, 2011).

To note, OC symptoms also predicted individuals’ perceptions of daily transgression occurrences (Study 4). Specifically, individuals high (vs. low) in OC symptoms experienced a greater number of transgressions in their daily lives. This finding is consistent with previous evidence that individuals with OCD are impaired in their interpersonal functioning (Hauschildt, Jelinek, Randjbar, Hottenrott, & Moritz, 2010; Painuly et al., 2012; Piacentini et al., 2003). In a recent survey by Hauschildt et al. (2010), for example, more than half of the patients (54.5%) reported that OCD causes conflicts in their partnerships. Furthermore, individuals with OCD tend to be distrustful of others (Moritz et al., 2011) and feel ambivalent toward them (Moritz et al., 2009). These findings may indicate that OCD also involves a heightened sensitivity toward negative events, such as criticism, hostility or other interpersonal harm. Indeed, it has been shown that OCD patients demonstrate a heightened sensitivity towards despising or critical facial expressions (see Mancini & Gangemi, 2011).

In sum, Study 5 represents a first step towards understanding how OC symptoms relate to revenge phenomena and the perception of interpersonal transgressions in everyday life. Our results may indicate that revenge phenomena need to be explicitly targeted in psychotherapy (Gäbler & Maercker, 2011). On the one hand, individuals with OCD need to accept their thoughts/feelings of revenge — they are part of normal cognition. On the other hand, after having acknowledged these thoughts/feelings, they need to develop effective coping strategies aimed at improving interpersonal functioning (Gäbler & Maercker, 2011; Moritz, Jelinek, Hauschildt, & Naber, 2010). Although promising, results need to replicate in a clinical OCD sample.

5.1.4. A neural mechanism of forgiveness (Study 6)

Study 6 used fMRI to identify a potential neural mechanism of forgiveness. Specifically, Study 6 aimed to examine how interpersonal closeness modulates behavioral and neural reactions to transgressions (in the present study, unfair monetary offers in an economic exchange game). On the behavioral level, participants were more forgiving of their partner (vs. an unknown person), as indexed by greater acceptance of unfair exchanges in the UG. On the neural level, when the partner (vs. the unknown person) behaved unfairly, we observed (1) increased MPFC activation and (2) heightened MPFC/DACC task-related functional connectivity. Most importantly, individuals with lower task-related MPFC/DACC connectivity were more forgiving of their partners' unfairness and scored higher on implicit (but not explicit) forgiveness.

Our behavioral results agree with the valuable-relationship hypothesis, positing that individuals are more likely to forgive their transgressor when he/she is a valuable partner (Burnette, McCullough, Van Tongeren, & Davis, 2012; McCullough, 2008). Individuals were more likely to accept unfair exchanges from the loved one (vs. an unknown person) and, importantly, this effect is unlikely to be driven by individual differences in resource sharing among couples. Specifically, controlling for resource sharing did not alter our results. In our view, this behavioral tendency — refraining from punishing unfair exchanges — rather reflects a forgiveness component, which also becomes evident in participants' emotion ratings: In response to unfair partner (vs. unfair unknown) treatment, participants reported increased feelings of forgiveness/benevolence. This is consistent with previous studies showing that closeness mitigates social-norm enforcement (i.e., punishing norm violations) in the context of close but not non-close others (Campanhã, Minati, Fregni, & Boggio, 2011; Diekhof, Wittmer, & Reimers, 2014).

Our fMRI results revealed that unfair exchanges from one's partner (vs. an unknown person) recruited the MPFC. Activity in this region is consistent with previous studies on mentalizing during strategic interactions (Hampton, Bossaerts, & O'Doherty, 2008) as well as mentalizing about personally relevant (e.g., close) others (Krienen et al., 2010). One interpretation might therefore be that unfair treatment by one's partner requires individuals to adopt his/her mental state in an attempt to understand the motivations underlying his/her behavior (e.g., "Why does he/she treat me like this?"). The reverse-inference analysis (Yarkoni et al., 2011) provided support in this regard, as demonstrated by an overlap between the present results and previous fMRI studies on "mentalizing".

Importantly, our results suggest a functional link between MPFC and DACC when participants think about whether or not to forgive their partner (vs. an unknown person). The DACC has been implicated performance/conflict monitoring (Botvinick et al., 2001) and may underlie the generation of an “alarm” signal, that calls for the resolution of conflict (Eisenberger & Lieberman, 2004; Spunt et al., 2012). Importantly, the DACC is well connected with the MPFC, which has been shown to bias behavior during conflict situations (Nakao et al., 2010). In our study, participants may have been faced with such a situation: Punishing unfair partner behavior may convey an important signal ("don't mess with me"; Funk, et al., 2014; see Study 4 discussed above), but at the same time threaten the relationship one has with the partner (e.g., signaling rejection). Hence, one hypothesis derived from these results is that the MPFC modulates conflict-related DACC activity via mentalizing (e.g., reinterpreting the intentions of one's partner) to resolve conflict due to competing response tendencies (i.e., forgive/accept vs. not forgive/reject).

Participants with lower MPFC/DACC connectivity showed increased forgiveness, as indexed by greater acceptance of unfair partner relative to unfair unknown exchanges. Interestingly, lower MPFC/DACC connectivity was also recently found among experienced meditators (Hasenkamp & Barsalou, 2012), suggesting disengagement from the present thought stream on the one hand, and greater acceptance of the situation on the other (Tang, Hölzel, & Posner, 2015). Crucially, the pattern of lower MPFC/DACC connectivity/higher forgiveness replicated with an individual-difference measure of forgiveness, as assessed by the forgiveness IAT administered after scanning: Lower MPFC/DACC connectivity predicted individuals' automatic (implicit, indirectly measured) but not controlled (explicit, self-reported) forgiveness. This finding parallels findings from the neuroscience of implicit social cognition, suggesting that brain responses are better predictors of automatic associations rather than self-reported traits and attitudes (Phelps et al., 2000; Stanley, Phelps, & Banaji, 2008). Furthermore, this finding provides further evidence for the validity of the forgiveness IAT. It also suggests that implicit forgiveness may indeed be malleable, that is, subject to context-specific changes that cannot be adequately captured by the TTF scale.

In conclusion, Study 6 provides the first evidence for a neural mechanism of forgiveness. Reduced MPFC/DACC functional connectivity may serve as a biomarker of forgiveness, thus providing fruitful avenues for a deeper understanding of lack of forgiveness in narcissism (see Studies 3 and 4) and/or OCD (see Study 5).

5.2. General discussion

5.2.1. A neurocognitive model of forgiveness

The proposed model aims to explain the processes that are involved when individuals face a transgressing event (see Fig. 14). The key assumption is that forgiveness happens both between and within people and incorporates two components (Newberg et al., 2000): First, the perception that some conflict has occurred (e.g., unreciprocated cooperation from a loved one) and, second, the attempt to resolve that conflict so as to restore a peace of mind.

On the interpersonal level, a transgressing event contains certain cues, which need to be readily perceived to be able to influence the evaluation of the transgression and how the transgressor should ultimately be treated (Fehr et al., 2010; Riek & Mania, 2012). These cues may refer to aspects of the relationship with the transgressor (e.g., interpersonal closeness; How close is he/she to me?) or to aspects of the transgression itself (e.g., presence vs. absence of apology/amends; Did he/she apologize for the wrongdoing?).

On the intrapersonal level, a network comprising the DACC and the MPFC mediates transgression-related information. The DACC is involved in discrepancy detection between a desired goal state (e.g., my partner should behave fairly) and an actual goal state (e.g., my partner behaves unfairly) and engages the MPFC to resolve such discrepancy by means of mentalizing-related computations (e.g., attributing a partner's unreciprocated cooperation as being due to financial problems instead of being due to selfishness). The less transgression-related information is online and readily available to the MPFC, the more an individual is to disengage from and accept the transgressing event at hand. Note that personality (e.g., narcissistic traits) and/or psychopathology (e.g., OC symptoms) can have distinct influences on the model components. Narcissists, which are characterized by lack of empathy and exploitative interpersonal behavior (Rhodewalt & Peterson, 2009), show a lack of forgiveness following interpersonal transgressions. Individuals high in OCD, which are characterized by obsessive and compulsive symptoms (Foa et al., 2002), similarly show a lack of forgiveness and perceive a greater frequency of interpersonal transgressions. Thus, narcissism and OCD may involve aberrant processing of transgression-related information.

Again, on the interpersonal level, different degrees of forgiveness are possible and these can be expressed in either direct (i.e., self-report) or indirect (i.e., IAT) measures. Whereas direct measures capture information processing that results from slow-acting reasoning processes, indirect measures tap into information processing that results from fast-acting associative processes (Back et al., 2009; Strack & Deutsch, 2004).

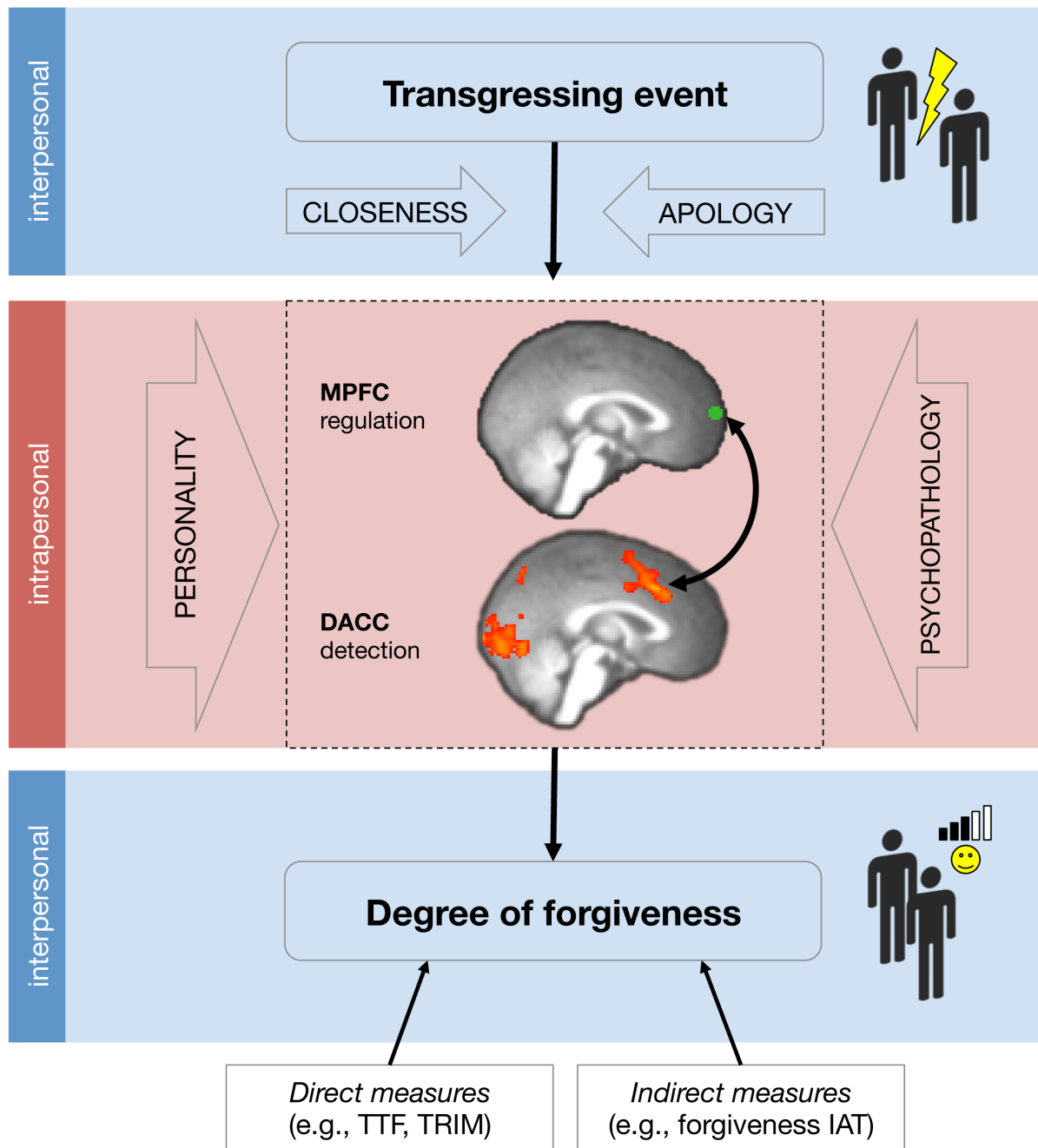


Figure 14. A neurocognitive model of forgiveness. The model distinguishes between inter- and intrapersonal processes. On the interpersonal level, a transgressing event that happens between two individuals can be modulated by relational and/or situational factors, such as closeness and apology. On the intrapersonal level, a network comprising the DACC and MPFC is suggested to be involved in discrepancy detection (e.g., being treated unfairly) and its resolution (e.g., mentalizing about potential reasons for the unfairness), respectively. Personality (e.g., narcissistic traits) and symptoms of psychopathology (e.g., OC symptoms) similarly exert their influences on the processing of transgression-related information. Finally, different degrees of forgiveness are possible, which can be measured directly (i.e., via verbal report) or indirectly (i.e., via the forgiveness IAT). TTF = Tendency to Forgive scale. TRIM = Transgression-Related Interpersonal Motivations. IAT = Implicit Association Test.

5.2.2. Implications for future research

The results from the six studies presented herein suggest that forgiveness is a multifaceted phenomenon. To capture this multifacetedness, multiple assessment formats (e.g., self-report, RT, IAT, behavior, neuroimaging) should be employed.

First, future studies might focus on potential factors that moderate the forgiveness IAT's predictive validity. Under time pressure or cognitive load, for example, people consider less information and behave more impulsively (Perugini et al., 2010). Individual differences in implicit forgiveness might therefore predict behavior under conditions of constrained processing time and/or self-regulation. Future research would also benefit from examining the construct of implicit forgiveness in contexts of health and well-being. Given that discrepancies between the explicit and implicit self-concept have been shown to account for psychopathology (e.g., Schröder-Abé, Rudolph, & Schütz, 2007; for the example of discrepant self-esteem), similar processes might exist for the forgiveness self-concept. Examining whether and to what extent discrepant forms of forgiveness influence health-related outcomes remains an important question for future research.

Second, consistent with the notion that RTs provide insights into dynamic aspects of information processing (Robinson & Neighbors, 2006), RT of state forgiveness ratings constitute a promising tool for tapping transient changes in evaluative judgments about one's forgiveness (Studies 1 and 2). With the advent of computerized testing, RT collection has become easier and provides an additional, yet understudied, source of diagnostic information. Whereas in the present studies RT has been used as dependent variable, future research is needed to examine whether it also possesses predictive (and incremental) validity for the assessment of forgiveness-related behavior. In addition, examination of whether RT of state forgiveness ratings captures individual differences in state versus trait scores can aid understanding of the context-specificity of this measure.

Third, the results indicate that narcissism is an important variable to pursue in future research on personality correlates of forgiveness (Studies 3 and 4). Future studies might focus on additional mediators of the narcissism/lack of forgiveness link, such as perceived intentionality of the transgression. Research would also benefit from testing further factors that might mitigate narcissists' lack of forgiveness, such as increased closeness to the transgressor. Also, the neural mechanisms that may underlie narcissists' lack of forgiveness should be explored. One testable prediction derived from the previously introduced model is that narcissists have a heightened sensitivity to transgression-related information, which might be reflected in an increased MPFC/DACC coupling. This contention is consistent with

the findings of a recent fMRI study demonstrating that narcissists show increased neural activation in social pain-related areas (in particular, the DACC) while experiencing interpersonal threat (Cascio, Konrath, & Falk, 2015).

Fourth, the results regarding the clinical correlates of forgiveness suggest that individuals high in OC symptoms may similarly have difficulty in forgiving interpersonal transgressions (Study 5). While the OC symptoms/vengeance link was consistent across several revenge-related phenomena (i.e., vengeful attitudes, dispositions, and motivations), the mechanism underlying this link remains to be explored. Given their heightened suspiciousness/distrust of others (Moritz et al., 2011; Moritz et al., 2012), individuals high in OC symptoms may attribute malevolent intentions on the side of the transgressor — that is, perceptions of increased intentionality of the transgression might function as a mediator of the OC symptoms/vengeance link. Future work using correlational and ideally experimental designs should explore this possibility. The finding that individuals high in OC symptoms more readily perceive negative interpersonal events may suggest a hypersensitivity to these events. This view is consistent with previous work that suggests a role of DACC hyperactivity in OCD (Gehring & Himle, 2000; Ursu, Stenger, Shear, Jones, & Carter, 2003) and should be further explored.

Fifth, the fMRI study suggests that a neural network comprising the MPFC/DACC may serve as a marker of forgiveness (Study 6). While previous fMRI studies exclusively focused on segregated brain areas, Study 6 highlights the importance of examining networks of interacting brain areas (Aue et al., 2009). Though it has been suggested that the DACC is implicated in discrepancy detection and the MPFC is involved in resolving discrepancy via mentalizing (also see Nakao et al., 2010), the PPI analysis is mute as to the assignment of causality. For example, it is possible that the MPFC/DACC connection is not direct but mediated by other brain regions (O'Reilly et al., 2012). Further studies using transcranial magnetic stimulation (TMS) alone or in combination with fMRI are needed to answer the question regarding the directionality of the observed coupling. Given that TMS to MPFC has been shown to impair mentalizing ability (Lev-Ran, Shamay-Tsoory, Zangen, & Levkovitz, 2012), it is possible that disruptions of the MPFC may interfere with individuals' tendency to forgive unfair exchanges in the context of close (vs. non-close) others.

6. Conclusion

The results of this dissertation advance our understanding of the psychological and neural correlates of forgiveness. In conclusion, the results suggest the following:

- The forgiveness IAT provides a new, reliable, and incrementally valid indirect measure of forgiveness.
- Individuals high (vs. low) in forgiveness are faster to describe and rate their thoughts and feelings toward a transgressor and they do so because they tend to ruminate less about the transgression at hand.
- The narcissism/lack of forgiveness link is better understood if we distinguish between two distinct facets of narcissism — admiration and rivalry — and consider potential mediating and moderating influences of this link.
- Socio-cognitive factors that are prompted by a transgressing event are key factors accounting for narcissists' lack of forgiveness: Whereas admiration is linked to less revenge/avoidance via higher empathy, rivalry is linked to greater revenge/avoidance via higher anger and rumination and lower empathy.
- Conciliatory gestures including a cost for the transgressor can, in part, mitigate lack of forgiveness in individuals high in rivalry: Whereas costly apology reduced the tendency to avoid the transgressor, the tendency to retaliate against him/her remained virtually unchanged.
- OC symptoms, especially obsessing, are linked to more positive attitudes toward revenge, higher trait revenge, and more frequent perceptions of interpersonal transgressions.
- Forgiveness is modulated by relationship closeness and involves reduced functional connectivity between two brain structures; the MPFC implementing mentalizing-related computations and the DACC modulating conflict between competing response tendencies (i.e., forgive vs. not forgive).
- Reduced MPFC/DACC connectivity may serve as a biomarker of forgiveness.

7. References

- Ackerman, R. A., Witt, E. A., Donnellan, M. B., Trzesniewski, K. H., Robins, R. W., & Kashy, D. A. (2011). What Does the Narcissistic Personality Inventory Really Measure? *Assessment, 18*(1), 67-87.
- Allemand, M., Amberg, I., Zimprich, D., & Fincham, F. D. (2007). The role of trait forgiveness and relationship satisfaction in episodic forgiveness. *Journal of Social and Clinical Psychology, 26*(2), 199-217. doi: 10.1521/jscp.2007.26.2.199
- Allemand, M., Sassin-Meng, A., Huber, S., & Schmitt, M. (2008). Entwicklung und Validierung einer Skala der Bereitschaft zu verzeihen (SBV). [Development and validation of a measure of willingness to forgive (Skala der Beretischaft zu verzeihen; SBV)]. *Diagnostica, 54*(2), 71-84. doi: 10.1026/0012-1924.54.2.71
- Amodio, D. M., & Frith, C. D. (2006). Meeting of minds: the medial frontal cortex and social cognition. *Nature Reviews Neuroscience, 7*(4), 268-277. doi: 10.1038/nrn1884
- Aue, T., Lavelle, L. A., & Cacioppo, J. T. (2009). Great expectations: what can fMRI research tell us about psychological phenomena? *International Journal Of Psychophysiology: Official Journal Of The International Organization Of Psychophysiology, 73*(1), 10-16. doi: 10.1016/j.ijpsycho.2008.12.017
- Back, M. D., Kufner, A. C. P., Dufner, M., Gerlach, T. M., Rauthmann, J. F., & Denissen, J. J. A. (2013). Narcissistic admiration and rivalry: Disentangling the bright and dark sides of narcissism. *Journal of Personality and Social Psychology, 105*(6), 1013-1037. doi: 10.1037/a0034431
- Back, M. D., Schmukle, S. C., & Egloff, B. (2009). Predicting actual behavior from the explicit and implicit self-concept of personality. *Journal of Personality and Social Psychology, 97*(3), 533-548. doi: 10.1037/a0016229
- Banse, R. (2001). Affective priming with liked and disliked persons: Prime visibility determines congruency and incongruency effects. *Cognition & Emotion, 15*(4), 501-520. doi: 10.1080/0269993004200213
- Berry, J. W., Worthington, E. L., Jr., Parrott, L., III, O'Connor, L. E., & Wade, N. G. (2001). Dispositional forgivingness: Development and construct validity of the Transgression Narrative Test of Forgivingness (TNTF). *Personality and Social Psychology Bulletin, 27*(10), 1277-1290. doi: 10.1177/01461672012710004
- Berry, J. W., Worthington, E. L., O'Connor, L. E., Parrott, L., & Wade, N. G. (2005). Forgivingness, Vengeful Rumination, and Affective Traits. *Journal of Personality, 73*(1), 183-226. doi: 10.1111/j.1467-6494.2004.00308.x
- Bornstein, R. F. (2002). A Process Dissociation Approach to Objective-Projective Test Score Interrelationships. *Journal of Personality Assessment, 78*(1), 47-68.
- Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychol Rev, 108*(3), 624-652.
- Brown, J. W. (2013). Beyond conflict monitoring: Cognitive control and the neural basis of thinking before you act. *Current Directions in Psychological Science, 22*(3), 179-185. doi: 10.1177/0963721412470685
- Brown, R. P. (2003). Measuring individual differences in the tendency to forgive: Construct validity and links with depression. *Personality and Social Psychology Bulletin, 29*(6), 759-771. doi: 10.1177/0146167203029006008
- Brown, R. P. (2004). Vengeance is mine: Narcissism, vengeance, and the tendency to forgive. *Journal of Research in Personality, 38*(6), 576-584. doi: 10.1016/j.jrp.2003.10.003

- Brown, R. P., & Phillips, A. (2005). Letting bygones be bygones: further evidence for the validity of the Tendency to Forgive scale. *Personality & Individual Differences*, 38(3), 627-638. doi: 10.1016/j.paid.2004.05.017
- Brüne, M., Juckel, G., & Enzi, B. (2013). "An eye for an eye"? Neural correlates of retribution and forgiveness. *Plos One*, 8(8), e73519-e73519. doi: 10.1371/journal.pone.0073519
- Burnette, J. L., McCullough, M. E., Van Tongeren, D. R., & Davis, D. E. (2012). Forgiveness results from integrating information about relationship value and exploitation risk. *Personality and Social Psychology Bulletin*, 38(3), 345-356. doi: 10.1177/0146167211424582
- Bushman, B. J., & Baumeister, R. F. (1998). Threatened egotism, narcissism, self-esteem, and direct and displaced aggression: Does self-love or self-hate lead to violence? *Journal of Personality and Social Psychology*, 75(1), 219-229. doi: 10.1037/0022-3514.75.1.219
- Camerer, C. F. (2003). *Behavioral game theory: Experiments in strategic interaction*. New York, NY, US: Russell Sage Foundation.
- Campanhã, C., Minati, L., Fregni, F., & Boggio, P. S. (2011). Responding to unfair offers made by a friend: Neuroelectrical activity changes in the anterior medial prefrontal cortex. *The Journal of Neuroscience*, 31(43), 15569-15574. doi: 10.1523/JNEUROSCI.1253-11.2011
- Cascio, C. N., Konrath, S. H., & Falk, E. B. (2015). Narcissists' social pain seen only in the brain. *Social Cognitive and Affective Neuroscience*, 10(3), 335-341. doi: 10.1093/scan/nsu072
- Clarke, I. E., Karlov, L., & Neale, N. J. (2014). The many faces of narcissism: Narcissism factors and their predictive utility. *Personality and Individual Differences*. doi: 10.1016/j.paid.2014.11.021
- Cogan, R., Ashford, D., Chaney, B., Embry, S., Emory, L., Goebel, H., Holstrom, N., Keithley, D., 3rd, Lawson, M., McPherson, J., Scott, B., & Tebbets, J., Jr. (2004). Obsessiveness and a thematic apperception test-based measure of aggression. *Psychological Reports*, 95(3 Pt 1), 828-830.
- Crockett, M. J., Apergis-Schoute, A., Herrmann, B., Lieberman, M., Müller, U., Robbins, T. W., & Clark, L. (2013). Serotonin modulates striatal responses to fairness and retaliation in humans. *The Journal of Neuroscience*, 33(8), 3505-3513. doi: 10.1523/JNEUROSCI.2761-12.2013
- De Houwer, J. (2006). What Are Implicit Measures and Why Are We Using Them? In R. W. Wiers & A. W. Stacy (Eds.), *Handbook of implicit cognition and addiction*. (pp. 11-28). Thousand Oaks, CA US: Sage Publications, Inc.
- DeMoss, N. L. (2008). *Choosing Forgiveness: Your Journey to Freedom* Chicago: Moody Publishers.
- Denny, B. T., Kober, H., Wager, T. D., & Ochsner, K. N. (2012). A Meta-analysis of Functional Neuroimaging Studies of Self- and Other Judgments Reveals a Spatial Gradient for Mentalizing in Medial Prefrontal Cortex. *Journal of Cognitive Neuroscience*, 24(8), 1742-1752.
- Diekhof, E. K., Wittmer, S., & Reimers, L. (2014). Does competition really bring out the worst? Testosterone, social distance and inter-male competition shape parochial altruism in human males. *Plos One*, 9(7), e98977-e98977. doi: 10.1371/journal.pone.0098977
- Dufner, M., Egloff, B., Hausmann, C. M., Wendland, L.-M., Neyer, F. J., & Back, M. D. (2014). Narcissistic Tendencies Among Actors: Craving for Admiration, But Not at

- the Cost of Others. *Social Psychological and Personality Science*. doi: 10.1177/1948550614564224
- Eaton, J., Struthers, C. W., & Santelli, A. G. (2006a). Dispositional and state forgiveness: The role of self-esteem, need for structure, and narcissism. *Personality and Individual Differences*, *41*(2), 371-380. doi: 10.1016/j.paid.2006.02.005
- Eaton, J., Struthers, C. W., & Santelli, A. G. (2006b). The Mediating Role of Perceptual Validation in the Repentance-Forgiveness Process. *Personality and Social Psychology Bulletin*, *32*(10), 1389-1401. doi: 10.1177/0146167206291005
- Egloff, B., & Schmukle, S. C. (2002). Predictive validity of an implicit association test for assessing anxiety. *Journal of Personality and Social Psychology*, *83*(6), 1441-1455. doi: 10.1037/0022-3514.83.6.1441
- Eisenberger, N. (2012a). The neural bases of social pain: Evidence for shared representations with physical pain. *Psychosomatic Medicine*, *74*(2), 126-135. doi: 10.1097/PSY.0b013e3182464dd1
- Eisenberger, N. (2012b). The pain of social disconnection: examining the shared neural underpinnings of physical and social pain. *Nature Reviews Neuroscience*, *13*(6), 421-434. doi: 10.1038/nrn3231
- Eisenberger, N. (2015). Social pain and the brain: Controversies, questions, and where to go from here. *Annual Review of Psychology*, *66*, 601-629. doi: 10.1146/annurev-psych-010213-115146
- Eisenberger, N., Lieberman, M., & Williams, K. (2003). Does rejection hurt? An FMRI study of social exclusion. *Science (New York, N.Y.)*, *302*(5643), 290-292.
- Eisenberger, N., & Lieberman, M. D. (2004). Why rejection hurts: A common neural alarm system for physical and social pain. *Trends in Cognitive Sciences*, *8*(7), 294-300. doi: 10.1016/j.tics.2004.05.010
- Emmons, R. A. (1992). *Revenge: Individual differences and correlates*. Paper presented at the Annual Meeting of the American Psychological Association, Washington.
- Emmons, R. A. (2000). Personality and forgiveness. In M. E. McCullough, K. I. Pargament & C. E. Thoresen (Eds.), *Forgiveness: Theory, research, and practice*. (pp. 156-175). New York, NY US: Guilford Press.
- Enright, R. D. (2001). *Forgiveness is a choice: A step-by-step process for resolving anger and restoring hope*. Washington, DC, US: American Psychological Association.
- Enright, R. D., & Fitzgibbons, R. P. (2000). *Helping clients forgive: An empirical guide for resolving anger and restoring hope*. Washington, DC US: American Psychological Association.
- Etkin, A., Egner, T., & Kalisch, R. (2011). Emotional processing in anterior cingulate and medial prefrontal cortex. *Trends in Cognitive Sciences*, *15*(2), 85-93. doi: 10.1016/j.tics.2010.11.004
- Exline, J. J., Baumeister, R. F., Bushman, B. J., Campbell, W. K., & Finkel, E. J. (2004). Too Proud to Let Go: Narcissistic Entitlement as a Barrier to Forgiveness. *Journal of Personality and Social Psychology*, *87*(6), 894-912. doi: 10.1037/0022-3514.87.6.894
- Farrow, T. F., Hunter, M. D., Wilkinson, I. D., Gouneea, C., Fawbert, D., Smith, R., Lee, K.-H., Mason, S., Spence, S. A., & Woodruff, P. W. R. (2005). Quantifiable change in functional brain response to empathic and forgiveness judgments with resolution of posttraumatic stress disorder. *Psychiatry Research: Neuroimaging*, *140*(1), 45-53.
- Farrow, T. F., Zheng, Y., Wilkinson, I. D., Spence, S. A., Deakin, J. F., Tarriner, N., Griffiths, P. D., & Woodruff, P. W. (2001). Investigating the functional anatomy of empathy and forgiveness. *Neuroreport*, *12*(11), 2433-2438.

- Fatfouta, R. (2015). How forgiveness affects processing time: Mediation by rumination about the transgression. *Personality and Individual Differences*, 82(0), 90-95. doi: 10.1016/j.paid.2015.03.016
- Fatfouta, R., & Gerlach, T. M. (under review). Can the shrew be tamed? Narcissism, lack of forgiveness, and the moderating role of conciliatory gestures.
- Fatfouta, R., Gerlach, T. M., Schröder-Abé, M., & Merkl, A. (2015). Narcissism and lack of interpersonal forgiveness: The mediating role of state anger, state rumination, and state empathy. *Personality and Individual Differences*, 75, 36-40. doi: 10.1016/j.paid.2014.10.051
- Fatfouta, R., Jacobs, A., & Merkl, A. (2013). Towards a multifaceted understanding of revenge and forgiveness. *The Behavioral And Brain Sciences*, 36(1), 21-21. doi: 10.1017/s0140525x12000374
- Fatfouta, R., & Merkl, A. (2014). Associations between obsessive-compulsive symptoms, revenge, and the perception of interpersonal transgressions. *Psychiatry Research*, 219(2), 316-321. doi: 10.1016/j.psychres.2014.05.038
- Fatfouta, R., Meshi, D., Merkl, A., & Heekeren, H. R. (under review). Forgiveness involves reduced connectivity between medial prefrontal and dorsal anterior cingulate cortex.
- Fatfouta, R., Schröder-Abé, M., & Merkl, A. (2014). Forgiving, Fast and Slow: Validity of the Implicit Association Test for Predicting Differential Response Latencies in a Transgression-Recall Paradigm. *Frontiers in Psychology*, 5, 1-14. doi: 10.3389/fpsyg.2014.00728
- Fazio, R. (1990a). Multiple Processes by which Attitudes guide Behavior: the MODE Model as an integrative framework. *Advances in experimental social psychology*, 23, 75-109.
- Fazio, R. (1990b). A practical guide to the use of response latency in social psychological research. In C. Hendrick & M. S. Clark (Eds.), *Research methods in personality and social psychology*. (pp. 74-97). Thousand Oaks, CA US: Sage Publications, Inc.
- Fehr, E., & Gächter, S. (2000). Fairness and Retaliation: The Economics of Reciprocity. *Journal of Economic Perspectives*, 14(3), 159-181.
- Fehr, R., Gelfand, M. J., & Nag, M. (2010). The road to forgiveness: A meta-analytic synthesis of its situational and dispositional correlates. *Psychological Bulletin*, 136(5), 894-914. doi: 10.1037/a0019993
- Fenichel, O. (1945). *The psychoanalytic theory of neurosis*. New York, NY, US: W W Norton & Co.
- Fincham, F. D. (2000). The kiss of the porcupines: From attributing responsibility to forgiving. *Personal Relationships*, 7(1), 1-23. doi: 10.1111/j.1475-6811.2000.tb00001.x
- Fincham, F. D. (2010). Forgiveness: Integral to a science of close relationships? In M. Mikulincer & P. R. Shaver (Eds.), *Prosocial motives, emotions, and behavior: The better angels of our nature*. (pp. 347-365). Washington, DC US: American Psychological Association.
- Fincham, F. D., & Beach, S. R. H. (2001). Forgiving in close relationships. In F. Columbus (Ed.), *Advances in psychology research, Vol. 7*. (pp. 163-197). Hauppauge, NY US: Nova Science Publishers.
- Finkel, E. J., Rusbult, C. E., Kumashiro, M., & Hannon, P. A. (2002). Dealing with betrayal in close relationships: Does commitment promote forgiveness? *Journal of Personality and Social Psychology*, 82(6), 956-974. doi: 10.1037/0022-3514.82.6.956
- Fitness, J., & Peterson, J. (2008). Punishment and forgiveness in close relationships: An evolutionary, social-psychological perspective. In J. P. Forgas, J. Fitness, J. P. Forgas & J. Fitness (Eds.), *Social relationships: Cognitive, affective, and motivational processes*. (pp. 255-269). New York, NY, US: Psychology Press.

- Foa, E. B., Huppert, J. D., Leiberg, S., Langner, R., Kichic, R., Hajcak, G., & Salkovskis, P. M. (2002). The Obsessive-Compulsive Inventory: Development and validation of a short version. *Psychological Assessment, 14*(4), 485-496. doi: 10.1037/1040-3590.14.4.485
- Friston, K. J., Buechel, C., Fink, G. R., Morris, J., Rolls, E., & Dolan, R. J. (1997). Psychophysiological and modulatory interactions in neuroimaging. *Neuroimage, 6*(3), 218-229.
- Frith, C. D., & Frith, U. (2006). The neural basis of mentalizing. *Neuron, 50*(4), 531-534.
- Funk, F., McGeer, V., & Gollwitzer, M. (2014). Get the message: Punishment is satisfying if the transgressor responds to its communicative intent. *Personality and Social Psychology Bulletin, 40*(8), 986-997. doi: 10.1177/0146167214533130
- Gäbler, I., & Maercker, A. (2011). Revenge after trauma: Theoretical outline. In M. Linden & A. Maercker (Eds.), *Embitment. Societal, psychological, and clinical perspectives* (pp. 42-69). Wien: Springer.
- Gehring, W. J., & Himle, J. (2000). Action-monitoring dysfunction in obsessive-compulsive disorder. *Psychological Science (Wiley-Blackwell), 11*(1), 1.
- Gollwitzer, M. (2009). Justice and revenge. In M. E. Oswald, S. Bieneck & J. Hupfeld-Heinemann (Eds.), *Social psychology of punishment of crime* (pp. 137-156). Chichester: Wiley Blackwell.
- Gollwitzer, M., Meder, M., & Schmitt, M. (2011). What gives victims satisfaction when they seek revenge? *European Journal of Social Psychology, 41*(3), 364-374. doi: 10.1002/ejsp.782
- 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*(1), 3-25. doi: 10.1037/0033-295x.109.1.3
- 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*(2), 197-216. doi: 10.1037/0022-3514.85.2.197
- Greve, D. N., & Fischl, B. (2009). Accurate and robust brain image alignment using boundary-based registration. *Neuroimage, 48*(1), 63-72. doi: 10.1016/j.neuroimage.2009.06.060
- Güth, W., Schmittberger, R., & Schwarze, B. (1982). An experimental analysis of ultimatum bargaining. *Journal of economic behavior & organization, 3*(4), 367-388.
- Hampton, A. N., Bossaerts, P., & O'Doherty, J. P. (2008). Neural correlates of mentalizing-related computations during strategic interactions in humans. *PNAS Proceedings of the National Academy of Sciences of the United States of America, 105*(18), 6741-6746.
- Harris, A. H. S., Luskin, F., Norman, S. B., Standard, S., Bruning, J., Evans, S., & Thoresen, C. E. (2006). Effects of a group forgiveness intervention on forgiveness, perceived stress, and trait-anger. *Journal of Clinical Psychology, 62*(6), 715-733. doi: 10.1002/jclp.20264
- Hasenkamp, W., & Barsalou, L. W. (2012). Effects of meditation experience on functional connectivity of distributed brain networks. *Frontiers in Human Neuroscience, 6*. doi: 10.3389/fnhum.2012.00038
- Hauschildt, M., Jelinek, L., Randjbar, S., Hottenrott, B., & Moritz, S. (2010). Generic and illness-specific quality of life in obsessive-compulsive disorder. *Behavioural and Cognitive Psychotherapy, 38*(4), 417-436. doi: 10.1017/S1352465810000275
- Hayashi, A., Abe, N., Ueno, A., Shigemune, Y., Mori, E., Tashiro, M., & Fujii, T. (2010). Neural correlates of forgiveness for moral transgressions involving deception. *Brain Research, 1332*, 90-99. doi: 10.1016/j.brainres.2010.03.045

- Hayes, A. F. (2013). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. New York, NY, US: Guilford Press.
- Heatherston, T. F., Wyland, C. L., Macrae, C. N., Demos, K. E., Denny, B. T., & Kelley, W. M. (2006). Medial prefrontal activity differentiates self from close others. *Social Cognitive and Affective Neuroscience, 1*(1), 18-25. doi: 10.1093/scan/nsi001
- Hepper, E. G., Hart, C. M., & Sedikides, C. (2014). Moving Narcissus: Can Narcissists Be Empathic? *Personality & Social Psychology Bulletin, 40*(9), 1079-1091.
- 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*(10), 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. doi: 10.1080/10463280500443228
- Horowitz, M. J. (2007). Understanding and ameliorating revenge fantasies in psychotherapy. *The American Journal of Psychiatry, 164*(1), 24-27. doi: 10.1176/appi.ajp.164.1.24
- Hoyt, W. T., & McCullough, M. E. (2005). Issues in the Multimodal Measurement of Forgiveness. In E. L. Worthington (Ed.), *Handbook of Forgiveness*. New York: Routledge.
- Jenkinson, M., Bannister, P., Brady, M., & Smith, S. (2002). Improved optimization for the robust and accurate linear registration and motion correction of brain images. *Neuroimage, 17*(2), 825-841.
- Jenkinson, M., & Smith, S. (2001). A global optimisation method for robust affine registration of brain images. *Med Image Anal, 5*(2), 143-156.
- Jonason, P. K., & Schmitt, D. P. (2012). What Have You Done For Me Lately? Friendship-Selection in the Shadow of the Dark Triad Traits. *Evolutionary Psychology, 10*(3), 400-421.
- Joormann, J., Levens, S. M., & Gotlib, I. H. (2011). Sticky Thoughts: Depression and Rumination Are Associated With Difficulties Manipulating Emotional Material in Working Memory. *Psychological Science (Sage Publications Inc.), 22*(8), 979-983. doi: 10.1177/0956797611415539
- Karremans, J. C., & Aarts, H. (2007). The role of automaticity in determining the inclination to forgive close others. *Journal of Experimental Social Psychology, 43*(6), 902-917. doi: 10.1016/j.jesp.2006.10.012
- Karremans, J. C., Regalia, C., Paleari, F. G., Fincham, F. D., Cui, M., Takada, N., Ohbuchi, K.-I., Terzino, K., Cross, S. E., & Uskul, A. K. (2011). Maintaining harmony across the globe: The cross-cultural association between closeness and interpersonal forgiveness. *Social Psychological and Personality Science, 2*(5), 443-451. doi: 10.1177/1948550610396957
- Karremans, J. C., & Van Lange, P. A. M. (2008). Forgiveness in personal relationships: Its malleability and powerful consequences. *European Review of Social Psychology, 19*, 202-241. doi: 10.1080/10463280802402609
- Karremans, J. C., & Van Lange, P. A. M. (2010). The malleability of forgiveness. In M. Mikulincer & P. R. Shaver (Eds.), *Prosocial motives, emotions, and behavior: The better angels of our nature*. (pp. 285-301). Washington, DC US: American Psychological Association.
- Kawamoto, T., Ura, M., & Nittono, H. (2015). Intrapersonal and interpersonal processes of social exclusion. *Frontiers In Neuroscience, 9*, 62-62. doi: 10.3389/fnins.2015.00062
- Kempke, S., & Luyten, P. (2007). Psychodynamic and cognitive-behavioral approaches of obsessive-compulsive disorder: is it time to work through our ambivalence? *Bulletin Of The Menninger Clinic, 71*(4), 291-311. doi: 10.1521/bumc.2007.71.4.291

- Koster, E. H. W., De Lissnyder, E., Derakshan, N., & De Raedt, R. (2011). Understanding depressive rumination from a cognitive science perspective: The impaired disengagement hypothesis. *Clinical Psychology Review, 31*(1), 138-145. doi: 10.1016/j.cpr.2010.08.005
- Krienen, F. M., Tu, P.-C., & Buckner, R. L. (2010). Clan mentality: Evidence that the medial prefrontal cortex responds to close others. *The Journal of Neuroscience, 30*(41), 13906-13915. doi: 10.1523/JNEUROSCI.2180-10.2010
- Krizan, Z., & Johar, O. (2014). Narcissistic Rage Revisited. *Journal of Personality and Social Psychology*. doi: 10.1037/pspp0000013
- Lane, K. A., Banaji, M. R., Nosek, B. A., & Greenwald, A. G. (2007). Understanding and Using the Implicit Association Test: IV: What We Know (So Far) about the Method. In B. Wittenbrink, N. Schwarz, B. Wittenbrink & N. Schwarz (Eds.), *Implicit measures of attitudes*. (pp. 59-102). New York, NY US: Guilford Press.
- Leckelt, M., Küfner, A., Nestler, S., & Back, M. D. (in press). Behavioral processes underlying the decline of narcissists' popularity over time. *Journal of Personality & Social Psychology*.
- Leckman, J. F., Denys, D., Simpson, H. B., Mataix-Cols, D., Hollander, E., Saxena, S., Miguel, E. C., Rauch, S. L., Goodman, W. K., Phillips, K. A., & Stein, D. J. (2010). Obsessive-compulsive disorder: a review of the diagnostic criteria and possible subtypes and dimensional specifiers for DSM-V. *Depression & Anxiety (1091-4269), 27*(6), 507-527. doi: 10.1002/da.20669
- Lev-Ran, S., Shamay-Tsoory, S. G., Zangen, A., & Levkovitz, Y. (2012). Transcranial magnetic stimulation of the ventromedial prefrontal cortex impairs theory of mind learning. *European Psychiatry, 27*(4), 285-289. doi: 10.1016/j.eurpsy.2010.11.008
- Lischetzke, T., Angelova, R., & Eid, M. (2011). Validating an indirect measure of clarity of feelings: Evidence from laboratory and naturalistic settings. *Psychological Assessment, 23*(2), 447-455. doi: 10.1037/a0022211
- Lischetzke, T., Cuccodoro, G., Gauger, A., Todeschini, L., & Eid, M. (2005). Measuring Affective Clarity Indirectly: Individual Differences in Response Latencies of State Affect Ratings. *Emotion, 5*(4), 431-445. doi: 10.1037/1528-3542.5.4.431
- Logothetis, N. K. (2008). What we can do and what we cannot do with fMRI. *Nature, 453*(7197), 869-878. doi: 10.1038/nature06976
- Mancini, F., & Gangemi, A. (2011). Fear of deontological guilt and fear of contamination in obsessive compulsive disorder. *Psicoterapia Cognitiva Comportamentale, 17*(3), 395-404.
- McCullough, M. E. (2008). *Beyond revenge: The evolution of the forgiveness instinct*. San Francisco: Jossey-Bass.
- McCullough, M. E., Bellah, C. G., Kilpatrick, S. D., & Johnson, J. L. (2001). Vengefulness: Relationships with forgiveness, rumination, well-being, and the Big Five. *Personality and Social Psychology Bulletin, 27*(5), 601-610. doi: 10.1177/0146167201275008
- McCullough, M. E., & Hoyt, W. T. (2002). Transgression-related motivational dispositions: Personality substrates of forgiveness and their links to the Big Five. *Personality and Social Psychology Bulletin, 28*(11), 1556-1573. doi: 10.1177/014616702237583
- McCullough, M. E., Hoyt, W. T., & Rachal, K. C. (2000). What We Know (and Need to Know) about Assessing Forgiveness Constructs. In M. E. McCullough, K. I. Pargament & C. E. Thoresen (Eds.), *Forgiveness: Theory, research, and practice*. (pp. 65-88). New York, NY US: Guilford Press.
- McCullough, M. E., Kurzban, R., & Tabak, B. A. (2011). Evolved mechanisms for revenge and forgiveness. In P. R. Shaver & M. Mikulincer (Eds.), *Human aggression and*

- violence: *Causes, manifestations, and consequences*. (pp. 221-239). Washington, DC US: American Psychological Association.
- McCullough, M. E., Kurzban, R., & Tabak, B. A. (2013). Cognitive systems for revenge and forgiveness. *Behavioral and Brain Sciences*, *36*(1), 1-15. doi: 10.1017/S0140525X11002160
- McCullough, M. E., Pedersen, E. J., Tabak, B. A., & Carter, E. C. (2014). Conciliatory gestures promote forgiveness and reduce anger in humans. *PNAS Proceedings of the National Academy of Sciences of the United States of America*, *111*(30), 11211-11216. doi: 10.1073/pnas.1405072111
- McCullough, M. E., Rachal, K. C., Sandage, S. J., Worthington, E. L., Jr., Brown, S. W., & Hight, T. L. (1998). Interpersonal forgiving in close relationships: II. Theoretical elaboration and measurement. *Journal of Personality and Social Psychology*, *75*(6), 1586-1603. doi: 10.1037/0022-3514.75.6.1586
- McCullough, M. E., Root, L. M., & Cohen, A. D. (2006). Writing about the benefits of an interpersonal transgression facilitates forgiveness. *Journal of Consulting and Clinical Psychology*, *74*(5), 887-897. doi: 10.1037/0022-006x.74.5.887
- McCullough, M. E., & Worthington, E. L., Jr. (1999). Religion and the forgiving personality. *Journal of Personality*, *67*(6), 1141-1164.
- McCullough, M. E., Worthington, E. L., Jr., & Rachal, K. C. (1997). Interpersonal forgiving in close relationships. *Journal of Personality and Social Psychology*, *73*(2), 321-336. doi: 10.1037/0022-3514.73.2.321
- Meyer, M. L., Williams, K. D., & Eisenberger, N. I. (2015). Why Social Pain Can Live on: Different Neural Mechanisms Are Associated with Reliving Social and Physical Pain. *Plos One*, *10*(6), e0128294-e0128294. doi: 10.1371/journal.pone.0128294
- Miller, G. F. (2007). Sexual selection for moral virtues. *Quarterly Review of Biology*, *82*(2), 97-125.
- Mitchell, J. P., Macrae, C. N., & Banaji, M. R. (2006). Dissociable medial prefrontal contributions to judgments of similar and dissimilar others. *Neuron*, *50*(4), 655-663.
- Moritz, S., Jelinek, L., Hauschildt, M., & Naber, D. (2010). How to treat the untreated: effectiveness of a self-help metacognitive training program (myMCT) for obsessive-compulsive disorder. *Dialogues In Clinical Neuroscience*, *12*(2), 209-220.
- Moritz, S., Kempke, S., Luyten, P., Randjbar, S., & Jelinek, L. (2011). Was Freud partly right on obsessive-compulsive disorder (OCD)? Investigation of latent aggression in OCD. *Psychiatry Research*, *187*(1/2), 180-184. doi: 10.1016/j.psychres.2010.09.007
- Moritz, S., Niemeyer, H., Hottenrott, B., Schilling, L., & Spitzer, C. (2012). Interpersonal Ambivalence in Obsessive-Compulsive Disorder. *Behavioural and Cognitive Psychotherapy*, 1-16.
- Moritz, S., Wahl, K., Ertle, A., Jelinek, L., Hauschildt, M., Klinge, R., & Hand, I. (2009). Neither Saints nor Wolves in Disguise: Ambivalent Interpersonal Attitudes and Behaviors in Obsessive-Compulsive Disorder. *Behavior Modification*, *33*(2), 274-292.
- Moscovitch, D. A., McCabe, R. E., Antony, M. M., Rocca, L., & Swinson, R. P. (2008). Anger experience and expression across the anxiety disorders. *Depression And Anxiety*, *25*(2), 107-113.
- Nakao, T., Osumi, T., Ohira, H., Kasuya, Y., Shinoda, J., Yamada, J., & Northoff, G. (2010). Medial prefrontal cortex-dorsal anterior cingulate cortex connectivity during behavior selection without an objective correct answer. *Neuroscience Letters*, *482*(3), 220-224. doi: 10.1016/j.neulet.2010.07.041
- Newberg, A. B., d'Aquili, E. G., Newberg, S. K., & deMarici, V. (2000). The neuropsychological correlates of forgiveness. In M. E. McCullough, K. I. Pargament

- & C. E. Thoresen (Eds.), *Forgiveness: Theory, research, and practice*. (pp. 91-110). New York, NY US: Guilford Press.
- Nosek, B. A., Greenwald, A. G., & Banaji, M. R. (2005). Understanding and using the Implicit Association Test: II. Method variables and construct validity. *Personality & Social Psychology Bulletin*, *31*(2), 166-180.
- 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 US: Psychology Press.
- O'Reilly, J. X., Woolrich, M. W., Behrens, T. E. J., Smith, S. M., & Johansen-Berg, H. (2012). Tools of the trade: Psychophysiological interactions and functional connectivity. *Social Cognitive and Affective Neuroscience*, *7*(5), 604-609. doi: 10.1093/scan/nss055
- Ochsner, K. N., Beer, J. S., Robertson, E. R., Cooper, J. C., Gabrieli, J. D. E., Kihlstrom, J. F., & D'Esposito, M. (2005). The neural correlates of direct and reflected self-knowledge. *Neuroimage*, *28*(4), 797-814.
- Ohbuchi, K.-i., Kameda, M., & Agarie, N. (1989). Apology as Aggression Control: Its Role in Mediating Appraisal of and Response to Harm. *Journal of Personality & Social Psychology*, *56*(2), 219-227.
- Ohtsubo, Y., & Watanabe, E. (2009). Do sincere apologies need to be costly? Test of a costly signaling model of apology. *Evolution and Human Behavior*, *30*(2), 114-123.
- Orth, U., Montada, L., & Maercker, A. (2006). Feelings of revenge, retaliation motive, and posttraumatic stress reactions in crime victims. *Journal of Interpersonal Violence*, *21*(2), 229-243. doi: 10.1177/0886260505282286
- Painuly, N. P., Grover, S., Mattoo, S. K., & Gupta, N. (2012). Anger attacks in obsessive compulsive disorder. *Industrial Psychiatry Journal*, *20*(2), 115-119. doi: 10.4103/0972-6748.102501
- Paulhus, D. L., & John, O. P. (1998). Egoistic and moralistic biases in self-perception: The interplay of self-deceptive styles with basic traits and motives. *Journal of Personality*, *66*(6), 1025-1060. doi: 10.1111/1467-6494.00041
- Penke, L., & Asendorpf, J. B. (2008). Beyond global sociosexual orientations: A more differentiated look at sociosexuality and its effects on courtship and romantic relationships. *Journal of Personality and Social Psychology*, *95*(5), 1113-1135. doi: 10.1037/0022-3514.95.5.1113
- Perugini, M. (2005). Predictive models of implicit and explicit attitudes. *British Journal of Social Psychology*, *44*(1), 29-45. doi: 10.1348/014466604x23491
- 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*. (pp. 255-277). New York, NY US: Guilford Press.
- Phelps, E. A., O'Connor, K. J., Cunningham, W. A., Funayama, E. S., Gatenby, J. C., Gore, J. C., & Banaji, M. R. (2000). Performance on Indirect Measures of Race Evaluation Predicts Amygdala Activation. *Journal of Cognitive Neuroscience*, *12*(5), 729-738. doi: 10.1162/089892900562552
- Piacentini, J., Bergman, R. L., Keller, M., & McCracken, J. (2003). Functional impairment in children and adolescents with obsessive-compulsive disorder. *Journal of Child and Adolescent Psychopharmacology*, *13*(2,Suppl), S61-S69. doi: 10.1089/104454603322126359
- Pincus, A. L., & Lukowitsky, M. R. (2010). Pathological narcissism and narcissistic personality disorder. *Annual Review of Clinical Psychology*, *6*, 421-446. doi: 10.1146/annurev.clinpsy.121208.131215

- Purdon, C. (2011). Assessing Comorbidity, Insight, Family and Functioning in OCD. In G. Steketee (Ed.), *The Oxford Handbook of Obsessive Compulsive and Spectrum Disorders*. ox: Oxford University Press.
- Radomsky, A. S., Ashbaugh, A. R., & Gelfand, L. A. (2007). Relationships between anger, symptoms, and cognitive factors in OCD checkers. *Behaviour Research and Therapy*, 45(11), 2712-2725. doi: 10.1016/j.brat.2007.07.009
- Raskin, R., & Hall, C. S. (1979). A narcissistic personality inventory. *Psychological Reports*, 45(2), 590-590.
- Raskin, R., & Novacek, J. (1991). Narcissism and the use of fantasy. *Journal Of Clinical Psychology*, 47(4), 490-499.
- Ratcliff, R. (1993). Methods for dealing with reaction time outliers. *Psychological Bulletin*, 114(3), 510-532. doi: 10.1037/0033-2909.114.3.510
- Rhodewalt, F., & Morf, C. C. (1998). On self-aggrandizement and anger: A temporal analysis of narcissism and affective reactions to success and failure. *Journal of Personality and Social Psychology*, 74(3), 672-685. doi: 10.1037/0022-3514.74.3.672
- Rhodewalt, F., & Peterson, B. (2009). Narcissism. In M. R. Leary & R. H. Hoyle (Eds.), *Handbook of individual differences in social behavior*. (pp. 547-560). New York, NY US: Guilford Press.
- Ricciardi, E., Rota, G., Sanil, L., Gentili, C., Gaglianese, A., Guazzelli, M., & Pietrini, P. (2013). How the brain heals emotional wounds: The functional neuroanatomy of forgiveness. *Frontiers in Human Neuroscience*, 7.
- Riek, B. M., & Mania, E. W. (2012). The antecedents and consequences of interpersonal forgiveness: A meta - analytic review. *Personal Relationships*, 19(2), 304-325.
- Robinson, M. D., & Neighbors, C. (2006). Catching the Mind in Action: Implicit Methods in Personality Research and Assessment. In M. Eid & E. Diener (Eds.), *Handbook of multimethod measurement in psychology*. (pp. 115-125). Washington, DC US: American Psychological Association.
- Rotge, J.-Y., Lemogne, C., Hinfrey, S., Huguet, P., Grynszpan, O., Tartour, E., George, N., & Fossati, P. (2015). A meta-analysis of the anterior cingulate contribution to social pain. *Social Cognitive and Affective Neuroscience*, 10(1), 19-27. doi: 10.1093/scan/nsu110
- Rowe, A. D., Bullock, P. R., Polkey, C. E., & Morris, R. G. (2001). 'Theory of mind' impairments and their relationship to executive functioning following frontal lobe excisions. *Brain: A Journal of Neurology*, 124(3), 600-616. doi: 10.1093/brain/124.3.600
- Rudolph, A., Schröder-Abé, M., Riketta, M., & Schütz, A. (2010). Easier when done than said! Implicit self-esteem predicts observed or spontaneous behavior, but not self-reported or controlled behavior. *Journal of Psychology*, 218(1), 12-19. doi: 10.1027/0044-3409/a000003
- Rye, M. S., Pargament, K. I., Ali, M. A., Beck, G. L., Dorff, E. N., Hallisey, C., Narayanan, V., & Williams, J. G. (2000). Religious perspectives on forgiveness. In M. E. McCullough, K. I. Pargament & C. E. Thoresen (Eds.), *Forgiveness: Theory, research, and practice*. (pp. 17-40). New York, NY US: Guilford Press.
- Saxe, R., & Wexler, A. (2005). Making sense of another mind: The role of the right temporoparietal junction. *Neuropsychologia*, 43(10), 1391-1399. doi: 10.1016/j.neuropsychologia.2005.02.013
- Schmukle, S. C., & Egloff, B. (2005). A latent state-trait analysis of implicit and explicit personality measures. *European Journal of Psychological Assessment*, 21(2), 100-107.

- Schnabel, K., & Asendorpf, J. B. (2010). The self-concept: New insights from implicit measurement procedures. In B. Gawronski & B. K. Payne (Eds.), *Handbook of implicit social cognition: Measurement, theory, and applications*. (pp. 408-425). New York, NY US: Guilford Press.
- Schnabel, K., Asendorpf, J. B., & Greenwald, A. G. (2008). Using implicit association tests for the assessment of implicit personality self-concept. In G. J. Boyle, G. Matthews & D. H. Saklofske (Eds.), *The Sage handbook of personality theory and assessment. Vol. 2: Personality measurement and testing* (pp. 508-528). Los Angeles: Sage.
- Schnabel, K., Banse, R., & Asendorpf, J. B. (2006). Assessment of implicit personality self-concept using the Implicit Association Test (IAT): Concurrent assessment of anxiousness and anger. *British Journal of Social Psychology*, *45*(2), 373-396. doi: 10.1348/014466605x49159
- Schröder-Abé, M., Rudolph, A., & Schütz, A. (2007). High implicit self-esteem is not necessarily advantageous: Discrepancies between explicit and implicit self-esteem and their relationship with anger expression and psychological health. *European Journal of Personality*, *21*(3), 319-339. doi: 10.1002/per.626
- Seebauer, L., Froß, S., Dubaschny, L., Schönberger, M., & Jacob, G. A. (2014). Is it dangerous to fantasize revenge in imagery exercises? An experimental study. *Journal of Behavior Therapy & Experimental Psychiatry*, *45*(1), 20-25. doi: 10.1016/j.jbtep.2013.07.003
- Shamay-Tsoory, S. G., Tomer, R., Berger, B. D., & Aharon-Peretz, J. (2003). Characterization of Empathy Deficits following Prefrontal Brain Damage: The Role of the Right Ventromedial Prefrontal Cortex. *Journal of Cognitive Neuroscience*, *15*(3), 324-337. doi: 10.1162/089892903321593063
- Shenhav, A., Botvinick, M. M., & Cohen, J. D. (2013). The expected value of control: an integrative theory of anterior cingulate cortex function. *Neuron*, *79*(2), 217-240. doi: 10.1016/j.neuron.2013.07.007
- Shenhav, A., Straccia, M. A., Cohen, J. D., & Botvinick, M. M. (2014). Anterior cingulate engagement in a foraging context reflects choice difficulty, not foraging value. *Nature Neuroscience*, *17*(9), 1249-1254. doi: 10.1038/nn.3771
- Sheth, S. A., Mian, M. K., Patel, S. R., Asaad, W. F., Williams, Z. M., Dougherty, D. D., Bush, G., & Eskandar, E. N. (2012). Human dorsal anterior cingulate cortex neurons mediate ongoing behavioural adaptation. *Nature*, *488*(7410), 218-221. doi: 10.1038/nature11239
- Smith, S. M. (2002). Fast robust automated brain extraction. *Hum Brain Mapp*, *17*(3), 143-155. doi: 10.1002/hbm.10062
- Smith, S. M., Jenkinson, M., Woolrich, M. W., Beckmann, C. F., Behrens, T. E., Johansen-Berg, H., Bannister, P. R., De Luca, M., Drobnjak, I., Flitney, D. E., Niazy, R. K., Saunders, J., Vickers, J., Zhang, Y., De Stefano, N., Brady, J. M., & Matthews, P. M. (2004). Advances in functional and structural MR image analysis and implementation as FSL. *Neuroimage*, *23 Suppl 1*, S208-219. doi: 10.1016/j.neuroimage.2004.07.051
- Somerville, L. H., Heatherton, T. F., & Kelley, W. M. (2006). Anterior cingulate cortex responds differentially to expectancy violation and social rejection. *Nature Neuroscience*, *9*(8), 1007-1008. doi: 10.1038/nn1728
- Spielberger, C. D. (1988). *State Trait Anger Expression Inventory (STAXI)*. Odessa, FL: Psychological Assessment Resources.
- Spunt, R. P., Lieberman, M. D., Cohen, J. R., & Eisenberger, N. I. (2012). The Phenomenology of Error Processing: The Dorsal ACC Response to Stop-signal Errors Tracks Reports of Negative Affect. *Journal of Cognitive Neuroscience*, *24*(8), 1753-1765.

- Stanley, D., Phelps, E., & Banaji, M. (2008). The Neural Basis of Implicit Attitudes. *Current Directions in Psychological Science*, 17(2), 164-170. doi: 10.1111/j.1467-8721.2008.00568.x
- Storch, E. A., Jones, A. M., Lack, C. W., Ale, C. M., Sulkowski, M. L., Lewin, A. B., De Nadai, A. S., & Murphy, T. K. (2012). Rage attacks in pediatric obsessive-compulsive disorder: phenomenology and clinical correlates. *Journal Of The American Academy Of Child And Adolescent Psychiatry*, 51(6), 582-592. doi: 10.1016/j.jaac.2012.02.016
- Storch, E. A., Ledley, D. R., Lewin, A. B., Murphy, T. K., Johns, N. B., Goodman, W. K., & Geffken, G. R. (2006). Peer Victimization in Children With Obsessive–Compulsive Disorder: Relations With Symptoms of Psychopathology. *Journal of Clinical Child & Adolescent Psychology*, 35(3), 446-455. doi: 10.1207/s15374424jccp3503_10
- Strack, F., & Deutsch, R. (2004). Reflective and Impulsive Determinants of Social Behavior. *Personality & Social Psychology Review (Lawrence Erlbaum Associates)*, 8(3), 220-247.
- Strang, S., Utikal, V., Fischbacher, U., Weber, B., & Falk, A. (2014). Neural Correlates of Receiving an Apology and Active Forgiveness: An fMRI Study. *PLoS ONE*, 9(2), 1-7. doi: 10.1371/journal.pone.0087654
- Tang, Y.-Y., Hölzel, B. K., & Posner, M. I. (2015). The neuroscience of mindfulness meditation. *Nature Reviews Neuroscience*, 16(4), 213-225. doi: 10.1038/nrn3916
- Tsang, J.-A., McCullough, M. E., & Fincham, F. D. (2006). The longitudinal association between forgiveness and relationship closeness and commitment *Journal of Social & Clinical Psychology*, 25(4), 448-472.
- Twenge, J. M., & Campbell, W. K. (2003). "Isn't it fun to get the respect that we're going to deserve?" Narcissism, social rejection, and aggression. *Personality & Social Psychology Bulletin*, 29(2), 261-272.
- Ursu, S., Stenger, V. A., Shear, M. K., Jones, M. R., & Carter, C. S. (2003). Overactive action monitoring in obsessive-compulsive disorder: evidence from functional magnetic resonance imaging. *Psychological Science*, 14(4), 347-353.
- Van Overwalle, F. (2009). Social cognition and the brain: A meta-analysis. *Human Brain Mapping*, 30(3), 829-858. doi: 10.1002/hbm.20547
- Van Overwalle, F. (2011). A dissociation between social mentalizing and general reasoning. *NeuroImage*, 54(2), 1589-1599. doi: 10.1016/j.neuroimage.2010.09.043
- Watson, P. J., Grisham, S. O., Trotter, M. V., & Biderman, M. D. (1984). Narcissism and Empathy: Validity Evidence for the Narcissistic Personality Inventory. *Journal of Personality Assessment*, 48(3), 301.
- Watson, P. J., & Morris, R. J. (1991). Narcissism, empathy and social desirability. *Personality and Individual Differences*, 12(6), 575-579. doi: 10.1016/0191-8869(91)90253-8
- Welborn, B. L., & Lieberman, M. D. (2015). Person-specific theory of mind in medial pFC. *Journal of Cognitive Neuroscience*, 27(1), 1-12. doi: 10.1162/jocn_a_00700
- Whiteside, S. P., & Abramowitz, J. S. (2004). Obsessive–Compulsive Symptoms and the Expression of Anger. *Cognitive Therapy & Research*, 28(2), 259-268.
- Whiteside, S. P., & Abramowitz, J. S. (2005). The expression of anger and its relationship to symptoms and cognitions in obsessive–compulsive disorder. *Depression & Anxiety (1091-4269)*, 21(3), 106-111. doi: 10.1002/da.20066
- Wilkowski, B. M., Hartung, C. M., Crowe, S. E., & Chai, C. A. (2012). Men don't just get mad; they get even: Revenge but not anger mediates gender differences in physical aggression. *Journal of Research in Personality*, 46(5), 546-555. doi: 10.1016/j.jrp.2012.06.001

- Wilkowski, B. M., Robinson, M. D., & Troop-Gordon, W. (2010). How does cognitive control reduce anger and aggression? The role of conflict monitoring and forgiveness processes. *Journal of Personality and Social Psychology, 98*(5), 830-840. doi: 10.1037/a0018962
- Will, G.-J., Crone, E. A., & Güroğlu, B. (2015). Acting on social exclusion: Neural correlates of punishment and forgiveness of excluders. *Social Cognitive and Affective Neuroscience, 10*(2), 209-218. doi: 10.1093/scan/nsu045
- Worthington, E. L. (2006). *Forgiveness and reconciliation: Theory and application*. New York, NY, US: Routledge/Taylor & Francis Group.
- Worthington, E. L., Witvliet, C. V. O., Pietrini, P., & Miller, A. J. (2007). Forgiveness, health, and well-being: a review of evidence for emotional versus decisional forgiveness, dispositional forgivingness, and reduced unforgiveness. *Journal Of Behavioral Medicine, 30*(4), 291-302.
- Yarkoni, T., Poldrack, R. A., Nichols, T. E., Van Essen, D. C., & Wager, T. D. (2011). Large-scale automated synthesis of human functional neuroimaging data. *Nat Methods, 8*(8), 665-670. doi: 10.1038/nmeth.1635
- Young, L., & Saxe, R. (2009). Innocent intentions: A correlation between forgiveness for accidental harm and neural activity. *Neuropsychologia, 47*(10), 2065-2072. doi: 10.1016/j.neuropsychologia.2009.03.020
- Ysseldyk, R., Matheson, K., & Anisman, H. (2007). Rumination: Bridging a gap between forgivingness, vengefulness, and psychological health. *Personality and Individual Differences, 42*(8), 1573-1584. doi: 10.1016/j.paid.2006.10.032

8. Appendix

Study 1

Reprinted from *Frontiers in Psychology*, Vol. 5, Fatfouta, R., Schröder-Abé, M., & Merkl, A., Forgiving, fast and slow: Validity of the implicit association test for predicting differential response latencies in a transgression-recall paradigm, 1-14, 2015, with permission from the authors.

Study 2

Reprinted from *Personality and Individual Differences*, Vol. 82, Fatfouta, R., How forgiveness affects processing time: Mediation by rumination about the transgression, 90-95, 2015, with permissions from Elsevier (License N° 3633131443663).

Study 3

Reprinted from *Personality and Individual Differences*, Vol. 75, Fatfouta, R., Gerlach, T. M., Schröder-Abé*, M., & Merkl*, A., Narcissism and lack of interpersonal forgiveness: The mediating role of state anger, state rumination, and state empathy, 36-40, 2015, with permissions from Elsevier (License N° 3633131259166). *equal contribution

Study 4

Preprint of Fatfouta*, R., & Gerlach*, T. M. (under review). Can the shrew be tamed? Narcissism, lack of forgiveness, and the moderating role of conciliatory gestures. *European Journal of Personality*. *equal contribution

Study 5

Reprinted from *Psychiatry Research*, Vol. 219, Fatfouta, R., & Merkl, A., Associations between obsessive-compulsive symptoms, revenge, and the perception of interpersonal transgressions, 316-321, 2014, with permissions from Elsevier (License N° 3633580608528).

Study 6

Preprint of Fatfouta, R., Meshi, D., Merkl, A., & Heekeren, H. R. (under review). Forgiveness involves reduced connectivity between medial prefrontal and dorsal anterior cingulate cortex. *Cerebral Cortex*.



Forgiving, fast and slow: validity of the implicit association test for predicting differential response latencies in a transgression-recall paradigm

Ramzi Fatfouta^{1,2*}, Michela Schröder-Abé³ and Angela Merkl^{1,4}

¹ Cluster of Excellence "Languages of Emotion," Freie Universität Berlin, Berlin, Germany

² Affective Neuroscience and Psychology of Emotion, Department of Education and Psychology, Freie Universität Berlin, Berlin, Germany

³ Department of Psychology, Technische Universität Darmstadt, Darmstadt, Germany

⁴ Department of Neurology with Experimental Neurology, Charité-Universitätsmedizin, Berlin, Germany

Edited by:

John M. Zelenski, Carleton University, Canada

Reviewed by:

Monika Fleischhauer, Technische Universität Dresden, Germany
Michael Dufner, University of Leipzig, Germany

*Correspondence:

Ramzi Fatfouta, Cluster of Excellence "Languages of Emotion," Freie Universität Berlin, Habelschwerdter Allee 45, 14195 Berlin, Germany
e-mail: ramzi.fatfouta@fu-berlin.de

This study examined the role of automaticity in forgiving a real-life offense. As an alternative to self-report, an Implicit Association Test (IAT) of forgiveness was developed. Implicit (IAT-measured) and explicit (self-reported) forgiveness predicted shorter response times of state forgiveness ratings. The forgiveness IAT was highly reliable, moderately stable over time, and demonstrated incremental validity. Results suggest that the newly introduced forgiveness IAT could advance personality research beyond what is known from self-report measures, further corroborating the notion of implicit forgiveness. Implications for personality assessment are discussed.

Keywords: forgiveness, implicit self-concept, automaticity, Implicit Association Test (IAT), response latency measures

INTRODUCTION

Individual differences in forgiveness have traditionally been assessed via self-report. Contemporary measures, however, can be biased in two fundamental ways. First, they are susceptible to social desirability (Hoyt and McCullough, 2005), and second, they are insensitive to processes outside of awareness (Bornstein, 2002; Moors and De Houwer, 2006). This latter point is problematic as one might explicitly state one has forgiven an offense, yet implicitly continue harboring a grudge (hollow forgiveness; Fincham, 2010). To remedy these limitations, indirect measures may prove useful.

Indirect measures tap into different psychological processes than direct measures (i.e., self-reports). In accordance with dual-process models (for an overview, see Strack and Deutsch, 2004; Back et al., 2009), human behavior is a joint function of both reflective and impulsive processes. Reflective processes operate slowly and refer to propositional or explicit representations of the self that result from reasoning. Impulsive processes operate fast and refer to associative or implicit representations of the self that are activated automatically (i.e., non-deliberatively) when encountering situational cues. While reflective processes can be assessed via direct measures, impulsive processes can be assessed indirectly. Notably, indirect measures have been found to provide increments in predictive validity beyond direct measures—which highlights their value for a more complete assessment of trait factors and for the prediction of trait-relevant behavior (e.g., Back et al., 2009; Rudolph et al., 2010; Fleischhauer et al., 2013).

The Implicit Association Test (IAT; Greenwald et al., 1998) is the most popular indirect measure of automatic associations

and has been shown to be relatively resistant to faking (Steffens, 2004; Röhner et al., 2011). Using a speeded categorization task, the IAT measures the (relative) association strength between a target concept (e.g., me–others) and an attribute dimension (e.g., forgiving–vengeful). The rationale is that when individuals strongly associate their self with congruent (e.g., me–forgiving) vs. incongruent (e.g., me–vengeful) attributes, categorization will be easier (i.e., faster). The IAT effect, defined as response-latency difference between congruent and incongruent pairings, thus reflects associative links between the self and a trait concept (in our case, forgiveness) and may be referred to as implicit self-concept of personality (Schnabel and Asendorpf, 2010). To date, however, no study has attempted to look at the implicit self-concept of forgiveness.

Existing research has conceptualized forgiveness as a deliberative process with a presumed endpoint, that is, one's decision to forgive a transgressor (Fincham, 2000; Fincham et al., 2005). Specifically, to forgive involves a prosocial change in individuals' thoughts, feelings, and motivations, whereby they become more positively disposed toward their transgressor (McCullough et al., 1998; Worthington and Wade, 1999). However, as Fincham et al. (2006, p. 422) noted, the focus on deliberative processes "may overlook aspects of forgiveness that occur outside of conscious awareness." Indeed, recent work has provided empirical support for this contention (for an overview, see Karremans and Lange, 2008). In close relationships, for instance, the inclination to forgive occurs automatically and without the individuals' intention. The reasoning is that individuals automatically associate forgiveness with beneficial outcomes (e.g., relationship

satisfaction), thereby reinforcing impulsive/automatic responding in a pro-relational manner (Karremans and Aarts, 2007, Studies 1 and 2). This is consistent with associative-learning processes in dual-process models (Strack and Deutsch, 2004; Back et al., 2009) and calls for a multimethod approach including indirect measures like the IAT. Consequently, it is necessary to examine both explicit *and* implicit aspects of forgiveness.

Based on these findings, the current study aimed to develop an IAT that taps individual differences in *implicit forgiveness* and to test its predictive validity over and above a corresponding self-report questionnaire. As forgiveness-related behavioral criteria, we used the response times (RTs) individuals need to rate their state forgiveness (i.e., their current thoughts and feelings toward an offender; McCullough et al., 1998). In past research, the measurement of response latency has been employed as an operative indicator of attitude strength (Fazio, 2001). The logic here is that RTs reflect the individual's mental effort to arrive at a response and therefore indicate the accessibility of information in memory. The more accessible the information is, the faster the individual responds (Fazio, 1990a,b). A personality-oriented adaptation of this approach has been provided by Lischetzke et al. (2005), who demonstrated that when individuals can easily access internal cues about their affective state they are faster to rate it.

With respect to forgiveness, a similar account has been proposed. Specifically, individuals who are consistently inclined to forgive are assumed to “have both a higher rate of forgiveness and a *shorter* latency of response” (Sutton and Thomas, 2006, p. 33; emphasis added). Hence, it appears plausible that when prompted (1) to recall an offense and then (2) to rate their state forgiveness, individuals who have already worked through the pain should more easily access internal cues about their forgiveness and, hence, exhibit faster responses. By contrast, individuals who have not yet achieved forgiveness (i.e., who find internal forgiveness cues to be rather difficult to access), should need comparatively more time to contemplate and report their responses. For example, responding to an item such as “I am finding it difficult to act warmly toward him/her” (McCullough et al., 2006) should be less effortful (i.e., RT should be faster) for a person who has come to terms with the hurt¹.

Response latencies reflect both deliberative and automatic processing (Fazio, 1990a). Therefore, we hypothesized that both the newly developed forgiveness IAT and a forgiveness self-report would predict RT of forgiveness ratings. Specifically, high-forgiving individuals—measured by self-report and indirectly—should be faster to rate their state forgiveness (*Hypothesis 1*). Given that incremental validity is an important psychometric property when evaluating indirect measures (Perugini and Banse,

2007), we tested whether the novel forgiveness IAT predicted RT of state forgiveness ratings beyond explicit (self-reported) forgiveness (*Hypothesis 2*).

MATERIALS AND METHODS

We report all data exclusions (if any), manipulations, and measures, and how we determined our sample size. The latter was determined a priori with the goal to obtain around 100 observations.

PARTICIPANTS

One hundred and four students (70 women; $M_{\text{age}} = 25.10$, $SD = 4.39$) of Freie Universität Berlin participated in exchange for money or course credit.

PROCEDURE

Data were collected individually or in sessions of two individuals. After providing written informed consent, participants completed the forgiveness IAT and a set of questionnaires on a computer screen described below. Each questionnaire item appeared separately on the screen. Participants answered by mouse click on the respective response category. They had to click a *next* button to proceed with the next item.

At the beginning of the computer-based questionnaire, participants answered questions to obtain a measure of their baseline speed (see Baseline speed). Participants then had (1) to think of a close person; (2) type in his/her name; (3) indicate the type of relationship involved, and (4) rate their perception of closeness to the self-selected other (Karremans and Aarts, 2007). Next, participants were instructed to bring to mind a real-life situation in which the other had hurt them and to rate their state forgiveness. RT (i.e., the time between item presentation and click on the *next* button) was unobtrusively recorded (Lischetzke et al., 2005). Subsequently, participants rated the severity of the transgression and completed measures of explicit forgiveness as well as social desirability. Finally, participants were thanked and debriefed.

MEASURES

Implicit forgiveness

Implicit forgiveness was measured with an IAT (Greenwald et al., 1998). The forgiveness IAT was specifically designed for this study and comprised five blocks (for procedural details, see **Table 1**). In each trial, a stimulus word was presented in the center of the screen. Participants then pressed either a left or a right key to categorize the stimulus as quickly and accurately as possible into one of the categories. The category labels appeared in the upper left and right hand corners of the screen.

The first two blocks consisted of a simple discrimination task, in which participants practiced correctly categorizing stimuli from the target category (me–others) and attribute category (forgiving–vengeful). The third block combined both discrimination tasks (i.e., me–forgiving for the left key; others–vengeful for the right key). In block 4, the labels of the attribute categories were reversed (vengeful–forgiving). Block 5 consisted of the reversed combined discrimination task (i.e., me–vengeful for the left key; others–forgiving for the right key).

¹To demonstrate that higher forgiveness is indeed related to faster responses, we performed a pilot study ($N = 120$, $M_{\text{age}} = 22.47$, $SD = 4.33$), instructing participants to recall an experienced hurt and to rate their motivation to seek revenge against the transgressor (while RT was unobtrusively measured). Participants then indicated their forgiveness (i.e., “I have forgiven what he/she did to me”). As predicted, RT was significantly negatively correlated with forgiveness ($r = -0.20$, $p = 0.031$).

Table 1 | Forgiveness IAT: Task Sequence and Stimuli.

Block	Task	Target concepts	Attribute concepts ^a	Key assignment ^b	
				Left (A)	Right (5)
1 (20)	Target discrimination	Me–others		Me, my, own, I, self	They, your, them, you, others
2 (20)	Attribute discrimination		Forgiving–Vengeful	Forgiving, conciliatory, merciful, gracious, lenient	Vengeful, punitive, hostile, merciless, unforgiving
3 (60)	Combined discrimination	Me–others	Forgiving–Vengeful	Me, my, own, I, self Forgiving, conciliatory, merciful, gracious, lenient	They, your, them, you, others Vengeful, punitive, hostile, merciless, unforgiving
4 (40)	Reversed attribute discrimination		Vengeful–Forgiving	Vengeful, punitive, hostile, merciless, unforgiving	Forgiving, conciliatory, merciful, gracious, lenient
5 (60)	Combined reversed discrimination	Me–others	Vengeful–Forgiving	Me, my, own, I, self Vengeful, punitive, hostile, merciless, unforgiving	They, your, them, you, others Forgiving, conciliatory, merciful, gracious, lenient

IAT, Implicit Association Test. The number of trials per block is in parentheses. Critical blocks are in bold. ^aStimuli were piloted ($N = 50$, $M_{age} = 24.96$, $SD = 4.43$) and matched according to familiarity, valence, similarity, and potency. ^bNumber 5 from the numeric keyboard.

The order of critical blocks (i.e., blocks 3 and 5) was held constant to maximize the reliability of person effects (Banse, 2001). IAT-effects were computed using the improved scoring algorithm (D_1 ; Greenwald et al., 2003). Higher scores reflect stronger automatic associations between me–forgiving (vs. me–vengeful), and hence, a more forgiving implicit self-concept.

To calculate the forgiveness IAT's reliability, we applied the D_1 -algorithm to two mutually exclusive subsets of the critical trials (Schmukle and Egloff, 2006). The Spearman-Brown corrected split-half correlation of this two-part measure was close to excellent ($r_{tt} = 0.89$). To provide an estimate of temporal stability, the last 30 participants (19 women; $M_{age} = 26.60$, $SD = 4.35$) were re-contacted after 1 month ($M_{time\ interval} = 32.70$ days, $SD = 3.70$) to complete the forgiveness IAT in a follow-up assessment. The temporal stability was moderate (0.50; 95% confidence interval = 0.34; 0.63) and similar to that previously reported for IATs (Egloff et al., 2005).

Baseline speed

Because we used RT as an individual difference measure, we controlled for differential baseline response speed (Fazio, 1990b). Baseline response speed was assessed by means of seven easy-knowledge questions [e.g., “What type of celestial body is the earth? (1) White giant; (2) Planet; (3) Asteroid; (4) Moon”; (Lisetzke et al., 2005)] to which the correct answer was evident, so that RT ($\alpha = 0.71$) tapped only differential reading speed and motor behavior (e.g., muscle speed). Mean item accuracy was high (99.36%, $SD = 0.03$), thus corroborating this idea. Consistent with Lisetzke and colleagues, we omitted the first (training) item and calculated the median of the remaining six RT

scores. Further validity evidence for this baseline speed measure can be found in Lisetzke et al. (2011).

Closeness

Closeness to the transgressor was assessed with the Inclusion of the Other in the Self scale (IOS; Aron et al., 1992), a pictorial measure comprising seven pairs of two increasingly overlapping circles, labeled “Self” and “Other.” Each pair was numbered and participants had to click on the respective response option to indicate their perceived closeness (1 = *not at all close*, 7 = *very close*). The IOS has demonstrated high alternate-form and test–retest stability as well as strong convergence with multi-item measures of relational closeness (Aron et al., 1992).

State forgiveness

State forgiveness was measured with the Transgression-Related Interpersonal Motivations Inventory (TRIM; McCullough et al., 1998, 2006). It assesses participants' motivational changes toward a specific transgressor and is divided into three subscales, with five items measuring revenge (e.g., “I'm going to get even”), seven items measuring avoidance (e.g., “I keep as much distance between us as possible”), and six items measuring benevolence (“I have given up my hurt and resentment”). Items were rated on a 7-point Likert scale (1 = *strongly disagree*, 7 = *strongly agree*). Past research has evidenced that the subscales have high internal consistency, moderate temporal stability, and evidence of construct validity (McCullough et al., 1998, 2006; McCullough and Hoyt, 2002). In this sample, internal consistency (α) was good for avoidance and benevolence (0.87 and 0.76, respectively), but somewhat modest for revenge (0.53). While this alpha is less than ideal, similar estimates have been reported with the revenge scale ($\alpha = 0.52$;

Allemand et al., 2013). Unlike previous research, however, our primary analyses focused on participants' RTs, and hence, this value is deemed reasonable.

RT of state forgiveness ratings

As an objective measure of state forgiveness, RT to each TRIM item was recorded. Again, the first item was regarded as a training item and its RT was excluded from the analyses. We then log-transformed ($\log_e[x]$) RTs for each person to correct for positive skewness (Fazio, 1990b) and calculated mean RT scores for each of the three TRIM subscales: RT_{Revenge} ($\alpha = 0.65$), $RT_{\text{Avoidance}}$ ($\alpha = 0.78$), and $RT_{\text{Benevolence}}$ ($\alpha = 0.70$). Using three regression equations, we computed—separately for each TRIM subscale—residual RT scores by partialling out baseline speed. Values greater (smaller) than zero indicate that an individual is slower (faster) than predicted by his or her baseline speed. Specifically, the residual (i.e., baseline-corrected) RT scores reflect the speed with which individuals rate their revenge, avoidance, and benevolence, respectively.

Transgression severity

On completion of the TRIM, participants indicated how severe they perceived their transgression to be on a face-valid, single-item scale (Ghaemmaghami et al., 2011). Specifically, they were asked: "How painful was the transgression to you at the time it occurred?" Responses were made on a 7-point scale (1 = *not painful at all*, 7 = *very painful*).

Explicit forgiveness

Explicit forgiveness was measured with the four-item Tendency to Forgive Scale (TTF; Brown, 2003; e.g., "I tend to get over it quickly when someone hurts my feelings," $\alpha = 0.76$). Participants responded to each item on a 7-point Likert scale (1 = *strongly disagree*, 7 = *strongly agree*). The TTF has consistently demonstrated good psychometric properties, including strong predictive validity across several studies (Brown, 2003; Brown and Phillips, 2005).

Social desirability

The Balanced Inventory of Desirable Responding (BIDR; Paulhus, 1994; German version by Musch et al., 2002) was used to assess both factors of social desirability, impression management (10 items; e.g., "I never swear," $\alpha = 0.76$) and self-deceptive enhancement (10 items; e.g., "My first impressions of people usually turn out to be right," $\alpha = 0.62$). Participants rated each item on a 7-point Likert scale (1 = *strongly disagree*, 7 = *strongly agree*). We computed continuous (vs. dichotomized) subscale scores due to their superior psychometric properties (Stöber et al., 2002). The BIDR is widely used and both subscales have provided evidence of reliability and validity in past research (Paulhus, 1984, 1994; Musch et al., 2002).

STATISTICAL ANALYSES

We analyzed bivariate correlations and conducted hierarchical multiple regression analyses using implicit and explicit forgiveness to predict residual (i.e., baseline-corrected) RT scores (i.e., RT_{Revenge} , $RT_{\text{Avoidance}}$, and $RT_{\text{Benevolence}}$). All statistical tests adopted a significance level of $\alpha = 0.05$ (two-tailed).

RESULTS

DESCRIPTIVE STATISTICS AND CORRELATIONS AMONG MEASURES

Participants reported transgressions pertaining to romantic partners (38.5%), close friends (37.5%), parents (10.6%), other family members or relatives (10.6%), coworker, acquaintance or neighbor (1.0%), or other person (1.9%). Table 2 details descriptive statistics for all measures.

Intercorrelations are presented in Table 3. Implicit and explicit forgiveness were not significantly correlated, a result that parallels findings on implicit-explicit relations in other research domains such as self-esteem (Bosson et al., 2000). Importantly, implicit forgiveness was unrelated to both social- desirability components. Explicit forgiveness, by contrast, correlated significantly positively with self-deceptive enhancement.

PREDICTING RT FROM IMPLICIT AND EXPLICIT FORGIVENESS

Our main predictions were that both implicit and explicit forgiveness would predict shorter RT of forgiveness ratings (*Hypothesis 1*) and that implicit forgiveness would demonstrate increments in predictive validity (*Hypothesis 2*). To test these hypotheses, we computed three hierarchical multiple regression analyses. Each residual RT score (i.e., RT_{Revenge} , $RT_{\text{Avoidance}}$, $RT_{\text{Benevolence}}$) was treated as a separate criterion and the trait forgiveness measures were entered in two steps (Step 1: explicit forgiveness; Step 2: implicit forgiveness). As such, we were able to estimate the

Table 2 | Means, Standard Deviations, and Actual Ranges of all Study Variables.

Variable	<i>M</i>	<i>SD</i>	Min	Max
BASELINE SPEED (S)^a				
RT to knowledge items	5.87	1.40	3.55	10.02
TRAIT FORGIVENESS				
Forgiveness IAT (D ₁)	0.47	0.25	-0.20	1.03
TTF	3.79	1.24	1.25	6.75
STATE FORGIVENESS (TRIM)				
Revenge	1.66	0.66	1.00	3.80
Avoidance	2.36	1.12	1.00	5.57
Benevolence	5.89	0.76	3.50	7.00
RT OF STATE FORGIVENESS RATINGS (S)^a				
RT_{Revenge}	4.89	1.37	3.00	8.75
$RT_{\text{Avoidance}}$	4.57	1.11	2.81	8.06
$RT_{\text{Benevolence}}$	5.68	1.32	3.48	9.49
TRANSGRESSION-RELATED CHARACTERISTICS^b				
Closeness (IOS)	5.23	1.17	3.00	7.00
Transgression severity	4.93	1.65	1.00	7.00
SOCIAL DESIRABILITY (BIDR)				
Self-deceptive enhancement	4.05	0.72	2.50	6.00
Impression management	3.75	0.88	1.60	6.20

IAT, Implicit Association Test; TTF, Tendency to Forgive Scale; TRIM, Transgression-Related Interpersonal Motivations Inventory; RT, response time; IOS, Inclusion of the Other in the Self Scale; BIDR, Balanced Inventory of Desirable Responding; Possible range of self-report scores: 1–7. ^aFor clarity of presentation, these values represent raw (mean) latencies. ^bSingle-item measures.

Table 3 | Intercorrelations Among Study Variables.

Variable	1	2	3	4	5	6	7	8	9	10	11	12
TRAIT FORGIVENESS												
1. Forgiveness IAT (D ₁)	–											
2. TTF	0.00	–										
STATE FORGIVENESS (TRIM)												
3. Revenge	–0.04	–0.29**	–									
4. Avoidance	–0.05	–0.26**	0.37***	–								
5. Benevolence	0.01	0.28**	–0.37***	–0.51***	–							
RESIDUAL RT OF STATE FORGIVENESS RATINGS^a												
6. Residual RT _{Revenge}	–0.22*	–0.24*	0.36***	0.22*	–0.07	–						
7. Residual RT _{Avoidance}	0.01	–0.17 [#]	0.12	0.40***	–0.15	0.58***	–					
8. Residual RT _{Benevolence}	–0.18 [#]	–0.27**	–0.02	0.10	–0.02	0.43***	0.54***	–				
TRANSGRESSION-RELATED CHARACTERISTICS^b												
9. Closeness (IOS)	–0.11	–0.02	–0.17 [#]	–0.20*	0.25*	–0.16	–0.15	–0.10	–			
10. Transgression severity	–0.18 [#]	–0.16	–0.09	–0.05	–0.12	0.14	0.02	0.02	0.04	–		
SOCIAL DESIRABILITY (BIDR)												
11. Self-deceptive enhancement	–0.10	0.23*	0.03	–0.01	0.06	–0.04	–0.05	–0.07	0.14	–0.07	–	
12. Impression management	–0.04	0.02	–0.08	–0.19 [#]	0.09	0.00	–0.01	0.00	0.14	0.01	0.15	–

IAT, Implicit Association Test; TTF, Tendency to Forgive Scale; TRIM, Transgression-Related Interpersonal Motivations Inventory; RT, response time; IOS, Inclusion of the Other in the Self Scale; BIDR, Balanced Inventory of Desirable Responding; Possible range of self-report scores: 1–7. ^aBaseline speed partialled out. ^bSingle-item measures. [#]*p* < 0.10. **p* < 0.05. ***p* < 0.01. ****p* < 0.001.

incremental proportion of variance in residual RT accounted for by implicit forgiveness, after controlling for the effect of explicit forgiveness. We obtained qualitatively identical results when controlling for closeness and transgression severity in the first step of the models. Also, the interaction between implicit and explicit forgiveness yielded no significant effects. In the interest of parsimony, these variables were trimmed from the final analyses. **Table 4** details the results.

The first regression model revealed that both implicit and explicit forgiveness significantly predicted residual RT_{Revenge}. As expected, individuals high in implicit and explicit forgiveness were faster to rate their revenge motivations. Introducing implicit forgiveness in Step 2 added significantly to the prediction of residual RT_{Revenge}, over and above explicit forgiveness. The full model explained 10.7% of variance in residual RT_{Revenge}.

The second regression model with residual RT_{Avoidance} as criterion revealed that explicit forgiveness predicted the time individuals needed to rate their avoidance motivations, but implicit forgiveness did not. Consequently, the addition of implicit forgiveness in Step 2 explained no variation in residual RT_{Avoidance}. The restricted model (i.e., the model containing only explicit forgiveness as predictor) explained a small (2.9%) and marginally significant proportion of variance in the RT measure.

The model predicting residual RT_{Benevolence} was examined last. As expected, high levels of implicit and explicit forgiveness were associated with shorter RT for items tapping benevolence. The addition of implicit forgiveness in Step 2, however, contributed only a marginally significant amount of variance to the prediction of residual RT_{Benevolence}, beyond explicit forgiveness. The full model explained 10.2% of variance in residual RT_{Benevolence}.

In summary, *Hypothesis 1* was generally supported for explicit forgiveness: among individuals high in explicit forgiveness, RT

Table 4 | Hierarchical Multiple Regression Analyses Predicting Response Time by Implicit and Explicit Forgiveness.

Predictor	Residual RT of State Forgiveness Ratings ^a					
	RT _{Revenge} ^b		RT _{Avoidance} ^c		RT _{Benevolence} ^d	
	ΔR ²	β	ΔR ²	β	ΔR ²	β
Step 1	0.060*		0.029 [#]		0.071**	
TTF		–0.24*		–0.17 [#]		–0.27**
Step 2	0.047*		0.000		0.031 [#]	
TTF		–0.24*		–0.17 [#]		–0.27**
Forgiveness IAT		–0.22*		0.01		–0.18 [#]

RT, response time; TTF, Tendency to Forgive Scale; IAT, Implicit Association Test; β, standardized beta coefficient; ΔR², change in explained variance from one step to the next. ^aBaseline speed partialled out. ^bStep 1: *F*_(1, 102) = 6.46, *p* = 0.013; Step 2: *F*_(2, 101) = 6.05, *p* = 0.003; ^cStep 1: *F*_(1, 102) = 3.03, *p* = 0.085; Step 2: *F*_(2, 101) = 1.50, *p* = 0.227; ^dStep 1: *F*_(1, 102) = 7.78, *p* = 0.006; Step 2: *F*_(2, 101) = 5.71, *p* = 0.004. [#]*p* < 0.10. **p* < 0.05. ***p* < 0.01.

was shorter for all three TRIM dimensions (i.e., revenge, avoidance, benevolence) and significantly so for revenge and benevolence. As regards implicit forgiveness, *Hypotheses 1* and *2* were supported for one of the three RT scores (i.e., residual RT_{Revenge}), providing suggestive evidence for the indirect forgiveness measure in terms of incremental predictive validity.

DISCUSSION

In the present study, we analyzed the predictive and incremental validity of a non-self-report (indirect) measure of forgiveness, the forgiveness IAT. As forgiveness-related behavioral criteria,

we examined differential response latencies of state forgiveness ratings.

The new forgiveness IAT provides an indirect assessment of forgiveness by measuring chronically accessible self-associations pertaining to forgiving vs. vengeful. Noteworthy, we found that the IAT and the TTF scale were virtually unrelated. This may indicate that both measures tap different modes of representation (i.e., automatic vs. controlled; Moors and De Houwer, 2006) and different content (i.e., the stimuli included in both instruments). Notably, such factors (i.e., measurement correspondence and processing mode) have been shown to reduce the correlation between direct and indirect measures (Hofmann et al., 2005).

As desired for personality assessment, the forgiveness IAT demonstrated high split-half reliability, comparable with coefficients in related research on the implicit personality self-concept (Egloff and Schmukle, 2002; Back et al., 2009) and similar to those of explicit forgiveness measures (Brown, 2003; Brown and Phillips, 2005). Concerning temporal stability (i.e., test–retest reliability), IAT scores demonstrated a moderate convergence across measurement occasions. The 1-month retest correlation in the current study ($r = 0.50$) was in the range typically found for IATs ($r = 0.25–0.69$, with mean and median estimates of 0.50; Lane et al., 2007). Given the small test–retest sample size, however, this estimate should be viewed with caution.

As to the prediction of behavior, results demonstrated that among individuals high in explicit forgiveness, RT was shorter for two of the three TRIM dimensions (i.e., revenge and benevolence). For implicit forgiveness, hypotheses were supported for one of the three RT scores (i.e., residual RT_{Revenge}). Lack of conceptual correspondence (Ajzen and Fishbein, 1977) might explain why implicit forgiveness did not predict residual RT_{Avoidance}, given that the concept of avoidance is not directly tapped with our IAT.

Overall, our findings are consistent with well-established evidence that indirect personality measures allow prediction of variance in behavior beyond what is already predicted by self-reports (Back et al., 2009; Greenwald et al., 2009). Although not significantly related to TTF scores, the forgiveness IAT provided incremental insights into trait-relevant criteria that would have been overlooked if our focus had been restricted to self-report. This parallels recent results from the fields of need for cognition or self-esteem (Rudolph et al., 2010; Fleischhauer et al., 2013), which demonstrated that indirect measures (although not significantly associated with direct measures) incrementally explained variance in relevant behavior.

Apart from those results, some methodological considerations need to be discussed. For example, our criterion measures were based on response latencies and not on *actual* forgiving behavior. Although important to mention, prior personality research using the IAT similarly operationalized behavior with RTs. In one study on implicit self-esteem, for example, the time individuals spent reading social feedback has been taken as an objective measure of defensiveness (Schröder-Abé et al., 2007; study 2). Additionally, unlike other personality traits (e.g., neuroticism) that are comparatively easy to observe (e.g., non-verbal nervousness; Back et al., 2009), forgiveness is an intrinsically processual phenomenon that unfolds over time (McCullough et al., 2003), thus not manifested

in a specific act (Fincham, 2000). Therefore, actual behavior might be less suited for further research on implicit forgiveness.

It might be valuable, however, to examine other indices, such as physiological reactivity. Johnston et al. (2013), for instance, have recently demonstrated that implicit—but not explicit—moral identity predicted increases in blood pressure and heart rate in response to moral transgressions. Given evidence that the same physiological indicants are associated with unforgiving responses (Witvliet et al., 2001), one testable prediction is that the forgiveness IAT should predict individual variation in physiological responses to *personal* transgressions.

Together, our research has important implications for personality theory and assessment. First, it suggests that an individual's implicit cognition may influence the way he or she responds to transgression. This resonates with recent work on the role of implicit processes in forgiveness (Karremans and Aarts, 2007). Second, our findings provide an example of how direct and indirect measures complement each other in targeting implicit *and* explicit personality features, thereby yielding a more comprehensive understanding of behavior (Bornstein, 2002). Future research needs to place the construct of *implicit forgiveness* within a broader nomological network, further establishing its construct validity with respect to related constructs (e.g., empathy for the offender). Third, the IAT's resistance to social desirability underscores its utility for assessing even constructs as desirable as forgiveness. Last, the results of this study demonstrate that implicit associations may contribute a valuable, albeit understudied, source of information in understanding the implicit self-concept of forgiveness. Therefore, we emphasize the importance of administering multiple assessment formats, especially in applied (e.g., clinical) contexts. Promising work has already begun on the differential treatment sensitivity of implicit (vs. explicit) personality measures (Gamer et al., 2008) and should be further explored.

CONCLUSION

The present study developed an indirect measure of forgiveness—the forgiveness IAT—and examined its incremental validity for the prediction of differential response latencies in a transgression-recall paradigm. We thereby complement previous work on the role of automatic processes in forgiveness and, importantly, provide an addition to existing self-reports. We found preliminary evidence that implicit forgiveness explained additional variance above and beyond a corresponding explicit measure. Together, our findings provide novel insights that will hopefully stimulate both research and practice of implicit personality functioning.

ACKNOWLEDGMENTS

We wish to thank our reviewers whose constructive comments helped to improve this article. The research was financially supported by the Cluster of Excellence “Languages of Emotion” at Freie Universität Berlin, funded by the DFG (German Research Foundation).

REFERENCES

- Ajzen, I., and Fishbein, M. (1977). Attitude-behavior relations: a theoretical analysis and review of empirical research. *Psychol. Bull.* 84, 888–918. doi: 10.1037/0033-2909.84.5.888

- Allemand, M., Steiner, M., and Hill, P. L. (2013). Effects of a forgiveness intervention for older adults. *J. Couns. Psychol.* 60, 279–286. doi: 10.1037/a0031839
- Aron, A., Aron, E. N., and Smollan, D. (1992). Inclusion of other in the self scale and the structure of interpersonal closeness. *J. Pers. Soc. Psychol.* 63, 596–612. doi: 10.1037/0022-3514.63.4.596
- Back, M. D., Schmukle, S. C., and Egloff, B. (2009). Predicting actual behavior from the explicit and implicit self-concept of personality. *J. Pers. Soc. Psychol.* 97, 533–548. doi: 10.1037/a0016229
- Banse, R. (2001). Affective priming with liked and disliked persons: prime visibility determines congruency and incongruency effects. *Cogn. Emot.* 15, 501–520. doi: 10.1080/02699930126251
- Bornstein, R. F. (2002). A process dissociation approach to objective-projective test score interrelationships. *J. Pers. Assess.* 78, 47–68. doi: 10.1207/s15327752JPA7801_04
- Bosson, J. K., Swann, W. B., and Pennebaker, J. W. (2000). Stalking the perfect measure of implicit self-esteem: the blind men and the elephant revisited? *J. Pers. Soc. Psychol.* 79, 631–643. doi: 10.1037/0022-3514.79.4.631
- Brown, R. P. (2003). Measuring individual differences in the tendency to forgive: construct validity and links with depression. *Pers. Soc. Psychol. Bull.* 29, 759–771. doi: 10.1177/0146167203029006008
- Brown, R. P., and Phillips, A. (2005). Letting bygones be bygones: further evidence for the validity of the tendency to forgive scale. *Pers. Individ. Diff.* 38, 627–638. doi: 10.1016/j.paid.2004.05.017
- Egloff, B., and Schmukle, S. C. (2002). Predictive validity of an implicit association test for assessing anxiety. *J. Pers. Soc. Psychol.* 83, 1441–1455. doi: 10.1037/0022-3514.83.6.1441
- Egloff, B., Schwerdtfeger, A., and Schmukle, S. C. (2005). Temporal stability of the implicit association test-anxiety. *J. Pers. Assess.* 84, 82–88. doi: 10.1207/s15327752jpa8401_14
- Fazio, R. (1990a). Multiple processes by which attitudes guide behavior: the MODE Model as an integrative framework. *Adv. Exp. Soc. Psychol.* 23, 75–109.
- Fazio, R. H. (1990b). “A practical guide to the use of response latency in social psychological research,” in *Research Methods in Personality and Social Psychology*, eds C. Hendrick and M. S. Clark (Thousand Oaks, CA: Sage Publications, Inc.), 74–97.
- Fazio, R. H. (2001). On the automatic activation of associated evaluations: an overview. *Cogn. Emot.* 15, 115–141. doi: 10.1080/02699930125908
- Fincham, F. D. (2000). The kiss of the porcupines: from attributing responsibility to forgiving. *Pers. Relat.* 7, 1–23. doi: 10.1111/j.1475-6811.2000.tb00001.x
- Fincham, F. D. (2010). “Forgiveness: integral to a science of close relationships?” in *Prosocial Motives, Emotions, and Behavior: The Better Angels of Our Nature*, eds M. Mikulincer and P. R. Shaver (Washington, DC: American Psychological Association), 347–365. doi: 10.1037/12061-018
- Fincham, F. D., Hall, J., and Beach, S. R. H. (2006). Forgiveness in marriage: current status and future directions. *Family Relat.* 55, 415–427. doi: 10.1111/j.1741-3729.2005.callf.x-i1
- Fincham, F. D., Hall, J. H., and Beach, S. R. H. (2005). “Til lack of forgiveness doth us part: forgiveness in marriage,” in *Handbook of Forgiveness* ed E. L. Worthington (New York, NY: Routledge), 207–226.
- Fleischhauer, M., Strobel, A., Enge, S., and Strobel, A. (2013). Assessing implicit cognitive motivation: developing and testing an implicit association test to measure need for cognition. *Eur. J. Pers.* 27, 15–29. doi: 10.1002/per.1841
- Gamer, J., Schmukle, S. C., Luka-Krausgrill, U., and Egloff, B. (2008). Examining the dynamics of the implicit and the explicit self-concept in social anxiety: changes in the implicit association test-anxiety and the social phobia anxiety inventory following treatment. *J. Pers. Assess.* 90, 476–480. doi: 10.1080/00223890802248786
- Ghaemmaghami, P., Allemand, M., and Martin, M. (2011). Forgiveness in younger, middle-aged and older adults: age and gender matters. *J. Adult Dev.* 18, 192–203. doi: 10.1007/s10804-011-9127-x
- Greenwald, A. G., McGhee, D. E., and Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: the implicit association test. *J. Pers. Soc. Psychol.* 74, 1464–1480. doi: 10.1037/0022-3514.74.6.1464
- Greenwald, A. G., Nosek, B. A., and Banaji, M. R. (2003). Understanding and using the implicit association test: I. An improved scoring algorithm. *J. Pers. Soc. Psychol.* 85, 197–216. doi: 10.1037/0022-3514.85.2.197
- Greenwald, A. G., Poehlman, T. A., Uhlmann, E. L., and Banaji, M. R. (2009). Understanding and using the Implicit Association Test: III. Meta-analysis of predictive validity. *J. Pers. Soc. Psychol.* 97, 17–41. doi: 10.1037/a0015575
- Hofmann, W., Gschwendner, T., Nosek, B. A., and Schmitt, M. (2005). What moderates implicit-explicit consistency? *Eur. Rev. Soc. Psychol.* 16, 335–390. doi: 10.1080/10463280500443228
- Hoyt, W. T., and McCullough, M. E. (2005). “Issues in the multimodal measurement of forgiveness,” in *Handbook of Forgiveness*, ed E. L. Worthington (New York, NY: Routledge), 109–123.
- Johnston, M. E., Sherman, A., and Grusec, J. E. (2013). Predicting moral outrage and religiosity with an implicit measure of moral identity. *J. Res. Pers.* 47, 209–217. doi: 10.1016/j.jrp.2013.01.006
- Karremans, J. C., and Aarts, H. (2007). The role of automaticity in determining the inclination to forgive close others. *J. Exp. Soc. Psychol.* 43, 902–917. doi: 10.1016/j.jesp.2006.10.012
- Karremans, J. C., and Lange, P. A. M. V. (2008). Forgiveness in personal relationships: its malleability and powerful consequences. *E. Rev. Soc. Psychol.* 19, 202–241. doi: 10.1080/10463280802402609
- Lane, K. A., Banaji, M. R., Nosek, B. A., and Greenwald, A. G. (2007). “Understanding and using the Implicit Association Test: IV: What We Know (So Far) about the Method,” in *Implicit Measures of Attitudes*, eds B. Wittenbrink, N. Schwarz, B. Wittenbrink, and N. Schwarz (New York, NY: Guilford Press), 59–102.
- Lischetzke, T., Angelova, R., and Eid, M. (2011). Validating an indirect measure of clarity of feelings: evidence from laboratory and naturalistic settings. *Psychol. Assess.* 23, 447–455. doi: 10.1037/a0022211
- Lischetzke, T., Cuccodoro, G., Gauger, A., Todeschini, L., and Eid, M. (2005). Measuring affective clarity indirectly: individual differences in response latencies of state affect ratings. *Emotion* 5, 431–445. doi: 10.1037/1528-3542.5.4.431
- McCullough, M. E., Fincham, F. D., and Tsang, J.-A. (2003). Forgiveness, forbearance, and time: the temporal unfolding of transgression-related interpersonal motivations. *J. Pers. Soc. Psychol.* 84, 540–557. doi: 10.1037/0022-3514.84.3.540
- McCullough, M. E., and Hoyt, W. T. (2002). Transgression-related motivational dispositions: personality substrates of forgiveness and their links to the big five. *Pers. Soc. Psychol. Bull.* 28, 1556–1573. doi: 10.1177/014616702237583
- McCullough, M. E., Rachal, K. C., Sandage, S. J., Worthington, E. L. Jr., Brown, S. W., and Hight, T. L. (1998). Interpersonal forgiving in close relationships: II. Theoretical elaboration and measurement. *J. Pers. Soc. Psychol.* 75, 1586–1603. doi: 10.1037/0022-3514.75.6.1586
- McCullough, M. E., Root, L. M., and Cohen, A. D. (2006). Writing about the benefits of an interpersonal transgression facilitates forgiveness. *J. Consult. Clin. Psychol.* 74, 887–897. doi: 10.1037/0022-006X.74.5.887
- Moors, A., and De Houwer, J. (2006). Automaticity: a theoretical and conceptual analysis. *Psychol. Bull.* 132, 297–326. doi: 10.1037/0033-2909.132.2.297
- Musch, J., Brockhaus, R., and Bröder, A. (2002). Ein inventar zur erfassung von zwei faktoren sozialer erwünschtheit [An inventory for the assessment of two factors of social desirability]. *Diagnostica* 48, 121–129. doi: 10.1026//0012-1924.48.3.121
- Paulhus, D. L. (1984). Two-component models of socially desirable responding. *J. Pers. Soc. Psychol.* 46, 598–609. doi: 10.1037/0022-3514.46.3.598
- Paulhus, D. L. (1994). *Reference Manual for BIDR Version 6*. Vancouver, BC: University of British Columbia.
- Perugini, M., and Banse, R. (2007). Personality, implicit self-concept and automaticity. *Eur. J. Pers.* 21, 257–261. doi: 10.1002/per.637
- Röhner, J., Schröder-Abé, M., and Schütz, A. (2011). Exaggeration is harder than understatement, but practice makes perfect! Faking success in the IAT. *Exp. Psychol.* 58, 464–472. doi: 10.1027/1618-3169/a000114
- Rudolph, A., Schröder-Abé, M., Riketta, M., and Schütz, A. (2010). Easier when done than said! Implicit self-esteem predicts observed or spontaneous behavior, but not self-reported or controlled behavior. *J. Psychol.* 218, 12–19. doi: 10.1027/0044-3409/a000003
- Schmukle, S. C., and Egloff, B. (2006). Assessing anxiety with extrinsic Simon tasks. *Exp. Psychol.* 53, 149–160. doi: 10.1027/1618-3169.53.2.149
- Schnabel, K., and Asendorpf, J. B. (2010). “The self-concept: new insights from implicit measurement procedures,” in *Handbook of Implicit Social Cognition: Measurement, Theory, And Applications*, eds B. Gawronski and B. K. Payne (New York, NY: Guilford Press), 408–425.
- Schröder-Abé, M., Rudolph, A., Wiesner, A., and Schütz, A. (2007). Self-esteem discrepancies and defensive reactions to social feedback. *Int. J. Psychol.* 42, 174–183. doi: 10.1080/00207590601068134

- Steffens, M. C. (2004). Is the implicit association test immune to faking? *Exp. Psychol.* 51, 165–179. doi: 10.1027/1618-3169.51.3.165
- Stöber, J., Dette, D. E., and Musch, J. (2002). Comparing continuous and dichotomous scoring of the balanced inventory of desirable responding. *J. Pers. Assess.* 78, 370–389. doi: 10.1207/S15327752JPA7802_10
- Strack, F., and Deutsch, R. (2004). Reflective and impulsive determinants of social behavior. *Pers. Soc. Psychol. Rev.* 8, 220–247. doi: 10.1207/s15327957pspr0803_1
- Sutton, G. W., and Thomas, E. K. (2006). Restoring christian leaders: how conceptualizations of forgiveness and restoration used in empirical studies can influence practice and research. *Am. J. Pastor. Counsel.* 8, 27–42. doi: 10.1300/J062v08n02_02
- Witvliet, C. V., Ludwig, T. E., and Vander Laan, K. L. (2001). Granting forgiveness or harboring grudges: implications for emotion, physiology, and health. *Psychol. Sci.* 12, 117–123. doi: 10.1111/1467-9280.00320
- Worthington, E. L. Jr., and Wade, N. G. (1999). The psychology of unforgiveness and forgiveness and implications for clinical practice. *J. Soc. Clin. Psychol.* 18, 385–418.

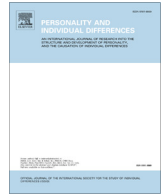
Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 04 April 2014; accepted: 23 June 2014; published online: 11 July 2014.

Citation: Fatfouta R, Schröder-Abé M and Merkl A (2014) Forgiving, fast and slow: validity of the implicit association test for predicting differential response latencies in a transgression-recall paradigm. *Front. Psychol.* 5:728. doi: 10.3389/fpsyg.2014.00728

This article was submitted to *Personality and Social Psychology*, a section of the journal *Frontiers in Psychology*.

Copyright © 2014 Fatfouta, Schröder-Abé and Merkl. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



How forgiveness affects processing time: Mediation by rumination about the transgression



Ramzi Fatfouta *

Cluster "Languages of Emotion", Freie Universität Berlin, Germany
Department of Education and Psychology, Freie Universität Berlin, Germany

ARTICLE INFO

Article history:

Received 15 December 2014
Received in revised form 4 March 2015
Accepted 7 March 2015

Keywords:

Forgiveness
Rumination
Response time
Interpersonal relationships
Social cognition

ABSTRACT

Previous work has demonstrated that individuals high (vs. low) in forgiveness are faster to rate their current thoughts and feelings toward their transgressor (i.e., their state forgiveness), but the underlying mechanism is still unclear. The present study examined whether individual differences in rumination about the transgression would mediate the association between forgiveness and response time. Participants ($N = 767$) completed measures of trait forgiveness, rumination about the transgression, and state forgiveness (while response time was unobtrusively recorded). Trait forgiveness was significantly negatively associated with response time and this effect was mediated by lower rumination about the transgression. Results support the use of response time as an unobtrusive measure of forgiveness-related processes and further clarify the role of rumination about the transgression in the forgiveness–response time link.

© 2015 Elsevier Ltd. All rights reserved.

1. Introduction

How can one tell if someone has forgiven an offense? One way to answer this question is to directly ask the individual for such information. Indeed, most forgiveness research has almost exclusively relied on self-reports, in which individuals (1) recall an interpersonal transgression and then (2) rate statements such as "I have given up my hurt and resentment" (McCullough, Root, & Cohen, 2006, p. 897). However, individuals' reports of their forgiveness can be affected by introspective limits and response factors (Greenwald et al., 2002). That is, individuals might be unable to accurately indicate their forgiveness, or they might present themselves favorably to others. In light of these limitations, researchers have called for a multimethod research strategy (McCullough & Witvliet, 2002). However, the use of alternative measures of forgiveness is still in its infancy.

To circumvent the problems intrinsic to self-report, researchers have begun to employ behavioral observations of forgiveness (Carlisle et al., 2012). For example, a recent study found that high-forgiving individuals, as compared to low-forgiving individuals, generated more positive characteristics that describe their transgressor in a free-listing task (Dorn, Hook, Davis, Van

Tongeren, & Worthington, 2014, Study 2). Another study demonstrated that individuals who had forgiven their transgressor used more plural pronouns (e.g., we, us) in an unrelated task (Karremans, Van Lange, & Holland, 2005, Study 1). Although promising, it is unclear whether these indices truly represent forgiveness-related behavior—or whether such indices are an epiphenomenal manifestation of some other factor (e.g., demand characteristics).

Another line of research employed a more subtle measure of response time that seeks to capture the processes underlying forgiveness–response decisions (Bast & Barnes-Holmes, 2014). The benefit of this methodology is that it does not require self-insight necessary for self-reports; instead, the relevant characteristic is targeted via performance (Robinson & Neighbors, 2006). For example, Fatfouta, Schröder-Abé, and Merkl (2014) had participants to relive a real-life transgression and then rate their current thoughts and feelings toward the offender (i.e., their state forgiveness). The latency for responding served as an unobtrusive indicator of forgiveness. The rationale behind this approach is that individuals should need less time to rate their state forgiveness when they have already worked through the pain (i.e., when they can easily access internal cues about their forgiveness) than when they have not yet achieved forgiveness (i.e., when they find internal forgiveness cues to be rather difficult to access). Indeed, the authors showed that high-forgiving individuals, as compared to low-forgiving individuals, exhibited faster responses. The authors

* Address: Department of Education and Psychology, Freie Universität Berlin, Habelschwerdter Allee 45, 14195 Berlin, Germany. Tel.: +49 (0)30 838 57 857.

E-mail address: ramzi.fatfouta@fu-berlin.de

further found that response time was unrelated to socially desirable responding. At present, however, there is no evidence as to how forgiveness affects response time. Thus, in the present research, the aim was to replicate and extend these initial findings by exploring the underlying mechanism that mediates this association.

Although several socio-cognitive variables have been shown to impact forgiveness, a critical role has been attributed to rumination (Ysseldyk, Matheson, & Anisman, 2007). Rumination is characterized by a “passive and repetitive focus on the negative and damaging features of a stressful transaction” (Skinner, Edge, Altman, & Sherwood, 2003, p. 242). Cross-sectional studies have demonstrated that forgiveness is negatively related to rumination (Berry, Worthington, Parrott, O’Connor, & Wade, 2001; Thompson et al., 2005). Specifically, individuals who report lower levels of forgiveness ruminate more after an interpersonal transgression (Kachadourian, Fincham, & Davila, 2005). These findings are corroborated by longitudinal data, demonstrating that changes in forgiveness are negatively correlated with changes in rumination (McCullough, Bellah, Kilpatrick, & Johnson, 2001; McCullough, Bono, & Root, 2007).

By mulling an aversive event such as an interpersonal transgression over in their minds, individuals often attempt to understand why it upsets them (Watkins & Baracaia, 2001). Such a thought process involves evaluative and analytical questions (e.g., “Why did he/she hurt me?”) that revolve around the (negative) causes and consequences of that event (Watkins, 2008; Watkins & Baracaia, 2002). Furthermore, when individuals ruminate, they tend to process information more thoroughly, carefully, and slowly (see Andrews & Thomson, 2009). Hence, when prompted to (1) call to mind a transgression and then (2) rate their state of forgiveness, individuals who ponder on their transgression more persistently than others do, should require more time to arrive at a response (and, hence, exhibit longer latencies). This is consistent with evidence that a deliberative mode of information processing takes time, as indexed by a slowing in response time (Baxter & Hinson, 2001; Fazio, 1990a).

To the extent that response times represent “overt behavioral indices of a covert mental process” (Holden, Woermke, & Fekken, 1993, p. 2), they may provide an appealing approach to the study of forgiveness. Moving beyond previous work, the present research sought to scrutinize the psychological mechanism underlying the association between forgiveness and the time with which individuals generated responses to state forgiveness ratings following the transgression recall. If individuals high (vs. low) in forgiveness are faster to rate their state forgiveness, it is perhaps their reduced amount of ruminative tendencies that explains this association. To test this hypothesis, a mediation model is proposed in which individuals high (vs. low) in forgiveness would display lower ruminative tendencies and this, in turn, might mediate the association between forgiveness and response time.

2. Methods

2.1. Participants

Seven hundred and sixty-seven individuals (76.6% women) of the general population, living in Germany (69.9%), Austria (28.3%), and Switzerland (1.8%), were recruited to complete an online study. Participants ranged in age from 18 to 60 ($M_{\text{age}} = 24.02$, $SD = 5.19$), and 23.3% held a college or university degree. All participants were German speaking. As an incentive, participants were offered entry into a raffle for a 25€ gift card.

2.2. Procedure

After consenting to participate, individuals completed the measures described below. Each item appeared one at a time on the computer screen. To proceed with the next item, participants had to click on the respective response category. Participants first answered questions to obtain a measure of their baseline speed (see Baseline speed) and then completed a measure of trait forgiveness. Next, participants were instructed to bring to mind a real-life situation in which someone had hurt them and to “indicate your current thoughts and feelings about the person who hurt you; that is, we want to know how you feel about that person right now.” Participants then completed measures of rumination about the transgression and state forgiveness. Response time was recorded as the time (in milliseconds) between item presentation and response selection. Participants were not informed that response time was being assessed so their behavior reflects their natural speed of responding (Lischetzke, Angelova, & Eid, 2011; Lischetzke, Cuccodoro, Gauger, Todeschini, & Eid, 2005).

3. Measures

3.1. Baseline speed

Given that response time was used as an individual-difference variable, it was particularly important to control for participants’ baseline speed of responding (Fazio, 1990b). Baseline speed was assessed by means of ten easy-knowledge questions (e.g., “What type of celestial body is the earth?”) (1) White giant; (2) Planet; (3) Asteroid; (4) Moon; Fatfouta et al., 2014; Lischetzke et al., 2005). The correct answer was always obvious, so that response time tapped only individual differences in reading ability and motor skills (e.g., muscle speed). Median item accuracy was high (100%), thus confirming this rationale. Following the procedures outlined by Ratcliff (1993), fixed-cutoff values were used to minimize the impact of response-time outliers; very fast (≤ 1000 ms) and very slow ($\geq 10,000$ ms) latencies were excluded on this base (4.6% of response times). The first question was regarded as a training item to familiarize participants with the response format and was therefore removed from the analysis (for a similar procedure, see Lischetzke et al., 2005, 2011). As a robust estimator of central tendency, the median rather than the mean of the remaining nine response time scores was calculated (Whelan, 2008).

3.2. Trait forgiveness

Trait forgiveness was measured with the four-item Tendency to Forgive Scale (TTF; Brown, 2003; e.g. “When people wrong me, my approach is just to forgive and forget”). Participants rated each item on a 7-point Likert-type scale from 1 (*strongly disagree*) to 7 (*strongly agree*).

3.3. Rumination about the transgression

Rumination about the transgression was assessed using an adapted version of the post-event rumination scale (McCullough et al., 2007). Participants responded to six items (e.g., “Images of the offense kept coming back to me”) on a 7-point Likert-type scale from 1 (*not at all true of me*) to 7 (*extremely true of me*).

3.4. State forgiveness

State forgiveness was assessed with the Transgression-Related Interpersonal Motivations Inventory (TRIM; McCullough et al., 1998, 2006). The questionnaire assesses individuals’ motivational

changes toward a specific transgressor and is divided into three subscales: revenge (5 items; e.g., “I want to see him/her hurt and miserable”), avoidance (7 items; e.g., “I don’t trust him/her”), and benevolence (6 items; e.g., “I have given up my hurt and resentment”). Items were rated on a 7-point Likert-type scale from 1 (*strongly disagree*) to 7 (*strongly agree*).

3.5. Response time

For each state forgiveness item, response time was unobtrusively recorded. Applying the same cutoff as described before removed 3.2% of response times. Again, the first item was treated as a training item and its response time was removed from the analysis. To make sure that the remaining 17 response times can be described by a single dimension, a principal component analysis with varimax rotation was performed. The scree plot revealed a strong first component (first seven eigenvalues: 4.5, 1.39, 1.19, 0.95, 0.91, 0.87, 0.81). Thus, the response time scores were combined by calculating the median for each person. Finally, baseline speed was partialled out from these response time scores.¹ The resulting residual (i.e., baseline-corrected) response time scores were then used in all subsequent analyses. To note, values greater (smaller) than zero indicate that a person is slower (faster) than predicted by his or her own baseline speed.

4. Results

4.1. Descriptive statistics, reliabilities, and correlations

The types of relationship targets who had committed transgressions against participants were diverse. Most transgressions were committed by romantic partners (47.5%), other family members or relatives (28.3%), and close friends (11.3%). A smaller proportion of participants reported on transgressions by parents (10.3%), coworker, acquaintance or neighbor (0.7%), and “others” (2.0%).

Table 1 details descriptive statistics and internal consistency estimates (Cronbach’s α) for all measures. Response time demonstrated adequate reliability (.85), similar in size to the self-report measures (alphas ranged from .69 to .92).

Intercorrelations are presented in Table 2. Post-hoc power estimates ranged from 0.95 to 1.0 ($N = 767$, $\alpha = .05$ [two-tailed]; Faul, Erdfelder, Buchner, & Lang, 2009), indicating excellent statistical power. Consistent with Fatfouta et al. (2014), individuals high (vs. low) in trait forgiveness were significantly faster to rate their state forgiveness. In addition, rumination about the transgression was positively correlated with response time. That is, those individuals who ruminated more about the transgression were slower to rate their state forgiveness (see Fig. 1 for a scatterplot of these associations).²

4.2. Meditation analysis

To test whether the effect of trait forgiveness (the independent variable) on response time (the dependent variable) was mediated by rumination about the transgression (the potential mediator), a path model was built. Following Hayes (2013), bootstrapping

Table 1

Means, standard deviations, Cronbach’s alphas (α), and actual ranges of all study variables.

Variable	M	SD	α	Min	Max
Baseline speed (ms) ^a	3827.86	942.20	.86	1313.38	7173.80
Trait forgiveness	3.81	1.14	.69	1.00	7.00
Rumination about the transgression	3.12	1.57	.92	1.00	7.00
TRIM-Revenge	1.39	0.64	.78	1.00	4.80
TRIM-Avoidance	2.01	1.06	.85	1.00	6.43
TRIM-Benevolence	5.58	1.16	.82	1.00	7.00
Response time (ms) ^a	3575.47	879.15	.85	1726.07	6825.75

Note: $N = 767$. TRIM, transgression-related interpersonal motivations. Potential range of self-report scores was 1–7.

^a For the sake of clarity, these values represent mean latencies.

Table 2

Intercorrelations among study variables.

Variable	1	2	3	4	5	6
1. Trait forgiveness	–					
2. Rumination about the transgression	–.26***	–				
3. TRIM-Revenge	–.22***	.45***	–			
4. TRIM-Avoidance	–.18***	.56***	.54***	–		
5. TRIM-Benevolence	.20***	–.29***	–.38***	–.49***	–	
6. Residual response time ^a	–.13***	.25***	.19***	.30**	–.13***	–

Note: $N = 767$. TRIM, transgression-related interpersonal motivations. To note, intercorrelations with the mean latencies were as follows: –.12, .20, .13, .24, and –.12 (all p ’s < .001, two-tailed) for variables 1–5, respectively.

^a Baseline speed partialled out.

*** $p < .001$, two-tailed.

analyses were performed using the PROCESS macro for SPSS; 5000 bootstrap samples with 95% bias-corrected confidence intervals (CIs) were applied. CIs that do not contain 0 indicate the presence of a mediation effect. Furthermore, kappa-squared (i.e., κ^2) is presented as a standardized effect size measure. κ^2 denotes the proportion of the maximum possible indirect effect (Preacher & Kelley, 2011), and can be interpreted as small (.01), medium (.09), or large (.25). Fig. 2 depicts the results of this analysis.

The results revealed that trait forgiveness was related to lower rumination about the transgression, $b_a \text{ path} = -0.36$, 95% CI (–0.46, –0.27). Results also demonstrated that rumination was related to slower response time, $b_b \text{ path} = 0.15$, 95% CI (0.11, 0.20). The indirect effect ab of trait forgiveness on response time through rumination was significant, 95% CI (–0.08, –0.04). The proportion of the maximum indirect effect observed was $\kappa^2 = .06$, 95% CI (0.04, 0.09), suggesting a small to medium effect size. Finally, the direct effect of trait forgiveness on response time was no longer significant, $b_c \text{ path} = -0.06$, 95% CI (–0.11, 0.01). Together, this means that, as hypothesized, rumination about the transgression mediated the association between trait forgiveness and response time.

5. Discussion

Individuals who have not yet forgiven an offense might be unwilling to disclose their thoughts and feelings to others even if prompted. To bypass this problem, the present study employed an unobtrusive measure of forgiveness, which is based on the time with which individuals generate responses to state forgiveness ratings. Furthermore, to shed light on the potential mechanism accounting for the link between forgiveness and response time, rumination about the transgression was proposed and tested as a mediator variable.

¹ In particular, a linear regression predicting median response time to state forgiveness items from median baseline speed was computed and the standardized residual saved.

² It was also examined whether a curvilinear relationship might exist between trait forgiveness as well as rumination about the transgression and response time. However, the data showed no support for the existence of a curvilinear relationship (p ’s > .277). In addition, it was tested whether the observed relationships would hold if all three TRIM subscales were controlled. Indeed, the partial correlations were still significant ($p = .041$ and $p < .001$, for trait forgiveness and rumination about the transgression, respectively).

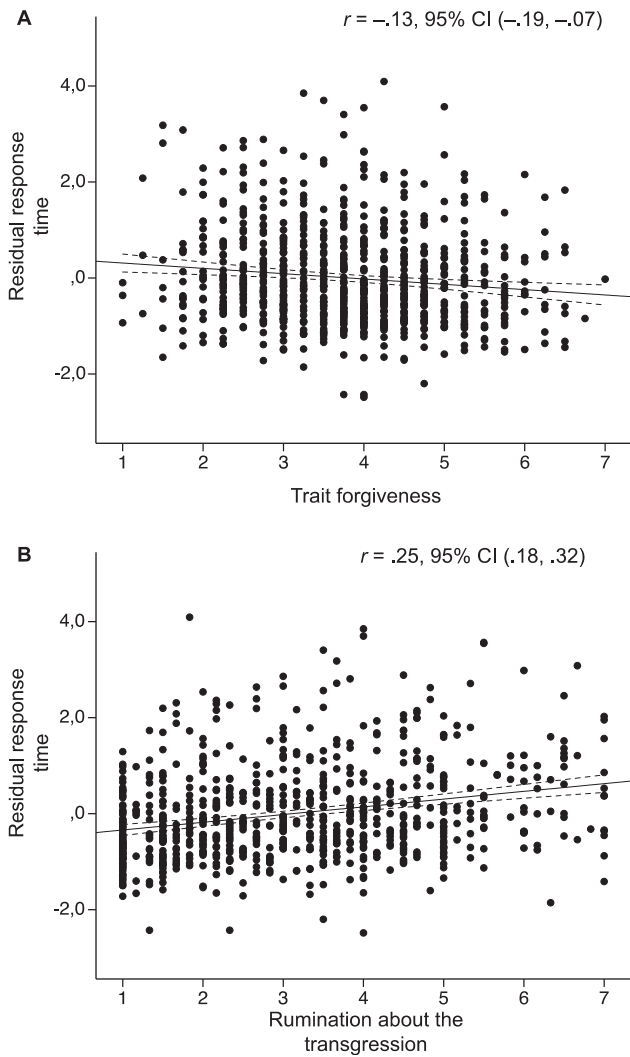


Fig. 1. Scatterplot of the relationship between trait forgiveness (panel a) as well as rumination about the transgression (panel b) and residual (i.e., baseline-corrected) response time. Dashed lines indicate the 95% confidence intervals (CIs) for the regression line.

Drawing on a large and heterogeneous sample, the present study found an inverse association between forgiveness and response time, thereby replicating the initial findings by Fatfouta et al. (2014). High-forgiving individuals, as compared to low-forgiving individuals, were faster to describe and rate their current thoughts and feelings toward their transgressor (i.e., state forgiveness). The obtained correlation was $r = -.13$, comparable with related research on implicit personality measures (Lischetzke et al., 2011). To note, when two measures pertain to different modalities (self-report vs. response time), small to moderate effects are to be expected (Lischetzke et al., 2005; Robinson & Neighbors, 2006). Importantly, because residual (i.e., baseline-corrected) response time was used, the present relationship holds after adjusting for individuals' general speed of responding.

The mediation analysis revealed a significant indirect effect of trait forgiveness on response time via rumination about the transgression. That is, as expected, individuals with high (vs. low) levels of forgiveness tended to ruminate less about the transgression, which in turn promoted faster responses. Thus, the present study moves beyond prior research by providing the first empirical support for a potential mechanism underlying the forgiveness–response time link. From this perspective, response time to state

forgiveness items may be better characterized as reflecting an individuals' diminished ruminative tendency rather than an increased accessibility of forgiveness. The positive association between rumination about the transgression and response time was anticipated based on past research demonstrating that ruminative tendencies are associated with careful information processing and, hence, longer latencies (Andrews & Thomson, 2009). Furthermore, less ruminative thoughts about the transgression may imply a better ability to inhibit unwanted memories related to the interpersonal transgression – an interpretation consistent with research showing that rumination is associated with deficits in inhibitory control (Denson, Pedersen, Friese, Hahm, & Roberts, 2011; Joormann, Levens, & Gotlib, 2011).

The results of the present research are in line with recent work on the role of implicit processes in forgiveness (Karremans & Aarts, 2007) and make a compelling case for the value of considering both implicit and explicit aspects of human behavior. Through the wider use of computerized testing, response times can be collected economically and provide a valuable, albeit underutilized, source of information. Hence, to allow for a multifaceted examination of forgiveness, the administration of multiple assessment formats is strongly recommended for future research. Moreover, future studies should scrutinize more thoroughly to what extent latencies of responses to forgiveness items reflect stable (trait) versus modifiable (state) aspects.

Before closing, some methodological considerations need to be discussed. First, though the large sample size allowed for detecting the observed effects with considerable precision (Schönbrodt & Perugini, 2013), small to medium effect sizes were obtained and thus warrant caution when interpreting the results. Second, the design of this study was cross-sectional. Although it is unlikely that response time influenced individuals' trait forgiveness, causal inferences cannot be drawn. Longitudinal and/or experimental research would provide a critical next step in examining this link in further detail. Third, this study focused on one specific real-life transgression. Using experience-sampling methodology, future studies might examine forgiveness-related processes in naturalistic settings (i.e., in individuals' daily lives). Demonstrating that individual differences in forgiveness are related to field-based response times to state forgiveness ratings and examining the

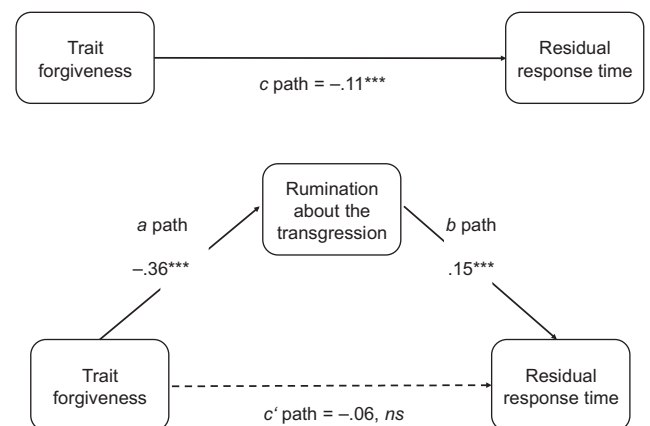


Fig. 2. Mediation model testing rumination about the transgression as a mediator of the relationship between trait forgiveness and response time. Path c is the total effect of trait forgiveness on response time (sum of direct and indirect effects: $c = c' + ab$); c' is the direct effect of trait forgiveness on response time; a is the direct effect of trait forgiveness on rumination; b is the direct effect of rumination on response time. When rumination was included in the model the direct effect of trait forgiveness on response time (dashed line) was no longer significant. Coefficients represent unstandardized regression coefficients. *** $p < .001$, two-tailed, ns = not significant.

mediating role of rumination therein would be an interesting avenue for future research. Furthermore, alternate measures of rumination about the transgression, beyond self-report, would constitute a useful addition to the multimethod approach in forgiveness research. Finally, future work might consider investigating potential moderators of the association between forgiveness and response time. For example, individuals experiencing anger toward their transgressor may yield faster responses times as a result of heightened arousal associated with this state (Bresin, Hilmert, Wilkowski, & Robinson, 2012). Alternatively, individuals who feel ambivalent toward their transgressor (e.g., if conflicted about whether or not to forgive him/her) might generate slowed responses to state forgiveness ratings (Kachadourian et al., 2005) – a notion consistent with evidence that ambivalence slows down response time (Van Harreveld, Van Der Pligt, De Vries, Wenneker, & Verhug, 2004). In future research, it might also be worthwhile to investigate the present relations with respect to unforgiving responses (e.g., harboring a grudge).³ Given evidence that individuals holding a grudge tend to rehearse memories of the hurtful event (Baumeister, Exline, & Sommer, 1998), one testable prediction is that these individuals should need comparatively more time to contemplate and report whether their transgressor deserves forgiveness.

To conclude, the present study supported the hypothesis that rumination about the transgression mediates the link between forgiveness and response time. A better understanding of the processes underlying forgiveness–response decisions provides a unique window into the study of forgiveness and, hence, awaits future research.

Acknowledgments

This research was financially supported by the Cluster “Languages of Emotion”, Freie Universität Berlin. I thank Michela Schröder-Abé, Angela Merkl, and Hauke R. Heekeren who influenced me in different ways. I am also indebted to the editors and two anonymous reviewers for their constructive comments during the review process.

References

- Andrews, P. W., & Thomson, J. A. Jr., (2009). The bright side of being blue: Depression as an adaptation for analyzing complex problems. *Psychological Review*, 116(3), 620–654. <http://dx.doi.org/10.1037/a0016242>.
- Bast, D., & Barnes-Holmes, D. (2014). A first test of the implicit relational assessment procedure as a measure of forgiveness of self and others. *Psychological Record*, 64(2), 253–260. <http://dx.doi.org/10.1007/s40732-014-0022-2>.
- Baumeister, R. F., Exline, J. J., & Sommer, K. L. (1998). The victim role, grudge theory, and two dimensions of forgiveness. In E. L. Worthington, Jr (Ed.), *Dimensions of forgiveness: Psychological research and theological perspectives* (pp. 79–104). Philadelphia, PA: Templeton Foundation Press.
- Baxter, B. W., & Hinson, R. E. (2001). Is smoking automatic? Demands of smoking behavior on attentional resources. *Journal of Abnormal Psychology*, 110(1), 59.
- Berry, J. W., Worthington, E. L., Jr., Parrott, L., III, O'Connor, L. E., & Wade, N. G. (2001). Dispositional forgiveness: Development and construct validity of the transgression narrative test of forgiveness (TNTF). *Personality and Social Psychology Bulletin*, 27(10), 1277–1290. <http://dx.doi.org/10.1177/01461672012710004>.
- Bresin, K., Hilmert, C. J., Wilkowski, B. M., & Robinson, M. D. (2012). Response speed as an individual difference: Its role in moderating the agreeableness–anger relationship. *Journal of Research in Personality*, 46(1), 79–86.
- Brown, R. P. (2003). Measuring individual differences in the tendency to forgive: Construct validity and links with depression. *Personality and Social Psychology Bulletin*, 29(6), 759–771. <http://dx.doi.org/10.1177/0146167203029006008>.
- Carlisle, R. D., Tsang, J.-A., Ahmad, N. Y., Worthington, E. L., Witvliet, C. V., & Wade, N. (2012). Do actions speak louder than words? Differential effects of apology and restitution on behavioral and self-report measures of forgiveness. *Journal of Positive Psychology*, 7(4), 294–305. <http://dx.doi.org/10.1080/17439760.2012.690444>.
- Denson, T. F., Pedersen, W. C., Friese, M., Hahm, A., & Roberts, L. (2011). Understanding impulsive aggression: Angry rumination and reduced self-control capacity are mechanisms underlying the provocation–aggression relationship (Impulsive Aggression verstehen: Zorniges Grübeln und verringerte Selbstkontrolle als zugrunde liegende Mechanismen der Beziehung zwischen Provokation und Aggression). *Personality and Social Psychology Bulletin*, 37(6), 850–862. <http://dx.doi.org/10.1177/0146167211401420>.
- Dorn, K., Hook, J. N., Davis, D. E., Van Tongeren, D. R., & Worthington, E. L. Jr., (2014). Behavioral methods of assessing forgiveness. *The Journal of Positive Psychology*, 9(1), 75–80. <http://dx.doi.org/10.1080/17439760.2013.844267>.
- Fatfouta, R., Schröder-Abé, M., & Merkl, A. (2014). Forgiving, fast and slow: Validity of the implicit association test for predicting differential response latencies in a transgression-recall paradigm. *Frontiers in Psychology*, 5, 1–14. <http://dx.doi.org/10.3389/fpsyg.2014.00728>.
- Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, 41(4), 1149–1160. <http://dx.doi.org/10.3758/brm.41.4.1149>.
- Fazio, R. (1990a). Multiple processes by which attitudes guide behavior: The MODE model as an integrative framework. *Advances in Experimental Social Psychology*, 23, 75–109.
- Fazio, R. (1990b). A practical guide to the use of response latency in social psychological research. In C. Hendrick & M. S. Clark (Eds.), *Research methods in personality and social psychology* (pp. 74–97). Thousand Oaks, CA US: Sage Publications Inc.
- 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(1), 3–25. <http://dx.doi.org/10.1037/0033-295x.109.1.3>.
- Hayes, A. F. (2013). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. New York, NY, US: Guilford Press.
- Holden, R. R., Woermke, C., & Fekken, G. C. (1993). Enhancing the construct validity of differential response latencies for personality test items. *Canadian Journal of Behavioural Science (Revue canadienne des sciences du comportement)*, 25(1), 1–11. <http://dx.doi.org/10.1037/h0078785>.
- Joermann, J., Levens, S. M., & Gotlib, I. H. (2011). Sticky thoughts: Depression and rumination are associated with difficulties manipulating emotional material in working memory. *Psychological Science*, 22(8), 979–983. <http://dx.doi.org/10.1177/0956797611415539> (Sage Publications Inc.).
- Kachadourian, L. K., Fincham, F., & Davila, J. (2005). Attitudinal ambivalence, rumination, and forgiveness of partner transgressions in marriage. *Personality and Social Psychology Bulletin*, 31(3), 334–342. <http://dx.doi.org/10.1177/0146167204271595>.
- Karremans, J. C., & Aarts, H. (2007). The role of automaticity in determining the inclination to forgive close others. *Journal of Experimental Social Psychology*, 43(6), 902–917. <http://dx.doi.org/10.1016/j.jesp.2006.10.012>.
- Karremans, J. C., Van Lange, P. A. M., & Holland, R. W. (2005). Forgiveness and its associations with prosocial thinking, feeling, and doing beyond the relationship with the offender. *Personality and Social Psychology Bulletin*, 31(10), 1315–1326. <http://dx.doi.org/10.1177/0146167205274892>.
- Lischetzke, T., Angelova, R., & Eid, M. (2011). Validating an indirect measure of clarity of feelings: Evidence from laboratory and naturalistic settings. *Psychological Assessment*, 23(2), 447–455. <http://dx.doi.org/10.1037/a0022211>.
- Lischetzke, T., Cuccodoro, G., Gauger, A., Todeschini, L., & Eid, M. (2005). Measuring affective clarity indirectly: Individual differences in response latencies of state affect ratings. *Emotion*, 5(4), 431–445. <http://dx.doi.org/10.1037/1528-3542.5.4.431>.
- McCullough, M. E., Bellah, C. G., Kilpatrick, S. D., & Johnson, J. L. (2001). Vengefulness: Relationships with forgiveness, rumination, well-being, and the big five. *Personality and Social Psychology Bulletin*, 27(5), 601–610. <http://dx.doi.org/10.1177/0146167201275008>.
- McCullough, M. E., Bono, G., & Root, L. M. (2007). Rumination, emotion, and forgiveness: Three longitudinal studies. *Journal of Personality and Social Psychology*, 92(3), 490–505. <http://dx.doi.org/10.1037/0022-3514.92.3.490>.
- McCullough, M. E., Rachal, K. C., Sandage, S. J., Worthington, E. L., Jr., Brown, S. W., & Hight, T. L. (1998). Interpersonal forgiving in close relationships: II. Theoretical elaboration and measurement. *Journal of Personality and Social Psychology*, 75(6), 1586–1603. <http://dx.doi.org/10.1037/0022-3514.75.6.1586>.
- McCullough, M. E., Root, L. M., & Cohen, A. D. (2006). Writing about the benefits of an interpersonal transgression facilitates forgiveness. *Journal of Consulting and Clinical Psychology*, 74(5), 887–897. <http://dx.doi.org/10.1037/0022-006x.74.5.887>.
- McCullough, M. E., & Witvliet, C. V. (2002). The psychology of forgiveness. In C. R. Snyder & S. J. Lopez (Eds.), *Handbook of positive psychology* (pp. 446–458). New York, NY US: Oxford University Press.
- Preacher, K. J., & Kelley, K. (2011). Effect size measures for mediation models: Quantitative strategies for communicating indirect effects. *Psychological Methods*, 16(2), 93–115. <http://dx.doi.org/10.1037/a0022658>.
- Ratcliff, R. (1993). Methods for dealing with reaction time outliers. *Psychological Bulletin*, 114(3), 510–532. <http://dx.doi.org/10.1037/0033-2909.114.3.510>.
- Robinson, M. D., & Neighbors, C. (2006). Catching the mind in action: Implicit methods in personality research and assessment. In M. Eid & E. Diener (Eds.), *Handbook of multimethod measurement in psychology* (pp. 115–125). Washington, DC US: American Psychological Association.
- Schönbrodt, F. D., & Perugini, M. (2013). At what sample size do correlations stabilize? *Journal of Research in Personality*, 47(5), 609–612. <http://dx.doi.org/10.1016/j.jrp.2013.05.009>.

³ I thank an anonymous reviewer for suggesting this point.

- Skinner, E. A., Edge, K., Altman, J., & Sherwood, H. (2003). Searching for the structure of coping: A review and critique of category systems for classifying ways of coping. *Psychological Bulletin*, *129*(2), 216–269. <http://dx.doi.org/10.1037/0033-2909.129.2.216>.
- Thompson, L. Y., Snyder, C. R., Hoffman, L., Michael, S. T., Rasmussen, H. N., Billings, L. S., et al. (2005). Dispositional forgiveness of self, others, and situations. *Journal of Personality*, *73*(2), 313–359. <http://dx.doi.org/10.1111/j.1467-6494.2005.00311.x>.
- Van Harreveld, F., Van Der Pligt, J., De Vries, N. K., Wenneker, C., & Verhue, D. (2004). Ambivalence and information integration in attitudinal judgment. *British Journal of Social Psychology*, *43*(3), 431–447.
- Watkins, E. (2008). Constructive and unconstructive repetitive thought. *Psychological Bulletin*, *134*(2), 163–206. <http://dx.doi.org/10.1037/0033-2909.134.2.163>.
- Watkins, E., & Baracaia, S. (2001). Why do people ruminate in dysphoric moods? *Personality and Individual Differences*, *30*(5), 723–734.
- Watkins, E., & Baracaia, S. (2002). Rumination and social problem-solving in depression. *Behaviour Research and Therapy*, *40*(10), 1179–1189.
- Whelan, R. (2008). Effective analysis of reaction time data. *Psychological Record*, *58*(3), 475–482.
- Ysseldyk, R., Matheson, K., & Anisman, H. (2007). Rumination: Bridging a gap between forgivingness, vengefulness, and psychological health. *Personality and Individual Differences*, *42*(8), 1573–1584. <http://dx.doi.org/10.1016/j.paid.2006.10.032>.



Contents lists available at ScienceDirect

Personality and Individual Differences

journal homepage: www.elsevier.com/locate/paid

Short Communication

Narcissism and lack of interpersonal forgiveness: The mediating role of state anger, state rumination, and state empathy[☆]



Ramzi Fatfouta^{a,b,*}, Tanja M. Gerlach^c, Michela Schröder-Abé^{d,1}, Angela Merkl^{b,e,1}

^a Department of Education and Psychology, Freie Universität Berlin, Germany

^b Cluster of Excellence "Languages of Emotion", Freie Universität Berlin, Germany

^c Institute of Psychology, Georg August University Göttingen, Germany

^d Department of Psychology, Technische Universität Darmstadt, Germany

^e Department of Neurology with Experimental Neurology, Charité Universitätsmedizin, Berlin, Germany

ARTICLE INFO

Article history:

Received 7 August 2014

Received in revised form 20 October 2014

Accepted 28 October 2014

Keywords:

Forgiveness

Revenge

Narcissism

Social relationships

Interpersonal transgressions

ABSTRACT

The present study investigated the unique contributions of two distinct dimensions of narcissism – admiration and rivalry – to two facets of unforgiveness: revenge and avoidance. In addition, we examined whether state anger, state rumination, and state empathy mediate this relationship. Using a large sample ($N = 1040$), we found that admiration was negatively related to revenge and avoidance via higher empathy for the transgressor. By contrast, rivalry was positively related to revenge and avoidance via greater anger and rumination and less empathy. Findings suggest that the mechanisms through which narcissism and lack of forgiveness are associated are better understood if we disentangle admiration and rivalry and consider both cognitive and affective antecedents of narcissists' unforgiving motivations.

© 2014 Elsevier Ltd. All rights reserved.

1. Introduction

A large body of research on forgiveness has demonstrated that individual difference variables affect one's response to transgressions (for an overview, see Fehr, Gelfand, & Nag, 2010). In particular, narcissism has been argued to have an "inhibitory effect on the forgiveness process" (Emmons, 2000, p. 164). Narcissism is characterized by feelings of entitlement, self-enhancement, exploitative interpersonal behavior, and lack of empathy for others (Rhodewalt & Peterson, 2009). Narcissists² tend to behave aggressively when faced with social rejection (Twenge & Campbell, 2003). Indeed, it has been demonstrated that narcissistic individuals have a lower tendency to react to interpersonal offenses with forgiveness (Eaton, Struthers, & Santelli, 2006). Specifically, narcissistic entitlement (i.e., a sense of deserving special treatment) has been

linked to reduced forgiveness (Exline, Baumeister, Bushman, Campbell, & Finkel, 2004).

Studies examining the narcissism–forgiveness link almost exclusively relied on one questionnaire, the Narcissistic Personality Inventory (NPI; Raskin & Hall, 1979). Although widely used, it has been criticized for conflating both adaptive and maladaptive dimensions of narcissism into one composite. Moreover, it has been suggested that narcissism will be better understood if we distinguish these dimensions (Ackerman et al., 2011). Specifically, as condensed in the Narcissistic Admiration and Rivalry Concept (NARC; Back et al., 2013), narcissism can be conceptualized as the interplay of two distinct social strategies: First, the propensity for assertive self-enhancement by means of self-promotion (*admiration*) and, second, the propensity for antagonistic self-protection by means of self-defense (*rivalry*). Empirically, admiration is related to adjustment indicators (e.g., self-assuredness, problem-focused reactions to transgressions), whereas rivalry is related to maladjustment (e.g., entitlement, more hostile reactions). Despite the well-established link between narcissism and reduced forgiveness, little is known about the facet-specific contributions of admiration and rivalry to lack of forgiveness. Consistent with the NARC, we reasoned that unforgiveness should be related to rivalry but less so to admiration.

Furthermore, the underlying mechanisms that mediate the association between narcissism and unforgiveness are largely

[☆] This research was financially supported by the Cluster of Excellence "Languages of Emotion" at Freie Universität Berlin.

* Corresponding author at: Department of Education and Psychology, Freie Universität Berlin, Habelschwerdter Allee 45, 14195 Berlin, Germany. Tel.: +49 (0)30 838 57 857.

E-mail address: ramzi.fatfouta@fu-berlin.de (R. Fatfouta).

¹ Shared senior authorship.

² In the present study, narcissism is considered as a continuous dimension of normal (vs. pathological) personality. Consequently, the term *narcissists* is used for individuals who score high on measures of normal narcissism.

unexplored. Here, we argue that three socio-cognitive variables – state anger, state rumination, and state empathy – might prove promising. Anger, rumination, and reduced empathy have been prominent variables in predicting a lack of forgiveness (Fehr et al., 2010). Narcissists report higher anger in the face of transgressions, such as an interpersonal rejection (Twenge & Campbell, 2003). Further, Krizan and Johar (in press, Study 3) found that narcissistic entitlement is associated with rumination. Finally, narcissism has been shown to predict low empathy (Watson & Morris, 1991). Thus, we hypothesized a multiple mediator model in which these socio-cognitive variables (state anger, state rumination, and state empathy) are key factors accounting for lack of forgiveness in narcissistic rivalry and admiration.

2. Method

2.1. Participants and procedure

A total of 1040 individuals ($M_{\text{age}} = 22.8$, $SD = 4.9$, range: 18–50; 81.1% female) from Germany (74.6%), Austria (24.4%), and Switzerland (1.0%) were recruited via advertising on a social networking site to complete an online study. As an incentive, they were given the option to enter a lottery for one of four 25 € gift cards. All participants provided informed consent.

Participants completed a measure of narcissism (see below) and were then instructed to bring to mind a real-life situation in which someone had hurt them. Next, they were requested to “indicate your current thoughts and feelings about the person who hurt you; that is, we want to know how you feel about that person *right now*.” Participants then completed the measures described below. All measures were administered in German language.

2.2. Measures

2.2.1. Narcissism

The Narcissistic Admirations and Rivalry Questionnaire (NARQ; Back et al., 2013) was used to assess both facets of narcissism, admiration (9 Items; e.g., “I am great”) and rivalry (9 Items; e.g., “I want my rivals to fail”). Items were rated from 1 (*do not agree at all*) to 6 (*agree completely*).

2.2.2. State anger

Participants were asked to rate their momentary anger toward the transgressor on a face-valid single item (“I am very angry about what he/she did to me”; 1 = *strongly disagree*, 5 = *strongly agree*).

2.2.3. State rumination

Participants also provided information about how much they currently ruminated about the transgression (“I can’t stop thinking about what he/she did to me”; 1 = *strongly disagree*, 5 = *strongly agree*).

2.2.4. State empathy

To measure empathic feelings for the transgressor, we used the empathy measure by Batson and Shaw (1991). Participants were instructed to rate the extent to which they currently experienced each feeling for their transgressor (8 items; e.g., “softhearted”; 1 = *not at all*, 5 = *very much*).

2.2.5. Lack of forgiveness

Lack of forgiveness was measured with the Transgression-Related Interpersonal Motivations Inventory (TRIM; McCullough et al., 1998). It assesses individuals’ current motivations toward a real-life transgressor and is divided into two subscales: revenge (5 items; e.g., “I’ll make him/her pay”) and avoidance (7 items; e.g., “I withdraw from him/her”). Items were rated from 1 (*strongly disagree*) to 5 (*strongly agree*). High levels of revenge and avoidance indicate a lack of forgiveness.

3. Results

3.1. Descriptive statistics and intercorrelations

Table 1 details descriptive statistics, reliabilities, and intercorrelations for all measures. Because we were interested in the facet-specific effects of admiration and rivalry on lack of forgiveness, we calculated uniqueness scores by regressing each of the two NARC facets on the other one. The resulting residual scores reflect the amount of variance of each facet that is not shared by the other one and, hence, provide a more accurate estimate of the unique contribution of each facet to the narcissism–forgiveness link (for a similar procedure, see Penke & Asendorpf, 2008).

The unique variance of the admiration facet showed no relations with revenge, avoidance, and anger, but it was negatively correlated with rumination and positively correlated with empathy. The unique variance of the rivalry facet, by contrast, was positively correlated with revenge, avoidance, anger, and rumination, but negatively correlated with empathy.

3.2. Mediation analyses

To test the possible mediating roles of state anger, state rumination, and state empathy, we performed bootstrapping analyses

Table 1
Means, standard deviations, and intercorrelations among measures.

Measure	<i>M</i>	<i>SD</i>	α	1	2	3	4	5	6	7	8	9
Admiration	3.08	.83	.84	–								
Rivalry	2.33	.83	.83	.39**/–.39**	–							
State anger ^a	3.80	1.00	–	.02/–.03	.11**/.12***	–						
State rumination ^a	3.30	1.12	–	–.04/–.07 [†]	–.06/.08**	.42***	–					
State empathy	2.24	.94	.93	.05/.09**	–.07/–.10***	–.25***	.01	–				
Revenge	2.12	.80	.84	.12**/–.02	.35**/.33***	.37***	.22***	–.38***	–			
Avoidance	3.43	.97	.90	–.01/–.04	.07/.08**	.27***	.09**	–.63***	.37***	–		
Gender	–	–	–	.20**/.13***	.21**/.15***	–.13***	–.13***	–.03	.12***	–.01	–	
Age	22.8	4.9	–	–.02/.02	–.09**/–.09**	–.02	–.04	.01	–.05	.03	.05	–

Note. $N = 1040$. For admiration and rivalry, uniqueness correlations (i.e., the amount of variance of each facet that is not shared by the other one) are shown next to the respective zero-order correlations with the other measures. Gender was coded as female = 0, male = 1.

^a Single-item measures.

[†] $p \leq .05$.

** $p \leq .01$.

*** $p \leq .001$ (two-tailed).

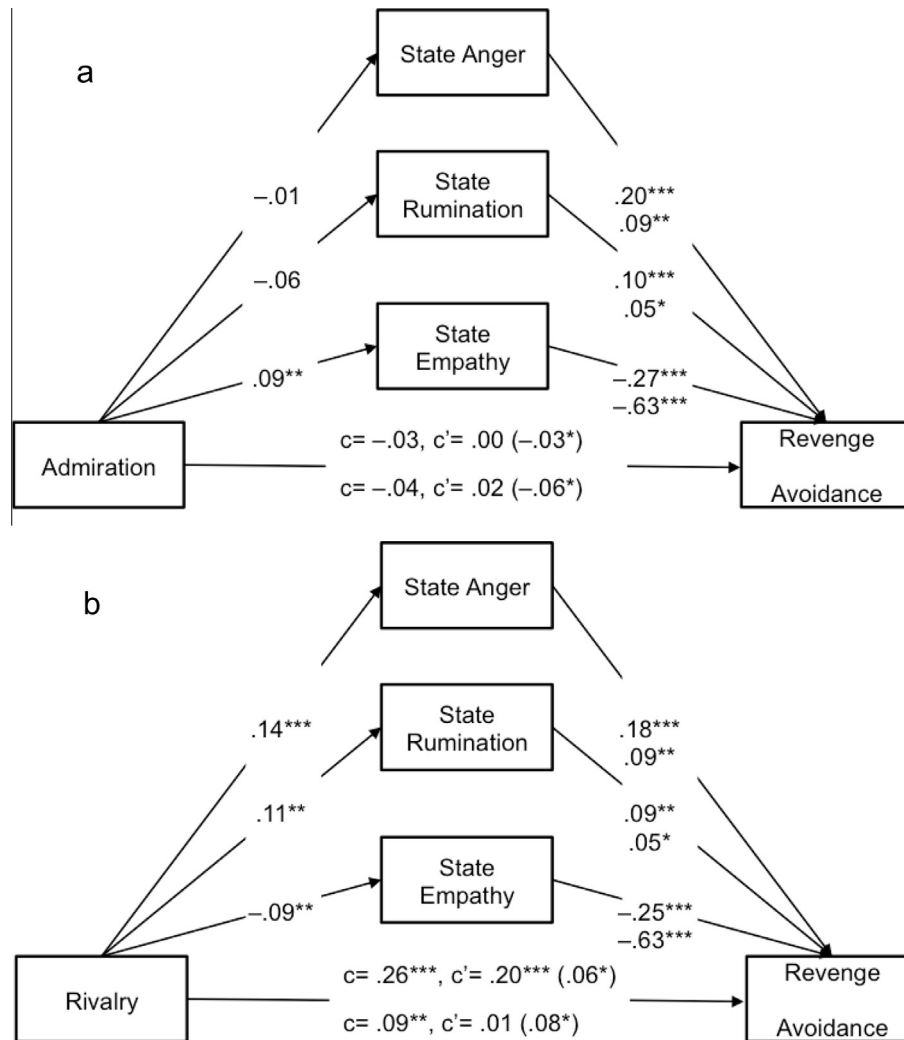


Fig. 1. Multiple mediation models for the unique effects of admiration (panel a) and rivalry (panel b) on unforgiveness facets (revenge and avoidance) via state anger, state rumination, and state empathy. Coefficients represent standardized regression coefficients while controlling for gender and age (not shown in figure); *c* is the total effect of each narcissism facet on revenge/avoidance, and *c'* is the direct effect after correcting for proposed mediators; total indirect effects are in parentheses. $N = 1040$. * $p < .05$. ** $p < .01$. *** $p < .001$ (two-tailed).

using the PROCESS macro for SPSS (Hayes, 2013); we applied 5000 bootstrap resamples with 95% bias-corrected confidence intervals (CIs). A significant indirect effect of a mediator is present for CIs that do not contain zero. A total of four models were computed with the NARQ uniqueness scores (admiration and rivalry) as independent variables and lack of forgiveness (revenge and avoidance) as dependent variables. All models controlled for gender (coded as female = 0, male = 1) and age given that past research has found that these variables are correlated with both narcissism and forgiveness (Foster, Campbell, & Twenge, 2003; Ghaemmaghami, Allemann, & Martin, 2011).

The total effects of admiration on revenge as well as avoidance were not significant, respectively (both $ps > .20$). However, an indirect effect of a mediator can still occur (Preacher & Hayes, 2004). Indeed, state empathy was a significant mediator (95% CI [-.040, -.008]), such that greater admiration was associated with less revenge through higher empathy for the transgressor (Fig. 1a). For avoidance as an outcome variable, the pattern of results was similar: Again, greater admiration was associated with less avoidance through higher empathy for the transgressor (95% CI [-.090, -.016]).

The total effect of rivalry on revenge was significant ($p < .001$). State anger (95% CI [.014, .039]), state rumination (95% CI [.004,

.019]), and state empathy (95% CI [.009, .040]) were significant mediators, such that rivalry was associated with higher revenge through greater anger, greater rumination, and less empathy (Fig. 1b). When avoidance was analyzed as outcome variable, results again revealed significant multiple mediation effects of state anger (95% CI [.005, .024]), state rumination (95% CI [.001, .014]), and state empathy (95% CI [.020, .094]), such that rivalry was associated with higher avoidance through greater anger and rumination and less empathy.

4. Discussion

Drawing on a large and heterogeneous sample, the current study examined the associations between two distinct dimensions of narcissism – admiration and rivalry – and lack of forgiveness (i.e., revenge and avoidance). Admiration was unrelated to revenge and avoidance, but the overall zero effects may be explained by the mitigating effect of state empathy: Individuals high in admiration experienced more empathic feelings for their transgressor, which in turn predicted less revenge and avoidance. Thus, empathy functioned as a suppressor variable such that its inclusion in mediation models yielded significant negative relationships between admiration and revenge/avoidance. Rivalry, by contrast, was related to

greater revenge and this association was mediated by state anger, state rumination, and state empathy. Conceptually, our results underscore the importance of distinguishing between two narcissistic dimensions when considering reactions to interpersonal hurt: one that entails responses that promote forgiveness (admiration) and another that entails adverse responses (rivalry).

In line with previous studies illustrating that narcissism is linked to negative interpersonal outcomes (Back et al., 2013; Eaton et al., 2006), we found that it is the antagonistic (i.e., rivalry) aspect of narcissism that is most robustly associated with revenge in response to real-life transgressions. Importantly, we conceptually replicated findings by Exline et al. (2004), who showed that narcissistic entitlement is associated with less forgiveness of transgressions. Our findings move beyond past work, however, by scrutinizing the mechanisms by which facets of narcissism and lack of forgiveness are related.

Our results may imply that the cognitive mechanisms mediating this link should be extended to include further variables, such as higher perceived intentionality of the transgression or the attribution of malevolent intentions on the side of the perpetrator. Given that individuals high in rivalry (but not admiration) are generally distrustful of others (Back et al., 2013), these attribution styles might be plausible. Indeed, attributing a transgression as being intentional has been found to be negatively related to forgiveness (Fehr et al., 2010). Similarly, inferring malevolent intentions even in seemingly regretful perpetrators has been shown to be linked to pronounced unforgiveness (Gerlach, Allemand, Agroskin, & Denissen, 2012).

Finally, our results also suggest that facets of narcissism might differ in their adaptiveness: Admiration was positively related to empathy for the transgressor but it was negatively related to rumination about the hurt. Rivalry, however, demonstrated inverse relations with these socio-cognitive variables; it was negatively related to empathy but positively related to rumination. Although consistent with recent conceptualizations of admiration as the bright side of narcissism and rivalry as its dark side (Back et al., 2013), it should also be noted that such classification is rather crude (Robins, Tracy, & Shaver, 2001) and awaits further qualification in future research.

4.1. Limitations and future directions

While our large sample allowed us to detect the observed effects with sufficient precision (Schönbrodt & Perugini, 2013), admittedly, some of the effects were rather modest in size (e.g., the significant suppression effects observed for empathy in the analysis of admiration and lack of forgiveness) and thus warrant interpretative caution. Another point refers to the use of single-item measures of state anger and state rumination. Although this has been shown to produce valid data (e.g., Allemand, Steiner, & Hill, 2013), future work should include multi-item measures of these constructs. We also encourage the use of non self-report measures, as individuals might not always be aware of their thoughts and feelings toward their transgressor.³ Such measures could include behavioral observations or indirect measures of forgiveness (Fatfouta, Schröder-Abé, & Merkl, 2014).

Given the cross-sectional nature of our study, causal pathways regarding the relations between admiration as well as rivalry and lack of forgiveness cannot be established. Longitudinal or experimental studies that investigate this more thoroughly are certainly an avenue for future work. Such studies might also use other conceptualizations and measures of narcissism (for an overview, see Rhodewalt & Peterson, 2009) to examine the generalizability of

the current results. In addition, our study focused on one specific transgression situation. Using experience-sampling designs, future studies could examine whether the present relations also hold on a day-to-day basis and with different types of transgressions. Previous research, for example, has shown that narcissists often present themselves as victims; that is, they report an increased number of interpersonal transgressions in their daily interactions (McCullough, Emmons, Kilpatrick, & Mooney, 2003). We would anticipate the antagonistic style of individuals high in rivalry to promote enhanced recollection of past transgressions rather than the assertive style of individuals high in admiration. Finally, it would be valuable to further examine the narcissism–forgiveness link in close (i.e., romantic) relationships and, specifically, narcissists' reactions to partner transgressions. As outlined in Back et al. (2013), using partner reports of relationship behaviors (e.g., accommodation) and perceptions (e.g., commitment) might be beneficial. Again, we would expect individuals high in rivalry to demonstrate less accommodative behavior and commitment toward their partner than individuals high in admiration. Future studies are necessary to test these predictions in more depth.

4.2. Conclusions

The current study adds to our limited understanding of the processes underlying the links between two distinct facets of narcissism – admiration and rivalry – and lack of forgiveness. We demonstrated that this association is differentially mediated by state anger, state rumination, and state empathy. Subsequent work is needed to identify further cognitive mechanisms of this link and, ultimately, to better clarify the causal pathways among these variables.

References

- Ackerman, R. A., Witt, E. A., Donnellan, M. B., Trzesniewski, K. H., Robins, R. W., & Kashy, D. A. (2011). What does the narcissistic personality inventory really measure? *Assessment, 18*(1), 67–87.
- Allemand, M., Steiner, M., & Hill, P. L. (2013). Effects of a forgiveness intervention for older adults. *Journal of Counseling Psychology, 60*(2), 279–286. <http://dx.doi.org/10.1037/a0031839>.
- Back, M. D., Küfner, A. C. P., Dufner, M., Gerlach, T. M., Rauthmann, J. F., & Denissen, J. J. A. (2013). Narcissistic admiration and rivalry: Disentangling the bright and dark sides of narcissism. *Journal of Personality and Social Psychology, 105*(6), 1013–1037. <http://dx.doi.org/10.1037/a0034431>.
- Batson, C. D., & Shaw, L. L. (1991). Evidence for altruism: Toward a pluralism of prosocial motives. *Psychological Inquiry, 2*(2), 107.
- Eaton, J., Struthers, C. W., & Santelli, A. G. (2006). Dispositional and state forgiveness: The role of self-esteem, need for structure, and narcissism. *Personality and Individual Differences, 41*(2), 371–380.
- Emmons, R. A. (2000). Personality and forgiveness. In M. E. McCullough, K. I. Pargament, & C. E. Thoresen (Eds.), *Forgiveness: Theory, research, and practice* (pp. 156–175). New York, NY, USA: Guilford Press.
- Exline, J. J., Baumeister, R. F., Bushman, B. J., Campbell, W. K., & Finkel, E. J. (2004). Too proud to let go: Narcissistic entitlement as a barrier to forgiveness. *Journal of Personality and Social Psychology, 87*(6), 894–912. <http://dx.doi.org/10.1037/0022-3514.87.6.894>.
- Fatfouta, R., Schröder-Abé, M., & Merkl, A. (2014). Forgiving, fast and slow: Validity of the implicit association test for predicting differential response latencies in a transgression-recall paradigm. *Frontiers in Psychology, 5*, 1–14. <http://dx.doi.org/10.3389/fpsyg.2014.00728>.
- Fehr, R., Gelfand, M. J., & Nag, M. (2010). The road to forgiveness: A meta-analytic synthesis of its situational and dispositional correlates. *Psychological Bulletin, 136*(5), 894–914. <http://dx.doi.org/10.1037/a0019993>.
- Foster, J. D., Campbell, W. K., & Twenge, J. M. (2003). Individual differences in narcissism: Inflated self-views across the lifespan and around the world. *Journal of Research in Personality, 37*(6), 469–486.
- Gerlach, T. M., Allemand, M., Agroskin, D., & Denissen, J. J. A. (2012). Justice sensitivity and forgiveness in close interpersonal relationships: The mediating role of mistrustful, legitimizing, and pro-relationship cognitions. *Journal of Personality, 80*(5), 1373–1413. <http://dx.doi.org/10.1111/j.1467-6494.2012.00762.x>.
- Ghaemmaghami, P., Allemand, M., & Martin, M. (2011). Forgiveness in younger, middle-aged and older adults: Age and gender matters. *Journal of Adult Development, 18*(4), 192–203. <http://dx.doi.org/10.1007/s10804-011-9127-x>.
- Hayes, A. F. (2013). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. New York, NY, USA: Guilford Press.

³ We thank an anonymous reviewer for pointing us to this possibility.

- Krizan, Z., & Johar, O. (in press). Narcissistic Rage Revisited. *Journal of Personality and Social Psychology*.
- McCullough, M. E., Emmons, R. A., Kilpatrick, S. D., & Mooney, C. N. (2003). Narcissists as "Victims": The role of narcissism in the perception of transgressions. *Personality and Social Psychology Bulletin*, 29(7), 885–893.
- McCullough, M. E., Rachal, K. C., Sandage, S. J., Worthington, E. L., Jr., Brown, S. W., & Hight, T. L. (1998). Interpersonal forgiving in close relationships: II. Theoretical elaboration and measurement. *Journal of Personality and Social Psychology*, 75(6), 1586–1603. <http://dx.doi.org/10.1037/0022-3514.75.6.1586>.
- Penke, L., & Asendorpf, J. B. (2008). Beyond global sociosexual orientations: A more differentiated look at sociosexuality and its effects on courtship and romantic relationships. *Journal of Personality and Social Psychology*, 95(5), 1113–1135. <http://dx.doi.org/10.1037/0022-3514.95.5.1113>.
- Preacher, K. J., & Hayes, A. F. (2004). SPSS and SAS procedures for estimating indirect effects in simple mediation models. *Behavior Research Methods, Instruments, & Computers: A Journal of the Psychonomic Society Inc.*, 36(4), 717–731.
- Raskin, R. N., & Hall, C. S. (1979). A narcissistic personality inventory. *Psychological Reports*, 45(2), 590.
- Rhodewalt, F., & Peterson, B. (2009). Narcissism. In M. R. Leary & R. H. Hoyle (Eds.), *Handbook of individual differences in social behavior* (pp. 547–560). New York, NY, USA: Guilford Press.
- Robins, R. W., Tracy, J. L., & Shaver, P. R. (2001). Shamed into self-love: Dynamics, roots, and functions of narcissism. *Psychological Inquiry*, 12(4), 230–236.
- Schönbrodt, F. D., & Perugini, M. (2013). At what sample size do correlations stabilize? *Journal of Research in Personality*, 47(5), 609–612. <http://dx.doi.org/10.1016/j.jrp.2013.05.009>.
- Twenge, J. M., & Campbell, W. K. (2003). "Isn't it fun to get the respect that we're going to deserve?" Narcissism, social rejection, and aggression. *Personality and Social Psychology Bulletin*, 29(2), 261–272.
- Watson, P. J., & Morris, R. J. (1991). Narcissism, empathy and social desirability. *Personality and Individual Differences*, 12(6), 575–579.

Can the shrew be tamed?

Narcissism, lack of forgiveness, and the moderating role of conciliatory gestures

Ramzi Fatfouta^{a,b*} & Tanja M. Gerlach^{c*}

^aDepartment of Education and Psychology, Freie Universität Berlin, Germany

^bCluster “Languages of Emotion”, Freie Universität Berlin, Germany

^cDepartment of Psychology, Georg August University Goettingen, Germany

*Shared first authorship

Author Note

Study 1 was supported by the Cluster “Languages of Emotion” at Freie Universität Berlin. Study 2 was supported by the International Max Planck Research School “The Life Course: Evolutionary and Ontogenetic Dynamics” (LIFE). We thank Angelina Bohlender for her help with data collection in Study 3. Please address correspondence to Ramzi Fatfouta, Department of Education and Psychology, Freie Universität Berlin, Habelschwerdter Allee 45, 14195 Berlin, Germany, Email: ramzi.fatfouta@fu-berlin.de or Tanja M. Gerlach, Institute of Psychology, Georg August University Goettingen, Gosslerstrasse 14, 37073 Goettingen, Germany, Email: tanjamgerlach@googlemail.com.

Abstract

Three studies (total $N = 2,127$) investigated how apology moderates the relationship between two facets of narcissism — admiration and rivalry — and lack of forgiveness (i.e., revenge, avoidance vs. benevolence). In Study 1 ($N = 767$), we assessed self-reported perceptions of apology following a real-life transgression. In Study 2 ($N = 930$), we used an experimental design to manipulate the presence (vs. absence) of an apology following a hypothetical transgression. In Study 3 ($N = 430$), we extended the experimental manipulation to include specific conciliatory gestures. Across all three studies, narcissistic rivalry (but less so admiration) was positively related to a lack of forgiveness. Further, the relation between narcissistic rivalry and revenge appeared to be robust against attempts of a transgressor to reconcile. Importantly, results of Study 3 indicated that the link between rivalry and avoidance could in fact be attenuated when the transgressor offered a costly apology in addition to verbal expressions of regret. Taken together, results indicate that narcissists can be partly appeased, provided that one is willing to incur a personal cost. Findings underscore the value of a two-dimensional conceptualization of narcissism and allow for a more nuanced view on narcissism's interplay with conciliatory gestures following interpersonal hurt.

Can the shrew be tamed?

Narcissism, lack of forgiveness, and the moderating role of conciliatory gestures

Interpersonal conflict pervades social life, yet many relationships retain their value despite the prior imposition of harm (McCullough, Kurzban, & Tabak, 2011, 2013). Our ability to forgive others' harmful actions may be key to dealing with relationship-threatening events (Fincham, 2000). At first sight it may be tempting to return like for like. However, relinquishing such desire can prevent continual cycles of retaliation (McCullough et al., 1998) — a notion consistent with evidence that forgiveness is positively associated with pro-relational responses (Fincham & Beach, 2002; Karremans & Van Lange, 2004). Aside from relational benefits, forgiveness is also associated with various benefits for individual health and well-being (for a review, see Worthington & Scherer, 2004), such as greater life satisfaction (Brown & Phillips, 2005), higher emotional stability (McCullough & Hoyt, 2002), and lower levels of depression (Brown, 2003).

Forgiveness involves a prosocial change in interpersonal motivation in response to a perceived wrong (McCullough, Fincham, & Tsang, 2003). On the one hand, one becomes less vengeful and less avoidant toward the transgressor (McCullough et al., 1998) and, on the other hand, more benevolent (i.e., motivated by conciliation and goodwill; McCullough, Root, & Cohen, 2006). Indeed, positive *and* negative dimensions of forgiveness are essential in a close-relationship context (Fincham, Hall, & Beach, 2005), and this notion is echoed in studies finding distinct correlates of these two dimensions (Fincham, Beach, & Davila, 2004).

Within the research on interindividual differences, one important question concerns dispositional correlates of forgiveness: Who forgives others and who does not? Conclusive evidence suggests that those high in neuroticism and low in agreeableness are less likely to forgive transgressions (Berry, Worthington, O'Connor, Parrott, & Wade, 2005; McCullough & Hoyt, 2002), whereas those who are high in empathy tend to be more forgiving

(McCullough et al., 1998; McCullough, Worthington, & Rachal, 1997). Importantly, narcissism — in its non-clinical conceptualization as an individual-difference variable — is considered to hinder forgiveness (Emmons, 2000). In particular, those high in narcissism have shown to respond in a vengeful manner when transgressed against (Brown, 2004; Eaton, Struthers, & Santelli, 2006a).

Narcissists (i.e., individuals scoring high on measures of non-clinical narcissism) have been characterized by a grandiose self-view, self-enhancement, exploitative interpersonal behavior, and lack of concern for others (Rhodewalt & Peterson, 2009) – in short, qualities that seem incommensurate with a forgiving stance. Furthermore, narcissists have often been described as “disagreeable extraverts” (Paulhus, 2001). In other words, starting new relationships is relatively easy for them (Bradlee & Emmons, 1992), but at the same time they exhibit aggressive and combative behavior in these relationships (Küfner, Nestler, & Back, 2013). As a consequence, initial positive impressions and others’ liking often fade after a series of interactions and interpersonal conflicts become more likely (Ackerman et al., 2011; Paulhus, 1998).

Although narcissism is frequently thought of as a unitary construct, the recently introduced Narcissistic Admiration and Rivalry Concept (NARC; Back et al., 2013) distinguishes two social strategies by which narcissists’ grandiose self-view can be maintained: First, the propensity for agentic self-enhancement by means of self-promotion (*admiration*) and, second, the propensity for antagonistic self-protection by means of self-defense (*rivalry*). Empirically, admiration is related to adjustment indicators (e.g., self-assured, sociable, and competent; problem-focused reactions to transgressions), whereas rivalry is related to maladjustment (e.g., aggressive, entitled, and unlikable; more hostile and unforgiving reactions; Back et al., 2013; Fatfouta, Gerlach, Schröder-Abé, & Merkl, 2015).

Narcissists' pronounced interpersonal problems (Ackerman et al., 2011; Back et al., 2013) raise the question about conditions that might change the positive association between narcissistic tendencies and lack of forgiveness. Accordingly, it seems particularly appealing to investigate potential moderators of this link. Conciliatory gestures have been posited to play a critical role in facilitating forgiveness — in particular, amends/apologies have been found to help victims let go of the hurt incurred (for a meta-analytic review, see Fehr, Gelfand, & Nag, 2010).

Apology can be defined as an acknowledgement of responsibility for the transgression, coupled with an expression of regret (Kim, Ferrin, Cooper, & Dirks, 2004). Such an acknowledgment can either be delivered by means of verbal (e.g., saying "I am sorry!"; Schlenker & Darby, 1981) or, more subtle, non-verbal communication (e.g., eye contact, "sad looks"; Kelley & Waldron, 2005; Waldron & Kelley, 2008). Furthermore, apology has been shown to result in more favorable impressions of and less destructive behaviors toward the transgressor (McCullough, Pedersen, Tabak, & Carter, 2014; Ohbuchi, Kameda, & Agarie, 1989; Tabak, McCullough, Luna, Bono, & Berry, 2012). Also, apology can serve as an effective signal that similar transgressions are less likely to occur in the future (McCullough et al., 2013).

We propose that narcissists may be more prone to forgive their transgressor when being offered an apology. Consistent with this, Eaton, Struthers, and Santelli (2006b) suggested that the act of apologizing might reduce potential ego-threats. If this were the case, apology could serve as a buffer against defensive reactions related to narcissistic rivalry (but less so to admiration). In other words, when the transgressor acknowledges the wrongdoing and expresses regret, this should directly attenuate the unforgiving responses of those high in rivalry. It therefore seems likely that perceptions of apology should moderate the link between rivalry and lack of forgiveness.

To date, no study has examined this possibility, although indirect evidence exists. In a series of six studies by Exline and colleagues (Exline, Baumeister, Bushman, Campbell, & Finkel, 2004), narcissistic entitlement as assessed with the Narcissistic Personality Inventory (NPI; Raskin & Hall, 1979) has been linked to greater insistence on concessions for the self before forgiving. According to Exline et al. (2004), entitled narcissists readily perceive that others owe them debts and also view these debts to be particularly large. Narcissistic entitlement, which involves expectations of special treatment (Campbell, Bonacci, Shelton, Exline, & Bushman, 2004), is more strongly correlated with rivalry than with admiration (Back et al., 2013). Thus, the fact that entitled narcissists make forgiveness contingent on concessions provides a plausible, albeit indirect, hint that perceptions of apology might indeed appease individuals high in rivalry.

Overview of the present research

In the present research, our main goal was to test whether and to what extent apology moderates the relationship between narcissism and lack of forgiveness. We addressed this goal by conducting one correlational and two experimental studies. Across these three studies, we examined narcissistic tendencies as outlined in the NARC: as two related yet distinguishable dimensions of narcissism. In line with the above reasoning, we hypothesized that narcissistic rivalry (but less so admiration) would be positively related to lack of forgiveness following an interpersonal transgression. Furthermore, we tested the following moderation hypothesis: With the perception that the transgressor apologized for his/her harmful actions, individuals high in rivalry should become more prone to forgive their transgressor. Specifically, we predicted that the link between narcissistic rivalry and lack of forgiveness would be attenuated when an apology is offered as compared to when this is not the case.

Study 1

Study 1 examined whether perceptions of an apology by the transgressor would moderate the relationship between narcissistic rivalry and forgiveness of a real-life transgression. Specifically, we hypothesized that narcissistic rivalry would interact with apology, in a way that among people high in rivalry greater perceptions of apology by the transgressor would be associated with increased forgiveness.

Method

Participants

A total of 767 participants ($M_{\text{age}} = 24.02$, $SD = 5.19$; 76.6% female) from Germany, Austria, and Switzerland were recruited via advertising on a social-networking site to complete an online study. As an incentive, participants were offered entry into a raffle for a 25€ gift card.¹

Procedure

Participants completed a measure of narcissism (see below) and were then instructed to bring to mind a real-life situation in which someone had hurt them. Next, participants were instructed to “indicate your current thoughts and feelings about that person; that is, we want to know how you feel about that person *right now*.” Following this instruction, participants were requested to indicate their degree of forgiveness and, then, the degree to which the transgressor had apologized for his/her behavior.

Measures

Narcissism was assessed using the Narcissistic Admiration and Rivalry Questionnaire (NARQ; Back et al., 2013) consisting of two subscales: admiration (9 Items; e.g., “I am

¹More details on this sample can be found in Fatfouta (in press).

great", $\alpha = .85$) and rivalry (9 Items; e.g., "I want my rivals to fail", $\alpha = .81$). Items were rated from 1 (*do not agree at all*) to 6 (*agree completely*).

Forgiveness was assessed using the Transgression-Related Interpersonal Motivations Inventory (McCullough et al., 1998; McCullough et al., 2006). The TRIM assesses individuals' motivations toward a real-life transgressor and consists of three subscales: revenge (5 items; e.g., "I wish that something bad would happen to him/her", $\alpha = .78$), avoidance (7 items; "I am trying to keep as much distance between us as possible", $\alpha = .86$), and benevolence (6 items; "I have given up my hurt and resentment", $\alpha = .82$). Items were rated from 1 (*strongly disagree*) to 7 (*strongly agree*). Low levels of revenge and avoidance and high levels of benevolence indicate a more forgiving stance toward the transgressor. Given that the factor structure of the TRIM-18 has never before been examined in a German-speaking sample, we performed a principal-components analysis (PCA; with oblimin rotation). After eliminating two items (items 7 and 15) with low factor loadings ($< .40$), the PCA yielded three factors with eigenvalues > 1 (avoidance, revenge, benevolence) that explained 35.36%, 11.17%, and 8.99% of the variance, respectively.

The perceived degree of apology was assessed with a scale previously used by McCullough et al. (1997). This scale consists of two items that were rated from 1 (*not at all*) to 5 (*very much*). Items tapped into the degree to which participants perceived that their transgressors showed signs of remorse and apologized (e.g., "To what degree was the person remorseful and apologized for his/her behavior?", $\alpha = .84$).

Results and discussion

Scores on the three subscales of the forgiveness measure were non-normally distributed ($1.12 > \text{skew} > 2.39$). To accommodate this, we conducted nonparametric tests in all analyses involving these subscales. In particular, we computed rank-based correlations

(Kendall's τ) instead of product-moment correlations (Pearson's r) and used bootstrap ordinary least squares (OLS) with 1,000 samples instead of standard OLS (Fox, 2008).

Descriptive statistics and intercorrelations

As we were interested in the facet-specific effects of admiration and rivalry on forgiveness, we calculated uniqueness scores by regressing each of the two NARC facets onto the other one (for a similar procedure, see Penke & Asendorpf, 2008). The resulting residual scores reflect the amount of variance of each narcissism facet that is not shared by the other one and, hence, provide a more accurate estimate of the unique contribution of each facet to the narcissism-forgiveness link (Fatfouta et al., 2015). Table 1 details descriptive statistics and intercorrelations for all measures.

--- Table 1 about here ---

The unique variance of the admiration facet showed no relations with revenge, but it was negatively correlated with avoidance and positively correlated with benevolence and apology. The unique variance of the rivalry facet, by contrast, was positively correlated with revenge and avoidance, but negatively correlated with benevolence and perceived apology.

Regression analyses

We conducted hierarchical bootstrap regression analyses to determine interactive contributions of narcissism and apology in predicting forgiveness. In all analyses, we included age and gender (coded as female = 0, male = 1) as covariates given meta-analytic evidence that these variables can be correlated with both narcissism (Grijalva et al., 2014) and forgiveness (Miller, Worthington, & McDaniel, 2008). However, results concerning the covariates were not of interest here and, hence, are not discussed further.

For each NARQ uniqueness score (i.e., admiration and rivalry), we analyzed a separate regression model in which the dependent variables were revenge, avoidance, and benevolence, respectively. Thus, a total of six models were computed with age and gender entered in Step 1, the respective NARQ uniqueness score entered in Step 2, apology entered in Step 3, and the respective two-way interaction between NARQ uniqueness score and apology entered in Step 4. Continuous variables were standardized and the interaction terms were computed based on the standardized variables (Aiken & West, 1991). Table 2 details summary statistics (see Appendix A for complete model statistics).

--- Table 2 about here ---

Results showed that, after accounting for covariates (i.e., age and gender), admiration uniquely predicted less avoidance and more benevolence, but it was not predictive of revenge. Rivalry, by contrast, uniquely predicted more revenge, avoidance, and less benevolence. In addition, perceptions of apology were unique predictors of forgiveness (i.e., low levels of revenge and avoidance and high levels of benevolence) above and beyond covariates and admiration/rivalry. However, the hypothesized interaction between rivalry and apology was not significant. Perceptions of apology thus did not moderate the effects of rivalry on forgiveness, indicating that such perceptions made virtually no difference in how narcissists high in rivalry responded to their transgressor.

Study 2

Although Study 1 shed new light on the relations between narcissism, apology, and forgiveness, it remained an open question as to whether apologies are ineffective for narcissists or whether design limitations precluded the possibility of detecting a moderating effect. Study 2 therefore extended Study 1 in two important ways: First, we used an

experimental design to test whether the presence (vs. absence) of an apology affects narcissists' responses to an interpersonal transgression. Second, we used standardized transgression incidents to control for potential contextual confounds. If the influence of narcissism on forgiveness is indeed moderated by apology, then we should observe the anticipated moderating effect. If, however, apology exerts no influence on how narcissists respond to their transgressor, we should again observe no interaction between narcissism and apology in predicting forgiveness.

Method

Participants

A total of 930 participants ($M_{\text{age}} = 27.76$, $SD = 8.92$; 74.5% female) from Germany were recruited via a mailing list to complete an online study as part of a larger research project examining associations between justice sensitivity and forgiveness (see Gerlach, Allemand, Agroskin, & Denissen, 2012b). As an incentive, participants obtained individualized feedback on their personality traits. Of interest to the present study were the analyses involving narcissism and forgiving motivations, which have not been reported before.

Procedure

As in Study 1, participants completed a measure of narcissism. They were then randomly assigned to read one of three transgression scenarios, each describing a close friend treating them badly (see below). After having committed the transgression, the friend either tried to pass over the situation by ignoring the topic when it came up in a conversation (i.e., did not apologize) or showed conciliatory behavior (i.e., apologized). Participants were instructed to imagine themselves as the transgressed party in the situation and to indicate their degree of forgiveness.

Measures

Narcissism was assessed using the brief version of the NARQ (Back et al., 2013) consisting of the same subscales as in Study 1. Each subscale had three items that were rated from 1 (*do not agree at all*) to 6 (*agree completely*). Internal consistency (α) was .67 and .65 for admiration and rivalry, respectively.

Forgiveness was assessed using an adapted version of the Marital Forgiveness Scale – Event (Fincham et al., 2004), rephrased to reflect close-friend relationships. Items assessed the degree to which individuals experienced specific reactions in response to their friend’s transgression. These reactions included retaliating against the friend (3 Items; “I thought about how I could get even”), avoiding the friend (3 Items; “I don’t want to have anything to do with her/him”), and/or reconciling with the friend (3 Items; “I want to forgive him/her quickly”). Items were rated from 1 (*not at all likely*) to 6 (*very much likely*). As in Study 1, we first inspected the dimensionality of the forgiveness measure. Here, the PCA (with oblimin rotation) yielded two factors with eigenvalues >1 (explaining 66.88% of the variance): The first factor contained the avoidance as well as the (reverse-coded) benevolence items and was therefore labeled “avoidance versus benevolence” (explaining 47.19% of the variance; $\alpha = .89$). The second factor contained the revenge items and was therefore labeled “revenge” (explaining 19.69% of the variance; $\alpha = .73$). To note, a similar two-factor solution has been proposed in previous research (e.g., McCullough et al., 2006). We created two composite indices accordingly, which we used as dependent variables.

Transgression scenarios

A detailed description of the scenarios as well as information on their pretesting can be found in Gerlach et al. (2012, p. 1394 f.). Briefly, scenarios involved “publicly disclosing a personal detail of the participant”, “talking badly about the participant to someone else”, and “giving away a carefully chosen gift”. After presenting the scenario, the post-

transgression behavior of the perpetrator was manipulated insofar as the friend was either said to have (a) taken up the topic briefly, but then turned the conversation to something else (condition: no apology) or (b) asked for pardon, said that he/she is very sorry, and offered amends (condition: apology) when the topic came up in a later conversation. Descriptions of perpetrator behavior were kept brief and comparable in length.

Results and discussion

Scores from the revenge subscale were non-normally distributed (skew = 1.51). Therefore, we used nonparametric tests in all analyses involving this subscale (i.e., Kendall's τ and bootstrapping regression models).

Descriptive statistics and intercorrelations

As in Study 1, we calculated uniqueness scores by regressing each of the two NARC facets onto the other one. Table 3 details descriptive statistics and intercorrelations for all measures.

--- Table 3 about here ---

Whereas the unique variance of the admiration facet had negligible effects on forgiveness, the unique variance of the rivalry facet was positively correlated with revenge and avoidance versus benevolence.

Regression analyses

As in Study 1, we conducted hierarchical multiple regression analyses with narcissism, apology condition, and their interaction as predictors of forgiveness (i.e., revenge and avoidance vs. benevolence). In all analyses, continuous variables were *z*-standardized and the interaction terms were computed based on the standardized variables (Aiken & West, 1991). A total of four models were computed with age and gender (coded as female = 0,

male = 1) entered in Step 1, scenario type (dummy-coded with “publicly disclosing a personal detail of the participant” as the reference category) entered in Step 2, the respective NARQ uniqueness score entered in Step 3, apology condition (coded as no apology = 0, apology = 1) entered in Step 4, and the respective two-way interaction between NARQ uniqueness score and apology condition entered in Step 5. Table 4 details summary statistics (see Appendix B for complete model statistics).

--- Table 4 about here ---

Results showed that, after accounting for covariates (i.e., age, gender, and scenario type), rivalry (but not admiration) uniquely predicted more revenge and avoidance versus benevolence. Moreover, the presence (vs. absence) of an apology was a unique predictor of forgiveness (i.e., low levels of revenge and avoidance and high levels of benevolence) above and beyond covariates and admiration/rivalry. As in Study 1, no interaction effects between admiration/rivalry and apology emerged. Stated differently, apology seemed to have no causal influence on narcissists’ lack of forgiveness.

Study 3

So far, our findings suggested that self-reported perceptions of apology (Study 1) as well as direct manipulations of apology (Study 2) do not affect the relationship between narcissism and forgiveness. However, in both studies we confronted participants with rather broad descriptions of “apology”. In Study 1, we assessed perceptions of apology in response to a real-life transgression, thus allowing for subjective definitions and idiosyncratic interpretations of events. In Study 2, we combined multiple conciliatory gestures into one apology condition (i.e., explicit acknowledgment/apology together with amends; see Schmitt, Gollwitzer, Förster, & Montada, 2004). Hence, the specific nature of each conciliatory

gesture's contribution to the apology remains unclear. We reasoned that there might exist a *specific* conciliatory gesture that facilitates narcissists' forgiveness. Given that the sincerity of a verbal apology is often uncertain (e.g., cheap talk; Farrell, 1987; Gerlach, Agroskin, & Denissen, 2012a), the payment of "penance" might be required to validate the apology (Bottom, Gibson, Daniels, & Murnighan, 2002; Ohtsubo & Watanabe, 2009). Consequently, narcissists may require more costly reparative acts (e.g., a specific compensation) to forgive their transgressors. Study 3 therefore extends Studies 1 and 2 by testing the prediction that apology will promote narcissists' forgiveness only when it comprises some sort of specific reparation for the individual.

Method

Participants

A total of 430 participants ($M_{\text{age}} = 25.88$, $SD = 7.54$; 67% female) from Germany were recruited via e-mail, word-of-mouth, and advertising on a social-networking site to participate in an online study. As an incentive, participants again obtained individualized feedback on their personality traits.

Procedure

Similar to Study 2, participants first completed a measure of narcissism (see below) and were then randomly assigned to read one of two transgression scenarios, each describing a close friend treating them badly. After having committed the transgression, the friend either tried to pass over the situation by ignoring the topic when it came up in conversation (i.e., did not apologize) or showed one of four conciliatory behaviors (which varied in expressed regret). Participants were invited to imagine themselves in the situation and to indicate their degree of forgiveness.

Measures

Again, narcissism was assessed using the NARQ (Back et al., 2013). Internal consistencies (α) were .84 and .82 for admiration and rivalry, respectively. Again, forgiveness was assessed using an adapted version of Fincham et al.'s (2004) forgiveness scale. As in Studies 1 and 2, we first analyzed dimensionality of the forgiveness scale. The PCA (with oblimin rotation) replicated the two-factor structure found in Study 2 (explaining 61.98% of the variance) consisting of "avoidance versus benevolence" (explaining 47.05% of the variance; $\alpha = .86$) and "revenge" (explaining 14.93% of the variance; $\alpha = .59$).

Transgression scenarios

Scenarios had been selected on the basis of a pretest. In that pretest, 36 independent raters ($M_{\text{age}} = 31.73$, $SD = 8.50$; 69.4% female) evaluated six scenarios (order randomized) with regard to two characteristics: severity (3 items; e.g. "How severe is the depicted behavior of your friend?", $\alpha = .88$) and perceived realism (2 items; e.g., "How realistic is the imagined situation?". $\alpha = .81$). Items were rated from 1 (*not at all severe/realistic*) to 6 (*very severe/realistic*). Across scenarios, mean severity was 4.33 ($SD = 1.16$) and mean realism was 3.64 ($SD = 1.22$). We selected the following two scenarios representing moderate severity and sufficient realism to ensure that participants would be able to properly imagine the scenario.

Scenario 1 ($M_{\text{severity}} = 3.90$, $SD = 1.11$; $M_{\text{realism}} = 3.61$, $SD = 1.22$): "You want to relax with a good friend of yours for a few days. As he/she lacks the necessary spare cash, you offer him/her to borrow the money. He/she gratefully accepts the offer and you both agree that he/she will pay you back no later than the 15th of the following month. On checking your account balance, however, you realize that your rent could not be debited from your account because your friend did not return the borrowed money in time."

Scenario 2 ($M_{\text{severity}} = 4.49$, $SD = 1.29$; $M_{\text{realism}} = 3.61$, $SD = 1.22$): “A good friend of yours wants to buy a new piece of furniture, but his/her car is too small to transport it. As you own a bigger car, your friend asks you to lend him/her your car. At first, you are uncertain, but you finally agree and ask him/her to be cautious. As agreed, he/she returns your car the following day. On your next car ride, you realize that there is a big scratch in the passenger door.”

Post-transgression conciliatory gestures were similarly selected on the basis of a pretest including the same raters as described above. These raters ranked five conciliatory gestures (order randomized) with regard to the credibility of expressed regret, from 1 (*least credible*) to 5 (*most credible*). The resulting rank order as well as detailed descriptions of conciliatory gestures are displayed in Table 5.

--- Table 5 about here ---

The first conciliatory gesture, no expression of regret, lacked any component of an apology; similar to the no-apology-condition in Study 2, the friend tried to pass over the situation by ignoring it when it came up in a later conversation. The second gesture, verbal expression of regret, contained an explicit acknowledgment/apology by the transgressor. The third gesture, verbal expression of regret and offer of compensation, additionally contained an apology cost that is incurred by the transgressor (see Bottom et al., 2002; Ohtsubo & Watanabe, 2009) — that is, the cancellation of one’s meeting to invite the victim to dinner. The fourth gesture, verbal and nonverbal expression of regret, contained both explicit and implicit accounts of acknowledgment/apology, the latter communicating negative perpetrator feelings via facial display. The fifth and last gesture subsumed all previously described components.

Results and discussion

Scores from the revenge subscale were non-normally distributed (skew = 1.28). To address this, we again used nonparametric tests in all analyses involving this subscale.

Descriptive statistics and intercorrelations

As in Studies 1 and 2, we calculated uniqueness scores by regressing each of the two NARC facets onto the other one. Table 6 details intercorrelations and descriptive statistics for all measures.

--- Table 6 about here ---

Consistent with Study 2, the unique variance of the admiration facet had negligible effects on revenge and avoidance versus benevolence. The unique variance of the rivalry facet, by contrast, was positively correlated with revenge and avoidance versus benevolence.

Regression analyses

We conducted hierarchical multiple regression analyses to determine the extent to which narcissism, conciliatory gestures, and their interactions were related to forgiveness (i.e., revenge and avoidance vs. benevolence). Prior to analyses, continuous variables were *z*-standardized and the interaction terms were computed based on the standardized variables (Aiken & West, 1991). As in Study 2, a total of four models were computed with age and gender (coded as female = 0, male = 1) entered in Step 1, scenario type (coded as money scenario = 0, car scenario = 1) entered in Step 2, the respective NARQ uniqueness score entered in Step 3, conciliatory gestures (dummy-coded with “no conciliatory gesture” as the reference category) entered in Step 4, and the respective two-way interactions between the NARQ uniqueness score and conciliatory gestures entered in Step 5. Table 7 details summary statistics (see Appendix C for complete model statistics).

--- Table 7 about here ---

Consistent with Study 2, rivalry (but not admiration) uniquely predicted more revenge and avoidance versus benevolence above and beyond covariates (i.e., age, gender, and scenario type). Moreover, conciliatory gestures were unique predictors of forgiveness above and beyond covariates and admiration/rivalry: Verbal expressions of regret with nonverbal displays of negative perpetrator feelings (verbal, nonverbal) as well as a combination of the latter with offers of compensation (verbal, nonverbal, costly) were unique predictors of less revenge and avoidance versus benevolence. Verbal expressions of regret alone were only predictive of less revenge but not avoidance versus benevolence. Finally, verbal expressions of regret with offers of compensation (verbal, costly) as such were not predictive of forgiveness. With regard to the interactions between conciliatory behaviors and narcissism facets, results were as follows: No interaction between admiration and all four conciliatory gestures emerged, revealing that any particular form of conciliatory gesture did not moderate the effects of admiration on forgiveness. Similarly, rivalry did not interact with any particular form of conciliatory gesture in predicting revenge. However, regarding avoidance versus benevolence, a significant two-way interaction between rivalry and verbal expressions of regret with offers of compensation emerged (Table 7; last column, third row from below). We tested simple slopes at values one standard deviation above and below the mean of rivalry to further explore this interaction. As Figure 1 illustrates, among individuals low in rivalry (-1 *SD*), verbal expressions of regret with offers of compensation did not moderate the tendency for avoidance versus benevolence ($\beta[SE] = .31 [.21], p = .136$). They did, however, among individuals high in rivalry ($+1$ *SD*); verbal expressions of regret with offers of compensation were associated with less avoidance versus benevolence, a tendency that was marginally significant ($\beta[SE] = -.35 [.19], p = .066$). That is, when the transgressor asked for pardon and made a costly apology, individuals high (but not low) in rivalry tended to be less motivated toward avoidance versus benevolence.

General discussion

To investigate how apology moderates the link between narcissism and lack of forgiveness, we conducted three studies — one correlational and two experimental ones. Across these studies, drawing on three diverse samples, we obtained the following main results: First, narcissistic rivalry (but less so admiration) was positively related to lack of forgiveness (i.e., greater revenge and avoidance versus benevolence). Second, apology was negatively related to lack of forgiveness (above and beyond rivalry and admiration, respectively). Third, the link between rivalry and avoidance versus benevolence (but not revenge) was attenuated when a costly apology was offered to the victim, in addition to a verbal expression of regret.

In line with prior work illustrating that facets of narcissism can have distinct social consequences (Back et al., 2013; Fatfouta et al., 2015), we found that it is the antagonistic (i.e., rivalry) rather than the agentic (i.e., admiration) facet of narcissism that was most consistently associated with lack of forgiveness in response to real-life (Study 1) and hypothetical (Studies 2 and 3) transgressions. Importantly, results of Study 3 conceptually replicated Exline and colleagues' (2004) finding that entitled narcissists make forgiveness contingent on concessions. Extending these findings, the results of our moderation analyses provided the first evidence that these concessions need to be costly in order to appease individuals high in rivalry. In particular, the tendency to avoid the transgressor could only be mitigated when a verbal expression of regret (i.e., saying one is sorry) was accompanied by a specific compensation for the individual (in the case of the present study, an invitation to dinner). Interestingly, the motivation to retaliate against the transgressor in those high in narcissistic rivalry remained virtually unaffected.

Our results suggest that individuals high in rivalry may experience an unmitigated motivation to seek revenge, regardless of whether and *how* the transgressor conveyed the

apology. At first glance and in line with previous work relating rivalry to social conflict (Back et al., 2013) this tendency might seem detrimental to relationship functioning of those high in rivalry. Revenge, however, must not necessarily be maladaptive: It can also be considered a functional response to disrespectful behavior (Fitness & Peterson, 2008; Gollwitzer, 2009) as it holds the potential to deter similar transgressions in the future (McCullough et al., 2013). In particular, revenge can communicate an important message to the transgressor ("don't mess with me"; e.g., Funk, McGeer, & Gollwitzer, 2014; Gollwitzer, Meder, & Schmitt, 2011). In this light, individuals high in rivalry may strategically pursue two revenge-related goals; namely, (1) to actively prevent others from exploiting the self and (2) to induce others to acknowledge one's entitlement for better treatment (McCullough et al., 2011, 2013).

The mitigating effect of costly apology on the link between narcissistic rivalry and lack of forgiveness (i.e., avoidance versus benevolence) is intriguing. On the one hand, it seems possible to appease individuals high in rivalry, in a way that they become more motivated toward approach motivation (rather than avoidance motivation) regarding the transgressor. Given that costly apology has been found to convey more sincerity (Ohtsubo & Watanabe, 2009), this costliness could signal that the transgressor holds benign intentions (Kim et al., 2004) and thus could directly serve a cue indicating similar transgressions to be unlikely to reoccur (McCullough et al., 2014). On the other hand, an alternative possibility would be that individuals high in rivalry might be willing to re-approach the transgressor because his/her costly apology signals potential resources, and hence, a desirable quality as a social partner. These resources must not be of a financial nature per se (e.g., invitation to a prestigious restaurant), but may also be of a psychological nature (e.g., commitment to the relationship, willingness to sacrifice). Future research should try to tease apart these possibilities.

From a transgressor's perspective, our results may imply that it could be particularly costly (in the true sense of the word) to seek forgiveness from an antagonistic narcissist. Not only would one have to invest resources (be it financial or psychological), but rather one should be prepared to still find the individual high in rivalry willing to get even.

From a victim's perspective, it is possible that narcissists' approach motivation toward the transgressor could serve as a means of recalibrating their relationship. For example, those high in rivalry might take the chance to impose restrictions within an ongoing relationship ("I forgive you as long as you don't do it again"; Gerlach et al., 2012b). In future research, it might be worthwhile to more closely examine narcissists' motivation to get along with the transgressor as well as the communicative goals pursued by those high in rivalry.

A noteworthy contribution of our research is that we differentiated between two distinct dimensions of narcissism (i.e., admiration and rivalry) to better understand its social consequences in the wake of interpersonal hurt. In this respect, a particular strength of the present research is that we employed the NARQ (Back et al., 2013) instead of the NPI (Raskin & Hall, 1979), which has been criticized for conflating various dimensions of personality into one composite and psychometric shortcomings, in particular with regard to its entitlement facet (Ackerman, Donnellan, Roberts, & Fraley, 2015). Another strength is that we examined narcissist's lack of forgiveness and the moderating role of apology therein using both correlational and experimental designs. The latter afforded more control through standardization of the transgression incidents and, further, helped to clarify the directionality of associations in the proposed moderation model.

To note, while the present research demonstrated for the first time that costly apology may be needed to appease individuals high in rivalry, it is mute to the psychological mechanism(s) underlying this moderation effect. In future research, it might therefore be worthwhile to scrutinize potential mediators of this effect, such as perceived sincerity of the

costly apology or the attribution of more benevolent intentions on the side of the transgressor. Given that individuals high in rivalry (but not admiration) have previously been found to be highly mistrustful of others (Back et al., 2013), changes in these attributions might plausibly underlie the observed mitigating effect of costly apologies.

Finally, though we piloted the conciliatory gestures regarding their credibility of expressed regret, it is possible that those high in rivalry perceived certain gestures as ambiguous. For example, a contrite look may either be interpreted as signaling the transgressor's remorse or as revealing the true defective nature of the transgression (for a similar account, see de Jong, Peters, & De Cremer, 2003). Thus, a nonverbal expression of regret might have created ambivalence regarding the intentionality of the transgression for those high in rivalry. Future work is necessary to understand how narcissists actually perceive and interpret differing conciliatory gestures.

To conclude, the present research sought to shed light on an important yet open question: Does apology hold the potential to attenuate narcissists' lack of forgiveness, that is, can the shrew be tamed? We suggest that this might well be possible – provided that one is willing to incur a personal cost. The finding that costly apology only moderated the relation between narcissistic rivalry and avoidance but not revenge further points to the importance of a differentiated examination of both, narcissism and forgiveness. We hope that the present studies as well as the insights gained from it may stimulate future research on these constructs.

References

- Ackerman, R. A., Donnellan, M. B., Roberts, B. W., & Fraley, R. C. (2015). The Effect of Response Format on the Psychometric Properties of the Narcissistic Personality Inventory: Consequences for Item Meaning and Factor Structure. *Assessment*.
- Ackerman, R. A., Witt, E. A., Donnellan, M. B., Trzesniewski, K. H., Robins, R. W., & Kashy, D. A. (2011). What Does the Narcissistic Personality Inventory Really Measure? *Assessment*, *18*(1), 67-87.
- Aiken, L. S., & West, S. G. (1991). *Multiple regression: Testing and interpreting interactions*. Thousand Oaks, CA US: Sage Publications, Inc.
- Back, M. D., Küfner, A. C. P., Dufner, M., Gerlach, T. M., Rauthmann, J. F., & Denissen, J. J. A. (2013). Narcissistic admiration and rivalry: Disentangling the bright and dark sides of narcissism. *Journal of Personality and Social Psychology*, *105*(6), 1013-1037. doi: 10.1037/a0034431
- Berry, J. W., Worthington, E. L., O'Connor, L. E., Parrott, L., & Wade, N. G. (2005). Forgivingness, Vengeful Rumination, and Affective Traits. *Journal of Personality*, *73*(1), 183-226. doi: 10.1111/j.1467-6494.2004.00308.x
- Bottom, W. P., Gibson, K., Daniels, S. E., & Murnighan, J. K. (2002). When talk is not cheap: Substantive penance and expressions of intent in rebuilding cooperation. *Organization Science*, *13*(5), 497-513. doi: 10.1287/orsc.13.5.497.7816
- Bradlee, P. M., & Emmons, R. A. (1992). Locating narcissism within the interpersonal circumplex and the Five-Factor model. *Personality and Individual Differences*, *13*(7), 821-830. doi: 10.1016/0191-8869(92)90056-U
- Brown, R. P. (2003). Measuring individual differences in the tendency to forgive: Construct validity and links with depression. *Personality and Social Psychology Bulletin*, *29*(6), 759-771. doi: 10.1177/0146167203029006008

- Brown, R. P. (2004). Vengeance is mine: Narcissism, vengeance, and the tendency to forgive. *Journal of Research in Personality, 38*(6), 576-584. doi: 10.1016/j.jrp.2003.10.003
- Brown, R. P., & Phillips, A. (2005). Letting bygones be bygones: further evidence for the validity of the Tendency to Forgive scale. *Personality & Individual Differences, 38*(3), 627-638. doi: 10.1016/j.paid.2004.05.017
- Campbell, W. K., Bonacci, A. M., Shelton, J., Exline, J. J., & Bushman, B. J. (2004). Psychological Entitlement: Interpersonal Consequences and Validation of a Self-Report Measure. *Journal of Personality Assessment, 83*(1), 29-45.
- de Jong, P. J., Peters, M. L., & De Cremer, D. (2003). Blushing May Signify Guilt: Revealing Effects of Blushing in Ambiguous Social Situations. *Motivation & Emotion, 27*(3), 225-249.
- Eaton, J., Struthers, C. W., & Santelli, A. G. (2006a). Dispositional and state forgiveness: The role of self-esteem, need for structure, and narcissism. *Personality and Individual Differences, 41*(2), 371-380. doi: 10.1016/j.paid.2006.02.005
- Eaton, J., Struthers, C. W., & Santelli, A. G. (2006b). The Mediating Role of Perceptual Validation in the Repentance-Forgiveness Process. *Personality and Social Psychology Bulletin, 32*(10), 1389-1401. doi: 10.1177/0146167206291005
- Emmons, R. A. (2000). Personality and forgiveness. In M. E. McCullough, K. I. Pargament & C. E. Thoresen (Eds.), *Forgiveness: Theory, research, and practice*. (pp. 156-175). New York, NY US: Guilford Press.
- Exline, J. J., Baumeister, R. F., Bushman, B. J., Campbell, W. K., & Finkel, E. J. (2004). Too Proud to Let Go: Narcissistic Entitlement as a Barrier to Forgiveness. *Journal of Personality and Social Psychology, 87*(6), 894-912. doi: 10.1037/0022-3514.87.6.894
- Farrell, J. (1987). Cheap talk, coordination, and entry. *RAND Journal of Economics (RAND Journal of Economics), 18*(1), 34-39.

- Fatfouta, R. (in press). How forgiveness affects processing time: Mediation by rumination about the transgression. *Personality & Individual Differences*. doi: 10.1016/j.paid.2015.03.016
- Fatfouta, R., Gerlach, T. M., Schröder-Abé, M., & Merkl, A. (2015). Narcissism and lack of interpersonal forgiveness: The mediating role of state anger, state rumination, and state empathy. *Personality and Individual Differences*, 75, 36-40. doi: 10.1016/j.paid.2014.10.051
- Fehr, R., Gelfand, M. J., & Nag, M. (2010). The road to forgiveness: A meta-analytic synthesis of its situational and dispositional correlates. *Psychological Bulletin*, 136(5), 894-914. doi: 10.1037/a0019993
- Fincham, F. D. (2000). The kiss of the porcupines: From attributing responsibility to forgiving. *Personal Relationships*, 7(1), 1-23. doi: 10.1111/j.1475-6811.2000.tb00001.x
- Fincham, F. D., & Beach, S. R. H. (2002). Forgiveness in Marriage: Implications for Psychological Aggression and Constructive Communication. *Personal Relationships*, 9(3), 239-251.
- Fincham, F. D., Beach, S. R. H., & Davila, J. (2004). Forgiveness and Conflict Resolution in Marriage. *Journal of Family Psychology*, 18(1), 72-81. doi: 10.1037/0893-3200.18.1.72
- Fincham, F. D., Hall, J. H., & Beach, S. R. H. (2005). 'Til lack of forgiveness doth us part: Forgiveness in marriage. In E. L. Worthington (Ed.), *Handbook of Forgiveness* (pp. 207–226). New York: Routledge.
- Fitness, J., & Peterson, J. (2008). Punishment and forgiveness in close relationships: An evolutionary, social-psychological perspective. In J. P. Forgas, J. Fitness, J. P. Forgas

- & J. Fitness (Eds.), *Social relationships: Cognitive, affective, and motivational processes*. (pp. 255-269). New York, NY, US: Psychology Press.
- Fox, J. (2008). *Applied regression analysis and generalized linear models (2nd ed.)*. Thousand Oaks, CA, US: Sage Publications, Inc.
- Funk, F., McGeer, V., & Gollwitzer, M. (2014). Get the message: Punishment is satisfying if the transgressor responds to its communicative intent. *Personality and Social Psychology Bulletin*, 40(8), 986-997. doi: 10.1177/0146167214533130
- Gerlach, T. M., Agroskin, D., & Denissen, J. J. A. (2012a). Forgiveness in close interpersonal relationships: A negotiation approach. In E. Kals & J. Maes (Eds.), *Justice and conflicts. Theoretical and empirical contributions* (pp. 377-390). Berlin: Springer.
- Gerlach, T. M., Allemand, M., Agroskin, D., & Denissen, J. J. A. (2012b). Justice Sensitivity and Forgiveness in Close Interpersonal Relationships: The Mediating Role of Mistrustful, Legitimizing, and Pro-Relationship Cognitions. *Journal of Personality*, 80(5), 1373-1413. doi: 10.1111/j.1467-6494.2012.00762.x
- Gollwitzer, M. (2009). Justice and revenge. In M. E. Oswald, S. Bieneck & J. Hupfeld-Heinemann (Eds.), *Social psychology of punishment of crime* (pp. 137-156). Chichester: Wiley Blackwell.
- Gollwitzer, M., Meder, M., & Schmitt, M. (2011). What gives victims satisfaction when they seek revenge? *European Journal of Social Psychology*, 41(3), 364-374. doi: 10.1002/ejsp.782
- Grijalva, E., Newman, D. A., Tay, L., Donnellan, M. B., Harms, P. D., Robins, R. W., & Yan, T. (2014). Gender Differences in Narcissism: A Meta-Analytic Review. *Psychological Bulletin*. doi: 10.1037/a0038231

- Karremans, J. C., & Van Lange, P. M. (2004). Back to caring after being hurt: the role of forgiveness. *European Journal of Social Psychology, 34*(2), 207-227. doi: 10.1002/ejsp.192
- Kelley, D. L., & Waldron, V. R. (2005). An Investigation of Forgiveness-seeking Communication and Relational Outcomes. *Communication Quarterly, 53*(3), 339-358. doi: 10.1080/01463370500101097
- Kim, P. H., Ferrin, D. L., Cooper, C. D., & Dirks, K. T. (2004). Removing the Shadow of Suspicion: The Effects of Apology Versus Denial for Repairing Competence- Versus Integrity-Based Trust Violations. *Journal of Applied Psychology, 89*(1), 104-118. doi: 10.1037/0021-9010.89.1.104
- Küfner, A. C. P., Nestler, S., & Back, M. D. (2013). The Two Pathways to Being an (Un-)Popular Narcissist. *Journal of Personality, 81*(2), 184-195. doi: 10.1111/j.1467-6494.2012.00795.x
- McCullough, M. E., Fincham, F. D., & Tsang, J.-A. (2003). Forgiveness, forbearance, and time: The temporal unfolding of transgression-related interpersonal motivations. *Journal of Personality and Social Psychology, 84*(3), 540-557. doi: 10.1037/0022-3514.84.3.540
- McCullough, M. E., & Hoyt, W. T. (2002). Transgression-related motivational dispositions: Personality substrates of forgiveness and their links to the Big Five. *Personality and Social Psychology Bulletin, 28*(11), 1556-1573. doi: 10.1177/014616702237583
- McCullough, M. E., Kurzban, R., & Tabak, B. A. (2011). Evolved mechanisms for revenge and forgiveness. In P. R. Shaver & M. Mikulincer (Eds.), *Human aggression and violence: Causes, manifestations, and consequences*. (pp. 221-239). Washington, DC US: American Psychological Association.

- McCullough, M. E., Kurzban, R., & Tabak, B. A. (2013). Cognitive systems for revenge and forgiveness. *Behavioral and Brain Sciences*, *36*(1), 1-15. doi: 10.1017/S0140525X11002160
- McCullough, M. E., Pedersen, E. J., Tabak, B. A., & Carter, E. C. (2014). Conciliatory gestures promote forgiveness and reduce anger in humans. *PNAS Proceedings of the National Academy of Sciences of the United States of America*, *111*(30), 11211-11216. doi: 10.1073/pnas.1405072111
- McCullough, M. E., Rachal, K. C., Sandage, S. J., Worthington, E. L., Jr., Brown, S. W., & Hight, T. L. (1998). Interpersonal forgiving in close relationships: II. Theoretical elaboration and measurement. *Journal of Personality and Social Psychology*, *75*(6), 1586-1603. doi: 10.1037/0022-3514.75.6.1586
- McCullough, M. E., Root, L. M., & Cohen, A. D. (2006). Writing about the benefits of an interpersonal transgression facilitates forgiveness. *Journal of Consulting and Clinical Psychology*, *74*(5), 887-897. doi: 10.1037/0022-006x.74.5.887
- McCullough, M. E., Worthington, E. L., Jr., & Rachal, K. C. (1997). Interpersonal forgiving in close relationships. *Journal of Personality and Social Psychology*, *73*(2), 321-336. doi: 10.1037/0022-3514.73.2.321
- Miller, A. J., Worthington, E. L., & McDaniel, M. A. (2008). Gender and forgiveness: A meta-analytic review and research agenda. *Journal of Social & Clinical Psychology*, *27*(8), 843-876.
- Ohbuchi, K.-i., Kameda, M., & Agarie, N. (1989). Apology as Aggression Control: Its Role in Mediating Appraisal of and Response to Harm. *Journal of Personality & Social Psychology*, *56*(2), 219-227.
- Ohtsubo, Y., & Watanabe, E. (2009). Do sincere apologies need to be costly? Test of a costly signaling model of apology. *Evolution and Human Behavior*, *30*(2), 114-123.

- Paulhus, D. L. (1998). Interpersonal and Intrapyschic Adaptiveness of Trait Self-Enhancement: A Mixed Blessing? *Journal of Personality & Social Psychology*, *74*(5), 1197-1208.
- Paulhus, D. L. (2001). Normal narcissism: Two minimalist accounts. *Psychological Inquiry*, *12*(4), 228-230.
- Raskin, R., & Hall, C. S. (1979). A narcissistic personality inventory. *Psychological Reports*, *45*(2), 590-590.
- Rhodewalt, F., & Peterson, B. (2009). Narcissism. In M. R. Leary & R. H. Hoyle (Eds.), *Handbook of individual differences in social behavior*. (pp. 547-560). New York, NY US: Guilford Press.
- Schlenker, B. R., & Darby, B. W. (1981). The use of apologies in social predicaments. *Social Psychology Quarterly*, *44*(3), 271-278. doi: 10.2307/3033840
- Schmitt, M., Gollwitzer, M., Förster, N., & Montada, L. (2004). Effects of Objective and Subjective Account Components on Forgiving. *Journal of Social Psychology*, *144*(5), 465-485.
- Tabak, B. A., McCullough, M. E., Luna, L. R., Bono, G., & Berry, J. W. (2012). Conciliatory Gestures Facilitate Forgiveness and Feelings of Friendship by Making Transgressors Appear More Agreeable. *Journal of Personality*, *80*(2), 503-536. doi: 10.1111/j.1467-6494.2011.00728.x
- Waldron, V. R., & Kelley, D. L. (2008). *Communicating forgiveness*. Thousand Oaks, CA, US: Sage Publications, Inc.
- Worthington, E. L., & Scherer, M. (2004). Forgiveness is an emotion-focused coping strategy that can reduce health risks and promote health resilience: theory, review, and hypotheses. *Psychology & Health*, *19*(3), 385-405. doi: 10.1080/0887044042000196674

Table 1

Study 1: Intercorrelations, descriptive statistics, and reliability for all measures

<i>Measure</i>	1	2	3	4	5	6
1. Admiration	—					
2. Rivalry	.33***/- .33***	—				
3. Revenge	.05*/-.01	.22***/.21***	—			
4. Avoidance	-.02/-.06*	.13***/.14***	.40***	—		
5. Benevolence	.06*/.09***	-.11***/-.14***	-.32***	-.38***	—	
6. Apology	.04/.10**	-.14***/-.17***	-.12***	-.19***	.17***	—
<i>M</i>	3.11	2.23	1.39	2.15	5.58	3.54
<i>SD</i>	.87	.77	.64	1.23	1.16	1.17

Note. $N = 767$. All correlations are Pearson's r , except for those involving forgiveness facets (i.e., revenge, avoidance, and benevolence), which are Kendall's τ . For admiration and rivalry, uniqueness correlations (i.e., the amount of each facet that is not shared by the other one) are shown next to the respective zero-order correlations with the other measures.

* $p < .05$. ** $p < .01$. *** $p < .001$ (2-tailed).

Table 2

Study 1: Summary of hierarchical regression analyses

Variable	Models for admiration			Models for rivalry		
	Revenge	Avoidance	Benevolence	Revenge	Avoidance	Benevolence
<i>Step 1:</i>						
Constant	-.01	.02	.04	-.01	.02	.04
Age	.02	.03	-.01	.02	.03	-.01
Gender	.06	-.07	-.16	.06	-.07	-.16
R^2	.001	.002	.005	.001	.002	.005
<i>Step 2:</i>						
Narcissism ^a	-.02	-.11**	.15****	.31****	.21****	-.16****
ΔR^2	.000	.012***	.022****	.093****	.043****	.026****
<i>Step 3:</i>						
Apology	-.13****	-.23****	.22****	-.08*	-.21****	.22****
ΔR^2	.017****	.052****	.050****	.007*	.043****	.045****
<i>Step 4:</i>						
Narcissism•Apology	-.06	.02	-.03	-.04	-.03	-.03
ΔR^2	.003	.000	.001	.002	.001	.001

Note. $N=767$. Bootstrapped standardized coefficients (β) based on 1,000 samples are presented. Gender was coded as female = 0, male = 1.

^a NARQ uniqueness scores (i.e., the amount of variance of each narcissism facet that is not shared by the other one).

* $p < .05$. ** $p < .01$. *** $p < .005$. **** $p < .001$ (2-tailed).

Table 3

Study 2: Intercorrelations, descriptive statistics, and reliability for all measures

<i>Measure</i>	1	2	3	4
1. Admiration	—			
2. Rivalry	.31***/-.31***	—		
3. Revenge	.14***/.07**	.27***/.23***	—	
4. Avoidance vs. Benevolence	-.02/-.09*	.19***/.21***	.20***	—
<i>M</i>	2.98	2.29	1.85	3.47
<i>SD</i>	1.08	1.04	0.96	1.24

Note. $N = 930$. All correlations are Pearson's r , except for those involving revenge, which are Kendall's τ . For admiration and rivalry, uniqueness correlations (i.e., the amount of each facet that is not shared by the other one) are shown next to the respective zero-order correlations with the other measures.

* $p < .05$. ** $p < .005$. *** $p < .001$ (2-tailed).

Table 4

Study 2: Summary of hierarchical regression analyses

Variable	Models for admiration				Models for rivalry			
	Revenge ^a		Avoidance vs. Benevolence		Revenge ^a		Avoidance vs. Benevolence	
	ΔR^2	β	ΔR^2	β	ΔR^2	β	ΔR^2	β
<i>Step 1:</i>	.013***		.018****		.013***		.018****	
Constant		-.07		.08*		-.07		.08*
Age		-.00		.04		-.00		.04
Gender		.26****		-.30****		.26***		-.30****
<i>Step 2:</i>	.007*		.126****		.007*		.126****	
Scenario “gossip”		-.06		.14		-.06		.14
Scenario “gift”		-.20**		-.67****		-.20*		-.67****
<i>Step 3:</i>	.004*		.007***		.115****		.060****	
Narcissism ^b		.06		-.09***		.35****		.25****
<i>Step 4:</i>	.017****		.015****		.008***		.010****	
Apology		-.26****		-.24****		-.19***		-.20****
<i>Step 5:</i>	.001		.001		.003		.003	
Narcissism•Apology		-.08		-.05		-.12		.10

Note. $N = 930$. Gender was coded as female = 0, male = 1; for scenario type, “publicly disclosed a personal detail of the participant” was used as the reference category; apology was coded as no apology = 0, apology = 1.

^a Bootstrapped standardized coefficients (β) based on 1,000 samples.

^b NARQ uniqueness scores (i.e., the amount of variance of each narcissism facet that is not shared by the other one).

* $p < .05$. ** $p < .01$. *** $p < .005$. **** $p < .001$ (2-tailed).

Table 5

Study 3: Post-transgression conciliatory gestures

Gesture	Condition	Phrasing
1	No expression of regret	He/she tries to simply pass over the situation by ignoring the issue when it comes up in a later conversation.
2	Verbal expression of regret	He/she says she is sorry and acknowledges that his/her behavior was inadequate.
3	Verbal expression of regret and offer of compensation	He/she says she is sorry and acknowledges that his/her behavior was inadequate. He/she spontaneously invites you to dinner, for which he/she cancels another meeting.
4	Verbal and nonverbal expression of regret	He/she gives you a contrite look and looks sad. He/she says he/she is sorry and acknowledges that his/her behavior was inadequate.
5	Verbal and nonverbal expression of regret and offer of compensation	He/she gives you a contrite look and looks sad. He/she says he/she is sorry and acknowledges that his/her behavior was inadequate. He/she spontaneously invites you to dinner, for which he/she cancels another meeting.

Note. Conciliatory gestures were ranked ($N=36$, $M_{age} = 31.73$, $SD = 8.50$; 69.4% female) in ascending order of credibility of expressed regret from 1 (*least credible*) to 5 (*most credible*).

Table 6

Study 3: Intercorrelations, descriptive statistics, and reliability for all measures

<i>Measure</i>	1	2	3	4
1. Admiration	—			
2. Rivalry	.28***/-.28***	—		
3. Revenge	.13***/.06	.28***/.24***	—	
4. Avoidance vs. Benevolence	.02/-.05	.22***/.25***	.34***	—
<i>M</i>	3.18	2.3	1.71	2.41
<i>SD</i>	.86	.81	.78	.99

Note. $N = 430$. All correlations are Pearson's r , except for correlations involving revenge, which are Kendall's τ . For admiration and rivalry, uniqueness correlations (i.e., the amount of each facet that is not shared by the other one) are shown next to the respective zero-order correlations with the other measures.

* $p < .05$. ** $p < .005$. *** $p < .001$ (2-tailed).

Table 7

Study 3: Summary of hierarchical regression analyses

Variable	Models for admiration				Models for rivalry			
	Revenge ^a		Avoidance vs. Benevolence		Revenge ^a		Avoidance vs. Benevolence	
	ΔR^2	β	ΔR^2	β	ΔR^2	β	ΔR^2	β
<i>Step 1:</i>	.014*		.012		.014*		.012	
Constant		-.08		.08		-.08		.08
Age		-.04		.01		-.04		.01
Gender		.25*		-.23*		.25*		-.23*
<i>Step 2:</i>	.010*		.007		.010*		.007	
Scenario type		.20*		.17		.20*		.17
<i>Step 3:</i>	.006		.001		.116*****		.055*****	
Narcissism ^b		.08		-.03		.34*****		.24*****
<i>Step 4:</i>	.040***		.022*		.029**		.016	
Verbal only (VO)		-.47**		-.28		-.40**		-.23
Verbal, costly (VC)		-.25		-.05		-.25		-.06
Verbal, nonverbal (VN)		-.54***		-.33*		-.46***		-.28†
Verbal, nonverbal, costly (VNC)		-.47***		-.36*		-.42***		-.32*
<i>Step 5:</i>	.007		.006		.010		.014	
Narcissism•VO		-.05		.15		-.18		-.25
Narcissism•VC		.21		.09		-.28		-.33*
Narcissism•VN		-.01		.18		-.12		-.22
Narcissism•VNC		.03		.21		-.20		-.20

Note. $N = 430$. Gender was coded as female = 0, male = 1; scenario type was coded as money scenario = 0, car scenario = 1; for conciliatory gestures, “no conciliatory gesture” was used as the reference category; † $p = .055$. * $p < .05$. ** $p < .01$, *** $p < .005$. ***** $p < .001$ (2-tailed).

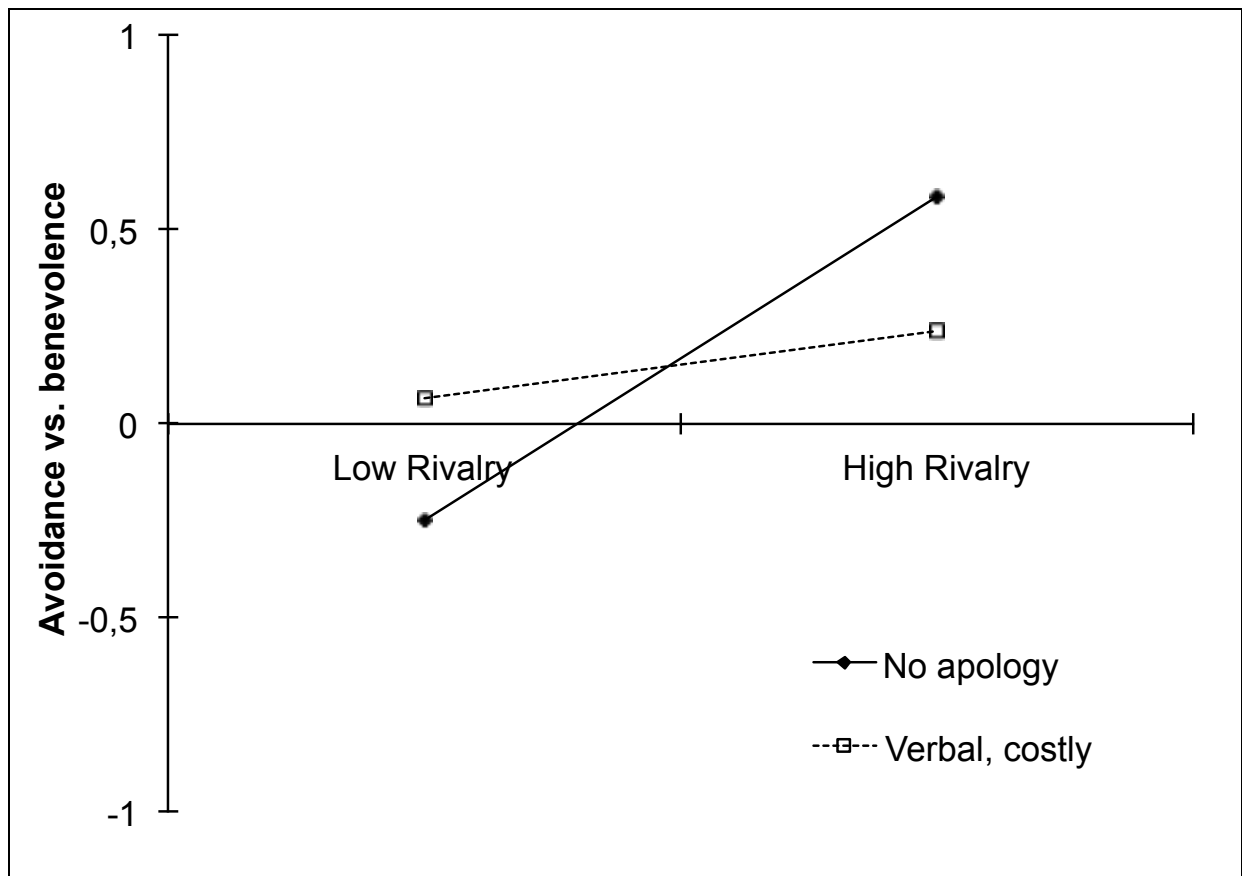
^a Bootstrapped standardized coefficients (β) based on 1,000 samples.

^b NARQ uniqueness scores (i.e., the amount of variance of each narcissism facet that is not shared by the other one).

Figure captions.

Figure 1. Predicted values for avoidance versus benevolence as a function of narcissistic rivalry (± 1 *SD* of the mean) and apology.

Figure 1



NARCISSISM AND APOLOGY

Appendix A

Study 1: Hierarchical regression analyses

Variable	Models for admiration			Models for rivalry		
	Revenge	Avoidance	Benevolence	Revenge	Avoidance	Benevolence
<i>Step 1:</i>						
Constant	-.01	.02	.04	-.01	.02	.04
Age	.02	.03	-.01	.02	.03	-.01
Gender	.06	-.07	-.16	.06	-.07	-.16
R^2	.001	.002	.005	.001	.002	.005
<i>Step 2:</i>						
Constant	-.02	.01	.04	.01	.03 (.04)	.02
Age	.02	.03	-.01	.04	.04 (.04)	-.02
Gender	.06	-.05	-.19*	-.05	-.15 (.08)	-.10
Narcissism ^a	-.02	-.11**	.15****	.31****	.21**** (.04)	-.16****
ΔR^2	.000	.012***	.022****	.093****	.043****	.026****
<i>Step 3:</i>						
Constant	-.01	.02	.04	.01	.04 (.04)	.02
Age	.02	.02	-.00	.04	.03 (.04)	-.01
Gender	.05	-.07	-.17	-.05	-.15 (.08)	-.10
Narcissism ^a	-.00	-.09*	.13***	.30****	.18**** (.04)	-.13***
Apology	-.13****	-.23****	.22****	-.08*	-.21**** (.04)	.22****
ΔR^2	.017****	.052****	.050****	.007*	.043****	.045****
<i>Step 4:</i>						
Constant	-.01	.02	.04	.01	.03 (.04)	.02
Age	.01	.02	-.01	.04	.03 (.04)	-.01
Gender	.06	-.08	-.16	-.05	-.15 (.08)	-.10
Narcissism ^a	-.00	-.09*	.13***	.29****	.17**** (.04)	-.13***
Apology	-.13****	-.23****	.22****	-.08*	-.21**** (.04)	.22****
Narcissism•Apology	-.06	.02	-.03	-.04	-.03 (.04)	-.03
ΔR^2	.003	.000	.001	.002	.001	.001

Note. $N=767$. Bootstrapped standardized coefficients (β) based on 1,000 samples are presented. Gender was coded as female = 0, male = 1.

^aNARQ uniqueness scores (i.e., the amount of variance of each narcissism facet that is not shared by the other one).

* $p < .05$. ** $p < .01$. *** $p < .005$. **** $p < .001$ (2-tailed).

NARCISSISM AND APOLOGY

Appendix B

Study 2: Hierarchical regression analyses

Variable	Models for admiration				Models for rivalry			
	Revenge ^a		Avoidance vs. Benevolence		Revenge ^a		Avoidance vs. Benevolence	
	ΔR^2	β	ΔR^2	β	ΔR^2	ΔR^2	β	
<i>Step 1:</i>	.013***		.018****		.013***	.018****		
Constant		-.07		.08*		-.07	.08*	
Age		-.00		.04		-.00	.04	
Gender		.26****		-.30****		.26***	-.30****	
<i>Step 2:</i>	.007*		.126****		.007*	.126****		
Constant		.02		.26****		.02	.25****	
Age		.00		.05		.00	.05	
Gender		.27****		-.26****		.27***	-.26****	
Scenario “gossip”		-.06		.14		-.06	.14	
Scenario “gift”		-.20**		-.67****		-.20*	-.67****	
<i>Step 3:</i>	.004*		.007***		.115****	.060****		
Constant		.02		.38****		.05	.27****	
Age		.00		.05		.05	.08**	
Gender		.27****		-.26****		.15	-.35****	
Scenario “gossip”		-.06		.14		-.03	.16*	
Scenario “gift”		-.21**		-.67****		-.21**	-.67****	
Narcissism ^b		.06		-.09***		.35****	.25****	
<i>Step 4:</i>	.017****		.015****		.008***	.010****		
Constant		.16*		.38****		.14*	.37****	
Age		.00		.05		.05	.08**	
Gender		.27****		-.26****		.15*	-.34****	
Scenario “gossip”		-.07		.14		-.04	.15*	
Scenario “gift”		-.21**		-.67****		-.21**	-.68****	
Narcissism ^b		.07*		-.08**		.34****	.24****	
Apology		-.26****		-.24****		-.19***	-.20****	
<i>Step 5:</i>	.001		.001		.003	.003		
Constant		.16*		.38****		.14*	.38****	
Age		.00		.05		.05	.08**	
Gender		.26****		-.26****		.16*	-.35****	
Scenario “gossip”		-.07		.14		-.04	.15*	
Scenario “gift”		-.22**		-.67****		-.22***	-.67****	
Narcissism ^b		.11		-.05		.39****	.19****	
Apology		-.26****		-.24****		-.19***	-.20****	
Narcissism•Apology		-.08		-.05		-.12	.10	

Note. $N=930$. Gender was coded as female = 0, male = 1; for scenario type, “publicly disclosed a personal detail of the participant” was used as the reference category; apology was coded as no apology = 0, apology = 1.

^a Bootstrapped standardized coefficients (β) based on 1,000 samples.

^b NARQ uniqueness scores (i.e., the amount of variance of each narcissism facet that is not shared by the other one).

* $p < .05$. ** $p < .01$. *** $p < .005$. **** $p < .001$ (2-tailed).

NARCISSISM AND APOLOGY

Appendix C

Study 3: Hierarchical regression analyses

Variable	Models for admiration				Models for rivalry			
	Revenge ^a		Avoidance vs. Benevolence		Revenge ^a		Avoidance vs. Benevolence	
	ΔR^2	β	ΔR^2	β	ΔR^2	β	ΔR^2	β
<i>Step 1:</i>	.014*		.012		.014*		.012	
Constant		-.08		.08		-.08		.08
Age		-.04		.01		-.04		.01
Gender		.25*		-.23*		.25*		-.23*
<i>Step 2:</i>	.010*		.007		.010*		.007	
Constant		-.18*		-.00		-.18*		-.00
Age		-.04		.02		-.04		.02
Gender		.25*		-.23*		.25*		-.23*
Scenario type		.20*		.17		.20*		.17
<i>Step 3:</i>	.006		.001		.116****		.055****	
Constant		-.17*		-.01		-.13		-.03
Age		-.05		.02		-.01		.04
Gender		.22*		-.22*		.16		-.30***
Scenario type		.19*		.17		.17		.17
Narcissism ^b		.08		-.03		.34****		.24****
<i>Step 4:</i>	.040***		.022*		.029**		.016	
Constant		.16		.19		.16		.20
Age		-.04		.02		-.01		.04
Gender		.23*		-.22*		.17		-.29***
Scenario type		.23*		.20*		.20*		.17
Narcissism ^b		.08		-.03		.33****		.23****
Verbal only (VO)		-.47**		-.28		-.40**		-.23
Verbal, costly (VC)		-.25		-.05		-.25		-.06
Verbal, nonverbal (VN)		-.54***		-.33*		-.46***		-.28†
Verbal, nonverbal, costly (VNC)		-.47***		-.36*		-.42***		-.32*
<i>Step 5:</i>	.007		.006		.010		.014	
Constant		.17		.18		.13		.17
Age		-.04		.02		-.01		.04
Gender		.22*		-.21*		.17		-.30***
Scenario type		.21*		.20*		.21*		.19*
Narcissism ^b		.05		-.16		.48****		.42****
Verbal only (VO)		-.47**		-.28		-.39*		-.22
Verbal, costly (VC)		-.25		-.05		-.22		-.02
Verbal, nonverbal (VN)		-.53***		-.34*		-.44**		-.27
Verbal, nonverbal, costly (VNC)		-.47***		-.35*		-.41**		-.30*
Narcissism•VO		-.05		.15		-.18		-.25
Narcissism•VC		.21		.09		-.28		-.33*
Narcissism•VN		-.01		.18		-.12		-.22
Narcissism•VNC		.03		.21		-.20		-.20

NARCISSISM AND APOLOGY

Note. $N=430$. Gender was coded as female = 0, male = 1; scenario type was coded as money scenario = 0, car scenario = 1; for conciliatory gestures, “no conciliatory gesture” was used as the reference category; † $p = .055$. * $p < .05$. ** $p < .01$, *** $p < .005$. **** $p < .001$ (2-tailed).

^a Bootstrapped standardized coefficients (β) based on 1,000 samples.

^b NARQ uniqueness scores (i.e., the amount of variance of each narcissism facet that is not shared by the other one).



ELSEVIER

Contents lists available at ScienceDirect

Psychiatry Research

journal homepage: www.elsevier.com/locate/psychres

Associations between obsessive–compulsive symptoms, revenge, and the perception of interpersonal transgressions



Ramzi Fatfouta^{a,b,*}, Angela Merkl^{b,c,d}

^a Affective Neuroscience and Psychology of Emotion, Department of Education and Psychology, Freie Universität Berlin, Habelschwerdter Allee 45, 14195 Berlin, Germany

^b Cluster of Excellence “Languages of Emotion”, Freie Universität Berlin, Germany

^c Department of Neurology with Experimental Neurology, Charité-Universitätsmedizin Berlin, Campus Virchow Klinikum, Germany

^d Department of Psychiatry, Charité-Universitätsmedizin Berlin, Campus Benjamin Franklin, Germany

ARTICLE INFO

Article history:

Received 17 December 2013

Received in revised form

8 April 2014

Accepted 20 May 2014

Available online 29 May 2014

Keywords:

Obsessive–compulsive disorder (OCD)

Obsessions

Revenge

Interpersonal transgressions

Latent aggression

ABSTRACT

Anger and aggression have only recently gained center stage in research on obsessive–compulsive disorder (OCD). An investigation of obsessive–compulsive (OC) symptoms focusing on the outcome of unresolved anger (i.e., revenge), however, is absent from the literature. The objective of the present research was therefore to provide a first step towards filling this gap and, hence, to systematically examine the associations between OC symptoms and different aspects of revenge (i.e., attitudes, dispositions, motivations). In three independent studies with nonclinical participants ($N=504$), we tested the hypothesis that OC symptoms relate to greater revenge. Individuals high in OC symptoms reported more positive attitudes toward revenge (Study 1), scored higher on a measure of trait revenge (Study 2), and reported increased revenge motivation regarding a real-life transgressor (Study 3). Furthermore, Study 4 ($N=175$) demonstrated that individuals high in OC symptoms perceived interpersonal transgressions more frequently in their daily lives. OC symptoms were positively related to the number of transgressions that respondents disclosed. Our results suggest that revenge and interpersonal hurt play a significant role in OCD.

© 2014 Elsevier Ireland Ltd. All rights reserved.

1. Introduction

In recent years, researchers have begun to explore how individuals with obsessive–compulsive disorder (OCD) deal with anger. Whiteside and Abramowitz (2004), for example, demonstrated that nonclinical individuals high in obsessive–compulsive (OC) symptoms experienced more anger, which however was inwardly suppressed (i.e., not openly displayed). Furthermore, relative to individuals low in OC symptoms, difficulties in controlling anger were present. Relatedly, two studies have evidenced that anger experience (but not anger expression) tends to be elevated in OCD (Moscovitch et al., 2008; Radomsky et al., 2007). These results cohere with recent findings that individuals with OCD score higher on measures of latent (covert) aggression (Moritz et al., 2011), as tapped with items such as “When a little misfortune has happened to friends/relatives, I feel a bit pleasure” (Moritz et al., 2009, p. 285).

* Corresponding author at: Affective Neuroscience and Psychology of Emotion, Department of Education and Psychology, Freie Universität Berlin, Habelschwerdter Allee 45, 14195 Berlin, Germany. Tel.: +49 30 838 57 857; fax: +49 30 838 55 778. E-mail address: ramzi.fatfouta@fu-berlin.de (R. Fatfouta).

Experiences of anger lead to a desire for revenge that does not cease until it is recognized and released (Fitzgibbons, 1986). However, as sketched above, individuals with OCD rarely give vent to their anger, which might perpetuate fantasies of revenge. Indeed, it has been suggested that symptomatic revenge fantasies (e.g., recurrent urges to get back on someone) are similar to intrusive thoughts (obsessions). That is, they are experienced as unwanted, uncontrollable, and anxiogenic (Horowitz, 2007). At present, despite its increasing importance for therapeutic interventions, the concept of “revenge” has been widely neglected in psychiatric research (Gäbler and Maercker, 2011; Horowitz, 2007). Specifically, it is unknown whether individuals high in OC symptoms (1) have more positive attitudes toward revenge, (2) have a more generalized propensity to react to personal offenses with revenge, and (3) experience higher motivations to seek revenge in response to an experienced hurt. The rationale of the present research was therefore to fill these gaps and, hence, to gain insight into the relations between OC symptoms and different aspects of revenge.

1.1. Anger, revenge, and negative interpersonal events in OCD

Anger and revenge are said to be related but distinct constructs (Wilkowski et al., 2012). Anger refers to a subjective, negative

emotional experience that varies in intensity, from mild irritation to fury and rage (Spielberger, 1988). Revenge, by contrast, pertains to individuals' beliefs about the morality and desirability of vengeful actions for attaining certain aims (e.g., restoring the moral balance) and their use as interpersonal problem-solving strategy (McCullough et al. 2001, p. 602). Hence, while in anger the focus is on experience, in revenge it is on motivation (Wilkowski et al., 2012). In this line, there is evidence to indicate that revenge is associated with greater motivations to retaliate against and avoid the perceived offender (McCullough et al., 2001). Furthermore, reductions in revenge and avoidance motivation over time are believed to underlie interpersonal forgiving (McCullough et al., 1998).

Might people high in OC symptoms experience higher levels of revenge? To our knowledge, this question has not been empirically examined. However, and consistent with our own clinical experience, some indirect evidence supports this contention. For example, Moritz et al. (2009) demonstrated that individuals with OCD exhibited increased latent aggression (i.e., disguised aggressive obsessions) compared to healthy controls. In line with psychodynamic views (Kempke and Luyten, 2007), latent aggression is concealed by erecting a benevolent façade, which in turn results in ambivalence towards others (Moritz et al., 2009, 2011). This suggestion resonates with survey data showing that interpersonal tensions are frequent in OCD (Hauschildt et al., 2010).

According to Moritz and colleagues, interpersonal conflicts in OCD may result from not obeying to the rules imposed by the individual (Moritz et al., 2012). As most individuals with OCD have high moral attitudes (Moritz et al., 2009), relational situations that compromise the maintenance of a rigid value system may induce anger, further fueling vengeful thoughts, images, and impulses. For example, a family member may not comply with OC-related routines (e.g., "I won't check the door lock for you again!"), treat the individual with OCD unjustly (e.g., "If you don't let me use the car, I'm going to rub my shoes all over your bed!"), or even mock him/her for adhering to idiosyncratic rules (cf. Purdon, 2011). In response to these negative interpersonal events, the angered OCD individual might wish to even the score. This is consistent with Hauschildt et al. (2010) who found that two-thirds of OCD patients fear becoming "mad" and half of them acknowledged behaving aggressively towards their partner.

2. Research overview: aims and predictions

Against this background, the aim of the present research was twofold: (a) to examine the relationships between OC symptoms and revenge and (b) to further explore the associations between OC symptoms and the perception of interpersonal transgressions in daily life. To address our first aim, we conducted three online studies, thereby drawing on three representative samples from various cities in Germany. In doing so, we pursued two goals: First, using independent samples from different cities allows for a heterogeneous distribution of participants' OC symptoms and personality. Second, using the internet for recruitment and data collection helps alleviate social desirability biases due to the sensitive nature of the investigated topics (Joinson, 1999; Rhodes et al., 2003). Importantly, we used different self-report measures, each assessing different aspects of revenge, to further increase the generalizability of results across studies. For Studies 1–3, it was hypothesized that OC symptoms would significantly predicts individual differences in revenge. To address the second aim, we designed a fourth study to provide initial empirical evidence that individuals with OCD perceive a greater frequency of negative interpersonal events in their daily social interactions. For Study 4,

it was hypothesized that OC symptoms would significantly predict individuals' perceptions of transgression occurrences.

3. Studies 1–3: revenge and OC symptoms

3.1. Method

3.1.1. Participants

A total of 504 participants (Study 1: $n=153$, $M_{age}=24.38$, $S.D.=5.52$, age range: 18–47; 72.5% female; Study 2: $n=178$, $M_{age}=25.39$, $S.D.=6.48$, age range: 18–50; 78.1% female; Study 3: $n=173$, $M_{age}=26.30$, $S.D.=6.75$, age range: 19–59; 82% female) voluntarily took part in an online study on "Conflict in everyday life". Participants were invited via mailing lists for psychology experiments hosted at different universities in Germany.

3.1.2. Procedure and measures

Invitation e-mails contained a link to the survey. After consenting to participate, participants provided basic demographic information (age, gender, academic subject) and completed a German version (Gonner et al., 2007) of the Obsessive-Compulsive Inventory-Revised (OCI-R; Foa et al., 2002). The OCI-R is an 18-item self-report scale assessing a wide range of OC symptoms with six subscales: washing, checking, ordering, obsessing, hoarding, and neutralizing. Participants rated each item on a 5-point scale (0 = not at all, 4 = extremely).

Next, participants completed different measures of revenge to allow for generalizability of results. In Study 1, participants completed the 20-item Vengeance Scale (Stuckless and Goranson, 1992; e.g., "It is important for me to get back at people who have hurt me"), which measures individual differences in revenge-seeking attitudes. Items were rated on a 7-point scale (1 = strongly disagree, 7 = strongly agree).

In Study 2, participants completed the four-item Tendency to Forgive Scale (TTF; Brown, 2003; e.g., "I tend to get over it quickly when someone hurts my feelings") measuring the extent to which individuals generally react to interpersonal offenses with forgiveness (i.e., trait forgiveness). To facilitate the focus on "revenge" in the present research, we reverse-scored two positively worded items so that higher scores indicated greater (trait) revenge. Participants responded to each item on a 7-point scale (1 = strongly disagree, 7 = strongly agree).

In Study 3, participants completed the 12-item Transgression-Related Interpersonal Motivations inventory (TRIM; McCullough et al., 1998). The TRIM is a situation-specific (i.e., state) measure of unforgiving motivations toward a real-life transgressor and is divided into two subscales, with five items measuring revenge motivation (e.g., "I'll make him/her pay") and seven items measuring avoidance motivation (e.g., "I am trying to keep as much distance between us as possible"). Participants were instructed to bring to mind a recent situation in which they felt hurt or treated unjustly by someone and to indicate how they felt about that person right now. Items were rated on a 5-point scale (1 = strongly disagree, 5 = strongly agree). Given that reliving an offense could induce a negative mood in participants (Karremans and Van Lange, 2008), we included a mood rating after the TRIM questionnaire. Participants rated their mood on a 5-point visual analog scale (1 = very sad, 5 = very happy).

3.2. Data analytic plan

Preliminary partial correlation analyses with participants' age and gender as potential control variables yielded almost identical results. For reasons of parsimony, these subject variables are not considered for the results reported below. For hypothesis testing, we analyzed bivariate correlations among the OCI-R and personality measures. Additionally, we computed stepwise multiple regression analyses to examine which subscale of the OCI-R uniquely predicts revenge (the entry and removal criteria were $P < 0.05$ and $P > 0.10$, respectively). All statistical tests adopted a significance level of $\alpha = 0.05$ (two-tailed).

4. Results and discussion

Table 1 details descriptive statistics, reliabilities, and intercorrelations by study. Cronbach's alpha coefficients were satisfactory, except for the neutralizing subscale of the OCI-R (α 's ranging from 0.41 to 0.57; see Table 1). This pattern is congruent with the low reliability of this subscale reported in previous research (Foa et al., 2002).

Our main hypothesis was that OC symptoms would be associated with greater revenge. In Study 1, the total OCI-R score and obsessing were significantly and positively related to revenge-seeking attitudes. When the OCI-R subscales were entered stepwise as predictors of revenge-seeking attitudes in a multiple regression analysis, obsessing was the sole significant predictor

Table 1
Descriptive statistics, reliabilities, and intercorrelations among variables in studies 1, 2, and 3.

Measure	M	S.D.	α	1	2	3	4	5	6	7	8	9	10
Study 1 ($n=153$)													
1. Revenge-seeking attitudes	2.85	0.75	0.86	–									
2. OCI-R washing	1.20	1.80	0.72	0.01	–								
3. OCI-R checking	2.45	2.22	0.71	0.141 [†]	0.46***	–							
4. OCI-R ordering	2.68	2.83	0.87	0.07	0.34***	0.58***	–						
5. OCI-R obsessing	2.57	2.85	0.86	0.19*	0.28***	0.42***	0.27***	–					
6. OCI-R hoarding	2.79	2.30	0.77	0.13	0.22**	0.29***	0.35***	0.37***	–				
7. OCI-R neutralizing	1.14	1.78	0.57	0.10	0.24**	0.55***	0.50***	0.42***	0.40***	–			
8. OCI-R total	12.82	9.62	0.88	0.16*	0.57***	0.79***	0.75***	0.69***	0.64***	0.72***	–		
Study 2 ($n=178$)													
1. Trait revenge	4.52	1.04	0.66	–									
2. OCI-R washing	1.29	2.13	0.84	0.18*	–								
3. OCI-R checking	2.34	2.23	0.72	0.04	0.31***	–							
4. OCI-R ordering	2.77	2.58	0.86	0.05	0.41***	0.39***	–						
5. OCI-R obsessing	2.28	2.06	0.72	0.24***	0.25***	0.36***	0.31***	–					
6. OCI-R hoarding	2.42	2.12	0.77	–0.00	0.19*	0.31***	0.27***	0.23**	–				
7. OCI-R neutralizing	0.95	1.37	0.43	0.10	0.32***	0.32***	0.34***	0.34***	0.23**	–			
8. OCI-R total	12.03	8.17	0.85	0.15*	0.64***	0.70***	0.74***	0.63***	0.58***	0.59***	–		
Study 3 ($n=173$)													
1. Revenge motivation	1.98	0.92	0.87	–									
2. Avoidance motivation	3.61	1.03	0.90	0.31***	–								
3. OCI-R washing	1.17	1.72	0.76	0.05	0.12	–							
4. OCI-R checking	2.22	2.20	0.79	0.28***	0.02	0.38***	–						
5. OCI-R ordering	2.74	2.54	0.86	0.18*	0.04	0.35***	0.38***	–					
6. OCI-R obsessing	2.69	2.10	0.76	0.151 [†]	0.141 [†]	0.23**	0.37***	0.25***	–				
7. OCI-R hoarding	2.92	2.03	0.74	0.10	–0.02	0.16*	0.23**	0.22**	0.17*	–			
8. OCI-R neutralizing	0.94	1.39	0.41	0.10	0.04	0.28***	0.26***	0.36***	0.29***	0.27***	–		
9. OCI-R total	12.68	7.62	0.84	0.23*	0.09	0.61***	0.71***	0.72***	0.61***	0.54***	0.59***	–	
10. Mood	3.58	0.94	n/a	–0.08	–0.10	–0.131 [†]	–0.11	–0.141 [†]	–0.25***	0.04	–0.13	–0.19*	–

Note. $N=504$. OCI-R=Obsessive–Compulsive Inventory–Revised.

[†] $P < 0.10$.

* $P < 0.05$.

** $P < 0.01$.

*** $P < 0.001$ (all P 's two-tailed).

($\beta=0.19$, $P=0.018$). The model explained 4% of the variance in participants' revenge-seeking attitudes, $F(1, 151)=5.73$, $P=0.018$.

In Study 2, the total OCI-R score, washing, and obsessing demonstrated a significant positive association with trait revenge. When all OCI-R subscales were entered stepwise as predictors of trait revenge in a multiple regression analysis, obsessing demonstrated a significant unique association with participants' revenge scores ($\beta=0.24$, $P=0.001$). The model accounted for 6% of the variance in participants' trait revenge, $F(1, 176)=10.96$, $P=0.001$.

In Study 3, the total OCI-R score, checking, and ordering were positively related to individuals' motivations to seek revenge against their transgressor.¹ The partial correlations controlled for mood (0.22, $P=0.003$; 0.27, $P < 0.001$; 0.17, $P=0.039$ for the total OCI-R score, checking, and ordering, respectively) were still significant. That is, individuals high (vs. low) in OC symptoms reported significantly higher revenge motivation, even when controlling for mood effects. When all OCI-R subscales were entered stepwise as predictors of revenge motivation in a multiple regression analysis, checking ($\beta=0.28$, $P < 0.001$) emerged as significant predictor. The model accounted for 8% of the variance in participants' revenge motivation, $F(1, 172)=14.62$, $P < 0.001$. When avoidance motivation was analyzed as outcome variable, there were no significant effects.

The effect of heightened revenge in OCD was conceptually replicated in three studies: The OCI-R total scale was consistently and significantly associated with more positive attitudes toward

vengeance (Study 1), greater trait revenge (Study 2), and greater motivation to seek revenge against an actual transgressor (Study 3).

5. Study 4: negative interpersonal events and OC symptoms

Studies 1–3 provided initial evidence for the prediction that OC symptoms are positively associated with revenge. Because revenge and/or vengeful motivations in OCD appear to arise in response to negative interpersonal events (e.g., denigrating remarks of others, relatives' noncompliance with circumscribed rules), we hypothesized that individuals with OCD perceive interpersonal transgressions more frequently. As an extension of Studies 1–3, Study 4 examined the hypothesis that OC symptoms were associated with individuals' self-assessed frequency of incurring negative interpersonal events in their daily lives. Specifically, it was predicted that OC symptoms would be positively associated with the number of transgressions individuals reported over a two-week period.

5.1. Method

5.1.1. Participants

A total of 175 participants ($M_{age}=25.39$, $S.D.=6.21$, age range: 18–54; 86.9% female) voluntarily completed an online questionnaire. Recruitment strategy was identical to that used in Studies 1–3.

5.1.2. Procedure and measures

Upon clicking on the survey link, participants provided informed consent and reported their gender, age, and academic subject. Similar to studies 1–3, they first

¹ Obsessing was at trend level correlated with revenge motivation ($P=0.054$).

Table 2
Descriptive statistics, reliabilities, and intercorrelations among variables in Study 4.

Measure	M	S.D.	α	1	2	3	4	5	6	7	8
1. TOM	0.66	0.62	0.90	–							
2. OCI-R washing	1.13	1.60	0.70	0.12	–						
3. OCI-R checking	2.38	2.12	0.71	0.15*	0.42***	–					
4. OCI-R ordering	2.60	2.26	0.83	0.17*	0.39***	0.41***	–				
5. OCI-R obsessing	2.81	2.66	0.83	0.43***	0.26***	0.36***	0.22**	–			
6. OCI-R hoarding	2.91	2.30	0.79	0.31***	0.24***	0.34***	0.21**	0.37***	–		
7. OCI-R neutralizing	1.08	1.67	0.63	0.18*	0.42***	0.42***	0.48***	0.28***	0.15	–	
8. OCI-R total	12.90	8.36	0.86	0.37***	0.64***	0.74***	0.67***	0.68***	0.61***	0.64***	–

Note. $N=175$. TOM=Transgression Occurrences Measure. OCI-R=Obsessive–Compulsive Inventory-Revised.

* $P < 0.05$.

** $P < 0.01$.

*** $P < 0.001$ (all P 's two-tailed).

completed the OCI-R (Foa et al., 2002; German version: Gonner et al., 2007). Subsequently, participants completed the German (Steiner et al., 2011) 20-item Transgression Occurrences Measure (TOM; McCullough et al., 2003), which assesses the frequency of a variety of interpersonal transgressions. Approximately half of the transgressions were “sins of commission” that would be relatively unambiguous and thus highly visible to a third party (e.g., “degraded you in public”) and approximately half were “sins of omission” that would be less visible to a third party (e.g., “took advantage of you”). On a 7-point scale (0=not at all, 6=constantly), participants were instructed to “indicate how frequently the following events have occurred to you in your relationships with other people in the last 2 weeks.”

5.2. Results and discussion

Table 2 details descriptive statistics, reliabilities, and intercorrelations among all study variables. We predicted that OC symptoms would relate to higher overall transgression occurrence. As hypothesized, individuals who scored higher on the OCI-R reported having incurred significantly higher numbers of transgressions in their daily lives, and this association was rather strong ($r=0.37$, $P < 0.001$). On closer inspection, five of the six OCI-R subscales (checking, ordering, obsessing, hoarding, and neutralizing) were significantly associated with participants' TOM scores, with obsessing demonstrating the strongest relationship ($r=0.43$, $P < 0.001$). When participants' TOM scores were regressed stepwise on the six OCI-R subscales, obsessing ($\beta=0.37$, $P < 0.001$) and hoarding ($\beta=0.18$, $P=0.016$) emerged as significant predictors. Together, the two OCI-R subscales explained 21.4% of the variance in participants' TOM scores, $F(2, 172)=23.44$, $P < 0.001$.

Table 3 presents the correlations between the OCI-R and each of the 20 specific transgressions on the TOM. The total OCI-R score was positively related with all but one of them (“stole from you”), and significantly so with 18 of them. Noteworthy, individuals higher in obsessing and hoarding endorsed all 20 transgressions more strongly than did individuals who were lower on these subscales. For obsessing, 18 of the 20 associations were statistically significant; for hoarding 16 were significant.²

6. General discussion

Extending previous work on angry affect in individuals with OCD, the present research aimed to better understand the relationships between OC symptoms and revenge, which we conceptualized as a motivational response in return for perceived wrong (McCullough et al., 2001). Furthermore, to shed light on the role of interpersonal transgressions in OCD (which is still largely unknown), we proposed and tested that OC symptoms are

associated with individuals' self-assessed frequency of incurring negative interpersonal events in their daily lives.

Our hypothesis that OC symptoms would significantly predict individual differences in revenge was generally supported. The results of three studies (using three independent samples) consistently revealed that OC symptoms were positively associated with individual differences in revenge-seeking attitudes (Study 1), trait revenge (Study 2), and motivations to retaliate against an actual transgressor (Study 3). Consistent with previous findings illustrating that obsessing is positively correlated with aggression (Moritz et al., 2012; Storch et al., 2012), the most consistent relationships appeared between revenge and obsessing. Given the conceptual similarity between revenge and anger (Wilkowski et al., 2012), this pattern of results is also in line with research reporting elevated levels of anger in individuals with OCD (Moscovitch et al., 2008; Painuly et al., 2012; Radomsky et al., 2007; Storch et al., 2012; Whiteside and Abramowitz, 2004). In addition, the observed obsession-revenge link (Studies 1 and 2) parallels findings from research on posttraumatic stress disorder (PTSD), demonstrating that re-experiencing/intrusion symptoms correlate positively with feelings of revenge (Orth et al., 2006). In line with Orth and colleagues, one might interpret the present relations in terms of their shared ruminative features. Indeed, there is evidence to indicate that rumination (i.e., recurrent negative thoughts) about interpersonal offenses is positively associated with revenge (McCullough et al., 1998, 2001).

Our hypothesis that OC symptoms would significantly predict individuals' perceptions of transgression occurrences was also supported. In Study 4, we found that individuals high in OC symptoms reported a significantly greater number of transgressions in their daily lives than did individuals low in OC symptoms. Again, the most robust association occurred with obsessing, suggesting that OCD might involve a heightened sensitivity to negative interpersonal events. This interpretation fits well with recent research indicating that most individuals with OCD incur interpersonal conflicts in their partnerships (Hauschildt et al., 2010), monitor their social environment with suspiciousness and distrust (Moritz et al., 2011), and hold ambivalent interpersonal attitudes toward significant others (Moritz et al., 2009). Furthermore, the heightened interpersonal sensitivity might manifest itself as higher demands for how individuals with OCD wish to be treated by others. For example, the person with OCD might expect others to comply with his or her rituals (e.g., doing something in a circumscribed way), so that refusing to accommodate these rules would be tantamount to a transgression (see Moritz et al., 2012, for a similar account).

From a clinical perspective, the different aspects of revenge may be targeted in therapy. One possibility could be to assist individuals (1) in acknowledging their fantasies and feelings of revenge and (2) to develop strategies to effectively cope with

² Similar to Studies 1–3, controlling for demographic variables (i.e., age and gender) did not lead to different conclusions.

Table 3
Means and standard deviations of transgression occurrences, and correlations with OCI-R subscales (Study 4).

Transgression	M	S.D.	OCI-R subscale						
			Washing	Checking	Ordering	Obsessing	Hoarding	Neutralizing	Total
1. Insulted you	0.92	1.19	−0.01	0.11	0.12	0.36***	0.18*	0.15*	0.25***
2. Took advantage of you	0.87	1.08	0.03	0.07	0.12	0.12	0.21**	0.06	0.17*
3. Betrayed you	0.45	1.01	0.00	−0.01	0.02	0.27***	0.18*	0.131;†	0.16*
4. Lied to you	0.95	1.25	0.07	0.02	0.07	0.24***	0.131;†	0.07	0.16*
5. Was unfaithful to you	0.18	0.55	0.12	0.09	0.17*	0.30***	0.24***	0.22**	0.29***
6. Hurt you physically	0.07	0.35	0.03	−0.02	0.06	0.16*	0.10	0.18*	0.13
7. Spread rumors or gossiped about you	0.47	1.04	0.14	0.07	0.09	0.20**	0.12	0.10	0.19*
8. Damaged something that belonged to you	0.19	0.67	0.07	0.02	0.10	0.25***	0.18*	0.10	0.20**
9. Stole from you	0.07	0.31	0.05	−0.03	−0.06	0.141;†	0.10	0.05	0.06
10. Failed to appreciate you adequately	1.77	1.41	0.00	0.07	0.09	0.32***	0.16*	0.09	0.21**
11. Told a secret that they promised not to tell	0.22	0.65	0.00	0.06	0.03	0.16*	0.25***	0.08	0.16*
12. Got even with you for something that happened previously	0.30	0.75	0.13	0.17*	0.18*	0.22**	0.23**	0.22**	0.29***
13. Benefited from your misfortune	0.59	1.17	0.20**	0.17*	0.22**	0.24***	0.20**	0.22**	0.32***
14. Teased you	0.37	0.70	0.07	0.07	0.12	0.27***	0.32***	0.131;†	0.26***
15. Degraded you in public	0.05	0.32	0.22**	0.27***	0.07	0.31***	0.19*	0.17*	0.31***
16. Was violent toward you	1.06	1.26	0.10	0.18*	0.13	0.38***	0.18*	0.12	0.29***
17. Was “two-faced” or insincere	1.29	1.44	0.03	0.06	0.08	0.34***	0.26***	0.06	0.23**
18. Got you in trouble	0.65	1.22	0.10	0.08	0.03	0.29***	0.23**	0.04	0.21**
19. Told you something that hurt you	1.04	1.35	0.08	0.14	0.13	0.30***	0.19*	0.07	0.25***
20. Failed to protect you or stick up for your rights	1.17	1.49	0.18*	0.19*	0.16*	0.32***	0.21**	0.15*	0.31***

Note. $N = 175$. OCI-R=Obsessive–Compulsive Inventory-Revised.

† $P < 0.10$.

* $P < 0.05$.

** $P < 0.01$.

*** $P < 0.001$ (all P 's two-tailed).

revenge in a socially competent way (Gäbler and Maercker, 2011). Of interest, promising work has already begun to treat interpersonal tensions in OCD (see the meta-cognitive treatment for OCD; Moritz et al., 2010) and should be further explored.

6.1. Limitations and future directions

Several limitations should be acknowledged. First, we did not control for individuals' general distress that might partly account for the revenge-OCD link and/or the perception of interpersonal transgressions. However, as noted by Moritz et al. (2011), partialling out distress could have removed valid variance since distress is intrinsic to OCD. Moreover, Radomsky et al. (2007) reported that initial correlations between anger and OC symptoms (checking in particular) retained significance after adjusting for distress. Noteworthy, the relationship between OC symptoms and revenge motivation (Study 3) persisted after controlling for mood effects. Still, future studies would profit from including a measure of negative affectivity to test whether revenge is uniquely related to OC symptoms. Second, the effect sizes in Studies 1–3 were modest, but promising given that the studies found predicted associations using different non-clinical samples and using different measures of revenge (suggesting that the effects are unlikely due to artifacts of a specific measure). An interesting avenue for future research may be to complement self-reports by adding behavioral observations of revengeful actions. Third, given the correlational nature of these studies, causal inferences regarding the relationship between revenge as well as transgression frequency and OC symptoms cannot be drawn. It remains to be investigated whether OC symptoms contribute to unforgiving motivations/higher frequency of encountering transgressions, or if unforgiving motivations/frequent exposures to interpersonal hurt exacerbate OC symptoms. Longitudinal designs that explore the directionality of these relationships would constitute a promising field for future research. Fourth, subsequent work is needed to examine the potential mechanisms underlying the relationships between OC symptoms and revenge. It might be worthwhile to investigate whether OC-related constructs such as anger or rumination tendencies moderate or even

mediate the associations between OCD symptomatology, revenge and the perception of transgressions. Finally, we used a heterogeneous sample of healthy college students who self-selected to be part of the studies. Clinical samples would be necessary to further generalize and substantiate the validity of our findings. However, our samples were similar to the ones in previous work on the anger-OCD link in that they consisted of nonclinical students. It should also be noted that OC symptoms are normally distributed in the general population (Mataix-Cols et al., 2005) and, hence, resemble the form and content of a clinical sample.

6.2. Conclusions

The present work provides an initial step towards understanding how OC symptoms relate to both individual differences in revenge and the incurrence of interpersonal transgressions in daily life. Replicating the current findings within a clinical sample of formally diagnosed OCD patients will be an important step for future work. When corroborated, consideration of revenge-seeking tendencies in psychotherapy may prove helpful (Horowitz, 2007).

Conflict of interest

None.

Acknowledgments

We wish to thank Professor Hans Reinecker and two anonymous reviewers whose constructive comments helped to improve this paper.

References

- Brown, R.P., 2003. Measuring individual differences in the tendency to forgive: construct validity and links with depression. *Personality and Social Psychology Bulletin* 29, 759–771.
- Fitzgibbons, R.P., 1986. The cognitive and emotive uses of forgiveness in the treatment of anger. *Psychotherapy: Theory, Research, Practice, Training* 23, 629–633.
- Foa, E.B., Huppert, J.D., Leiberg, S., Langner, R., Kichic, R., Hajcak, G., Salkovskis, P.M., 2002. The obsessive–compulsive inventory: development and validation of a short version. *Psychological Assessment* 14, 485–496.
- Gäbler, I., Maercker, A., 2011. Revenge after trauma: theoretical outline. In: Linden, M., Maercker, A. (Eds.), *Embitterment. Societal, Psychological, and Clinical Perspectives*. Springer, Wien, pp. 42–69.
- Gonner, S., Leonhardt, R., Ecker, W., 2007. The German version of the obsessive–compulsive inventory-revised: a brief self-report measure for the multidimensional assessment of obsessive–compulsive symptoms. *Psychotherapie, Psychosomatik, Medizinische Psychologie* 57, 395–404.
- Hauschildt, M., Jelinek, L., Randjbar, S., Hottenrott, B., Moritz, S., 2010. Generic and illness-specific quality of life in obsessive–compulsive disorder. *Behavioural and Cognitive Psychotherapy* 38, 417–436.
- Horowitz, M.J., 2007. Understanding and ameliorating revenge fantasies in psychotherapy. *The American Journal of Psychiatry* 164, 24–27.
- Joinson, A., 1999. Social desirability, anonymity, and internet-based questionnaires. *Behavior Research Methods, Instruments, & Computers* 31, 433–438.
- Karremans, J.C., Van Lange, P.A.M., 2008. The role of forgiveness in shifting from “Me” to “We”. *Self & Identity* 7, 75–88.
- Kempke, S., Luyten, P., 2007. Psychodynamic and cognitive–behavioral approaches of obsessive–compulsive disorder: is it time to work through our ambivalence? *Bulletin of the Menninger Clinic* 71, 291–311.
- Mataix-Cols, D., Rosario-Campos, M.C.d., Leckman, J.F., 2005. A multidimensional model of obsessive–compulsive disorder. *The American Journal of Psychiatry* 162, 228–238.
- McCullough, M.E., Bellah, C.G., Kilpatrick, S.D., Johnson, J.L., 2001. Vengefulness: relationships with forgiveness, rumination, well-being, and the big five. *Personality and Social Psychology Bulletin* 27, 601–610.
- McCullough, M.E., Emmons, R.A., Kilpatrick, S.D., Mooney, C.N., 2003. Narcissists as “Victims”: the role of narcissism in the perception of transgressions. *Personality & Social Psychology Bulletin* 29, 885–893.
- McCullough, M.E., Rachal, K.C., Sandage, S.J., Worthington Jr., E.L., Brown, S.W., Hight, T.L., 1998. Interpersonal forgiving in close relationships: II. Theoretical elaboration and measurement. *Journal of Personality and Social Psychology* 75, 1586–1603.
- Moritz, S., Jelinek, L., Hauschildt, M., Naber, D., 2010. How to treat the untreated: effectiveness of a self-help metacognitive training program (myMCT) for obsessive–compulsive disorder. *Dialogues in Clinical Neuroscience* 12, 209–220.
- Moritz, S., Kempke, S., Luyten, P., Randjbar, S., Jelinek, L., 2011. Was Freud partly right on obsessive–compulsive disorder (OCD)? Investigation of latent aggression in OCD. *Psychiatry Research* 187, 180–184.
- Moritz, S., Niemeyer, H., Hottenrott, B., Schilling, L., Spitzer, C., 2012. Interpersonal ambivalence in obsessive–compulsive disorder. *Behavioural and Cognitive Psychotherapy* 41, 1–16.
- Moritz, S., Wahl, K., Ertle, A., Jelinek, L., Hauschildt, M., Klinge, R., Hand, I., 2009. Neither saints nor wolves in disguise: ambivalent interpersonal attitudes and behaviors in obsessive–compulsive disorder. *Behavior Modification* 33, 274–292.
- Moscovitch, D.A., McCabe, R.E., Antony, M.M., Rocca, L., Swinson, R.P., 2008. Anger experience and expression across the anxiety disorders. *Depression and Anxiety* 25, 107–113.
- Orth, U., Montada, L., Maercker, A., 2006. Feelings of revenge, retaliation motive, and posttraumatic stress reactions in crime victims. *Journal of Interpersonal Violence* 21, 229–243.
- Painuly, N.P., Grover, S., Mattoo, S.K., Gupta, N., 2012. Anger attacks in obsessive compulsive disorder. *Industrial Psychiatry Journal* 20, 115–119.
- Purdon, C., 2011. Assessing comorbidity, insight, family and functioning in OCD. In: Steketee, G. (Ed.), *The Oxford Handbook of Obsessive Compulsive and Spectrum Disorders*. Oxford University Press, Oxford.
- Radomsky, A.S., Ashbaugh, A.R., Gelfand, L.A., 2007. Relationships between anger, symptoms, and cognitive factors in OCD checkers. *Behaviour Research and Therapy* 45, 2712–2725.
- Rhodes, S.D., Bowie, D.A., Hergenrather, K.C., 2003. Collecting behavioural data using the world wide web: considerations for researchers. *Journal of Epidemiology and Community Health* 57, 68–73.
- Spielberger, C.D., 1988. *State Trait Anger Expression Inventory (STAXI)*. Psychological Assessment Resources, Odessa, FL.
- Steiner, M., Allemann, M., McCullough, M.E., 2011. Age differences in forgivingness: the role of transgression frequency and intensity. *Journal of Research in Personality* 45, 670–678.
- Storch, E.A., Jones, A.M., Lack, C.W., Ale, C.M., Sulkowski, M.L., Lewin, A.B., De Nadai, A.S., Murphy, T.K., 2012. Rage attacks in pediatric obsessive-compulsive disorder: phenomenology and clinical correlates. *Journal of the American Academy of Child and Adolescent Psychiatry* 51, 582–592.
- Stuckless, N., Goranson, R., 1992. The vengeance scale: development of a measure of attitudes toward revenge. *Journal of Social Behavior & Personality* 7, 25–42.
- Whiteside, S.P., Abramowitz, J.S., 2004. Obsessive–compulsive symptoms and the expression of anger. *Cognitive Therapy & Research* 28, 259–268.
- Wilkowski, B.M., Hartung, C.M., Crowe, S.E., Chai, C.A., 2012. Men don't just get mad; they get even: revenge but not anger mediates gender differences in physical aggression. *Journal of Research in Personality* 46, 546–555.

Forgiveness involves reduced connectivity between medial prefrontal and dorsal anterior cingulate cortex

Ramzi Fatfouta^{1,2}, Dar Meshi^{1,2}, Angela Merkl^{2,3}, & Hauke R. Heekeren^{1,2}

¹Department of Education and Psychology, Freie Universität Berlin, Habelschwerdter Allee 45, 14195 Berlin, Germany

²Cluster “Languages of Emotion”, Freie Universität Berlin, Habelschwerdter Allee 45, 14195 Berlin, Germany

³Department of Neurology with Experimental Neurology, Charité-Universitätsmedizin, Augustenburger Platz 1, 13353 Berlin, Germany

Correspondence should be addressed to:

Ramzi Fatfouta

Habelschwerdter Allee 45, 14195 Berlin, Germany

Phone: 00493083857857

E-mail: ramzi.fatfouta@fu-berlin.de

Running title: Forgiveness involves reduced MPFC/DACC connectivity

Number of pages: 38

Number of figures: 6

Number of tables: 5

Number of words for Abstract: 163

Number of words for Introduction: 549

Number of words for Discussion: 1135

The authors declare that they have no competing financial interests.

Abstract

Conflict is a ubiquitous feature of interpersonal relationships, yet many of these relationships preserve their value following conflict. Our ability to motivate post-conflict reconciliatory behavior is a characteristic of forgiveness. Using a combination of behavioral and neuroimaging techniques, we show that forgiveness is modulated by relationship closeness. In an iterated social exchange, participants were more likely to forgive their partner compared to an unknown person by accepting unfair exchanges. Importantly, this effect was independent of how resources were actually being shared with one's partner. The medial prefrontal cortex (MPFC) was activated when the partner, rather than the unknown person, behaved unfairly and, in the same context, the MPFC demonstrated greater functional connectivity with the dorsal anterior cingulate cortex (DACC). MPFC-DACC connectivity was inversely associated with participants' tendency to forgive their partner for unfairness as well as performance outside the scanner on a behavioral measure of forgiveness. We conclude that a neural network comprising the MPFC/DACC may serve as a neurocognitive marker of forgiveness.

Keywords: forgiveness, interpersonal relationships, ultimatum game, social cognition, medial prefrontal cortex (MPFC)

Introduction

In a world where conflict pervades social life, humans must be able to successfully negotiate with others. Essential to adaptive behavior is our ability to maintain interpersonal relationships even after conflict (De Waal and Pokorny 2005). This capacity to refrain from retaliation and, instead, to motivate reconciliation with conspecifics despite their harmful prior actions is known as forgiveness (McCullough et al 2013). Forgiveness is deeply rooted in the world's major religious (and spiritual) groups (McCullough and Worthington 1999; Rye et al 2000) and plays a central role in social life across cultures (Ho and Fung 2011); it helps individuals resolve resentment, relieve depression and embitterment, and restore peace of mind (Enright and Fitzgibbons 2015). From the perspective of natural selection, forgiveness likely evolved because it allowed the repair of valuable relationships — that is, relationships whose repair would be anticipated to yield potential long-term benefits (McCullough 2008). Research on nonhuman primates, for instance, indicates that forgiving behavior is most common among genetic relatives and close associates, given their potential contributions to each other's fitness (De Waal and Pokorny 2005; Koski et al 2007). Similarly, in humans, behavioral studies suggest that we are more inclined to forgive someone to whom we feel close and committed (Karremans and Aarts 2007; Karremans et al 2011).

Neuroimaging studies of forgiveness (Farrow et al 2005; Farrow et al 2001; Hayashi et al 2010; Ricciardi et al 2013; Young and Saxe 2009) have found activation in areas previously implicated in mental-state reasoning — known as mentalizing — and cognitive conflict (medial prefrontal and anterior cingulate cortical areas). However, none of these studies examined how these areas influence our decision to forgive actual transgressions — in particular, in the context of close relationships. Furthermore, forgiveness is a complex psychological function

(Worthington 2006) and, hence, unlikely to be implemented by a single brain area, but rather by a network of interacting brain areas. Explaining such a neurobiological mechanism of forgiveness remains an outstanding challenge.

To fill this gap, we extend previous work in three important ways: First, we examine brain regions involved when humans face transgressions by a close other (i.e., their partner) and by a non-close other (i.e., an unknown person). Second, we investigate individuals' changes in brain activity as they interact with these people in a real, consequential social scenario. Third, we analyze the relation between brain activation and individual differences in forgiveness, as assessed by a commonly used forgiveness scale and a newly developed behavioral test (Fatfouta et al 2014). We hypothesized that relationship closeness modulates behavioral and neural responses to transgressions, resulting in greater forgiveness during interactions with one's partner compared to interactions with an unknown person. More specifically, the medial prefrontal cortex (MPFC) is sensitive to relationship closeness in the context of unfair behavior (Campanhã et al 2011) and responds to personally relevant (i.e., close) others (Krienen et al 2010). We therefore expected increased MPFC activity in a comparison of unfair treatment by one's partner versus an unknown person. Notably, conflict is detected in the dorsal anterior cingulate cortex (DACC), with the MPFC playing a role in regulating such conflict (Nakao et al 2010). We hypothesized that forgiveness depends on interactions between the MPFC and regions modulated by conflict between competing responses (i.e., forgive vs. not forgive), particularly the DACC.

Materials and Methods

Overview

To test these hypotheses, we adopted a game-theoretic approach that allows the investigation of forgiveness in an interactive social exchange. More specifically, we had 23 romantically involved couples (46 participants; see below) play a version of the Ultimatum Game (UG; Güth et al 1982). In the UG, one individual (the proposer) proposes how to split a sum of money with another individual (the responder). If the responder accepts, the sum is divided according to the proposal. If the responder rejects, neither individual receives anything. In such a social-exchange game, transgressions are encoded as inequitable divisions of monetary rewards and, hence, entail real consequences for each individual. Typically, responders reject offers less than 30% of the proposer's endowment. In other words, they retaliate for the low offers they receive (Camerer 2003; Crockett et al 2013). By contrast, responders can also accept those offers — that is, they refrain from retaliation (i.e., forgive) and act prosocially toward the proposer (McCullough et al 1997). In our task, participants decided whether to accept or reject offers from their partner and an unknown person (Fig. 1A). We also included a third condition, in which participants received offers from a computer (not relevant to the present research question, details available upon request). All offers were actually predetermined so as to control the number and size of offers made, which ranged from fair (proposing 5€ or 4€ out of 10€) to unfair (proposing 3€, 2€, or 1€ out of 10€) (see Koenigs and Tranel 2007; Mehta and Beer 2010). While playing the UG, one member of each couple ($n = 23$) was scanned using functional magnetic resonance imaging (fMRI), and we were specifically interested in comparing brain activations in the context of unfair offers from one's partner relative to unfair offers from an unknown person.

Upon completion of the UG, all participants reported their subjective impressions by rating how fair, attractive, and trustworthy these proposers were. Participants also rated another person whom they had not interacted with but who had participated in the previous experimental session. This was done to confirm that the partners and the unknown persons were comparable regarding the aforementioned characteristics. Furthermore, we assessed participants' affective responses to unfairness, and importantly, individual differences in forgiveness (i.e., trait forgiveness), which were used to determine whether these individual differences were predicted by participants' brain activation. To explain, in addition to using a standard self-report (explicit) measure of forgiveness we employed a computerized response-time task designed to assess individuals' automatic (implicit) associations regarding forgiveness. This task was an adaptation of the Implicit Association Test (IAT; Greenwald et al 1998), in which participants had to categorize different combinations of self-related words and forgiveness- and retaliation-related words, as fast as possible. Automatic mental associations were inferred from participants' behavior (i.e., response times) on different trial types of this task (Fatfouta et al 2014). The rationale for including such a behavioral test is that spontaneous brain responses are predictive of individuals' automatic associations rather than their explicit (i.e., self-reported) attitudes and beliefs (Stanley et al 2008). Finally, we assessed resource sharing among couples, so as to account for individual differences in how resources were actually being shared in real life.

Participants

Twenty-five heterosexual non-cohabiting couples took part in the study ($N = 50$). On average, couples had been together for 2.22 years ($SD = 1.99$); they maintained separate bank accounts, were unmarried, and had no children. All participants completed the experiment but only one member of each couple was scanned while playing the UG (see below). Four

participants accepted every offer from both their partner and the unknown person and, hence, used a rule-based (vs. fairness-based) strategy (i.e., accepting every offer). These participants were removed from all analyses. Thus, behavioral analyses included 46 participants ($M_{\text{age}} = 24.02$ years, $SD = 4.20$; 22 male) and neuroimaging analyses included 23 participants ($M_{\text{age}} = 24.35$ years, $SD = 3.80$; 15 male). All participants gave written informed consent in accord with local ethics.

Experimental procedure

The experiment was conducted on two separate days. On the first day, each member of the couple separately provided a biometric photograph and rated the perceived level of closeness to his/her romantic partner by using the “Inclusion of Other in the Self” (IOS) scale (Aron et al 1992) (1 = *not at all close*, 7 = *very close*). Around two weeks later, couples returned to play a version of the UG. The experimenter explained the rules of the UG, and informed participants that the roles of proposer and responder within the couple would only be revealed when the experiment began. Participants were told they would be interacting with three different players via an online network: (a) their romantic partner, (b) an unknown person, and (c) a computer. Participants were further told that one round out of all trials would be randomly selected and paid out to them. After the experimenter made sure that all participants understood the instructions, participants were taken to two separate but adjacent rooms to perform the experiment; one participant went into the scanner and the other to a behavioral testing room. To enhance the interactive aspect of our UG task, it was emphasized that both the scanned and the non-scanned participant must start the game simultaneously. While in separate rooms, participants were told to prepare to play the game, and they then heard the experimenter give them both the countdown sequence “ready, set, go”. On “go”, participants were requested to commence the game by button

press, after which they believed a computer algorithm determined their role in the UG. In reality, all participants received an on-screen message that they had been selected to play the game as the responder, and all offers were pre-determined similar to other studies using the UG (Sanfey et al 2003). Participants were then cued to press the button again, at which time the phrase “Connection is established...” was presented, followed by the first experimental trial.

Participants played 180 UG trials in the role of responder. Each trial began with a fixation cross for an average of 5.5 s (variable range: 2–9 s in 1 s increments), followed by the proposer’s photograph (4 s), the offer (4 s), and the outcome for that trial (4 s). While the offer was displayed, participants decided whether to accept or reject it by button press (left/right side counterbalanced across participants). Offers varied across three conditions: In the PARTNER condition (75 trials), participants received offers from their romantic partner. In the UNKNOWN condition (75 trials), participants received offers from an unknown person. The unknown person’s gender was matched to that of the participants’ romantic partner and this identity was held constant across participants. In the COMPUTER condition (30 trials), participants responded to offers from a computer proposer. Each condition contained equal proportions of fair (5€ [offered]: 5€ [kept], 4€:6€) and unfair (3€:7€, 2€:8€, 1€:9€) offers. Offers were equally divided across three functional runs (60 trials/run, 18 min 18 s each) and each run had an equivalent number of trials from the three conditions (25 PARTNER trials, 25 UNKNOWN trials, 10 COMPUTER trials). Trials were presented in a pseudo-random order, with the restriction that no proposer and no offer would be presented on three consecutive trials.

We took several steps to control for potential confounds related to the photographs. First, before the experiment, we selected 29 opposite-sex photographs with a neutral facial expression from the FACES database (Ebner et al 2010) to serve as the photographs of the unknown

persons, and we pretested them. In that pretest, 200 independent raters ($M_{\text{age}} = 23.77$ years, $SD = 3.69$; 100 male) evaluated the photos on attractiveness and trustworthiness (scales ranged from 1 = *not at all attractive/trustworthy*, 7 = *extremely attractive/trustworthy*). These raters were consistent in their attractiveness (Cronbach's $\alpha_{\text{males}} = .95$; $\alpha_{\text{females}} = .94$) and trustworthiness ($\alpha_{\text{males}} = .87$; $\alpha_{\text{females}} = .90$) ratings. For each gender, one male and one female model of both average attractiveness and trustworthiness were selected (FACES picture codes: 069_y_f_n_a, 016_y_m_n_a). Second, to maximize uniformity across participants and study photographs, all images were converted to grayscale and adjusted for size, luminance, and contrast. Images were then placed on a standardized gray background (RGB code: 204, 204, 204) and re-sized to 474 x 591 pixels.

Post-scan measures

Upon completion of the UG task, both participants, while still in separate rooms, completed the following measures (in the order specified).

Fairness, attractiveness, and trustworthiness ratings. Participants rated the perceived fairness, attractiveness, and trustworthiness of both their romantic partner and the unknown person (Likert scales ranged from 1 = *not at all fair/attractive/trustworthy*, 7 = *extremely fair/attractive/trustworthy*). Consistent with research on positive illusions in romantic relationships (Murray and Holmes 1997), we expected participants' ratings of their romantic partner to be positively biased (i.e., more favorable than those made by objective observers). Therefore, a third (control) condition required rating the perceived fairness, attractiveness, and trustworthiness of another gender-matched person (OTHER) who had participated in the previous session. We used these other-ratings as a reality benchmark to demonstrate that, on average, the romantic partners and the unknown persons were comparable regarding the above

characteristics.

Affective responses to unfairness. Participants also rated their subjective affective state for unfair offers made by the romantic partner and the unknown person (in a counterbalanced order across participants). This was performed by asking participants about the following emotions: surprise, anger, happiness, disgust, revenge, pain, confusion, forgiveness, benevolence, and disappointment (randomized within participants; Likert scales ranged from 1 = *not at all*, 7 = *very much*).

Implicit forgiveness. Participants performed a behavioral test of forgiveness using a variant of the implicit association test (IAT; Greenwald et al 1998), the forgiveness IAT (Fatfouta et al 2014). The forgiveness IAT is a computerized, speeded categorization task that measures mental associations of self with forgiveness. Participants were asked to categorize stimuli (words) into one of two categories, which varied across five blocks of trials. The first two blocks consisted of a simple discrimination task, in which participants practiced correctly categorizing stimuli from the target category (me–others) and attribute category (forgiving–vengeful). The third block combined both discrimination tasks (i.e., me–forgiving share one response button; others–vengeful share the other response button). In block 4, the labels of the attribute category were reversed (vengeful–forgiving). Block 5 consisted of the reversed combined discrimination task (i.e., me–vengeful share one response button; others–forgiving share the other response button). Automatic mental associations between “forgiving” and “me” were computed using the recommended scoring algorithm (so-called D_1 ; Greenwald, Banaji, & Nosek, 2003); higher scores reflect stronger associations between self and forgiveness (i.e., faster responses to the me–forgiving block relative to the me–vengeful block).

Explicit forgiveness. Participants' explicit forgiveness was measured using the Tendency to Forgive Scale (Brown 2003), a four-item self-report questionnaire ($\alpha = 0.71$). An example item: "I tend to get over it quickly when someone hurts my feelings" (Likert scales ranged from 1 = *strongly disagree*, 7 = *strongly agree*).

Resource sharing. Finally, we asked participants to indicate how much they share resources in their relationship on a single-item Likert scale (1 = *not at all*, 7 = *extremely*).

MRI data acquisition

Imaging was performed on a 3 T scanner (Trio; Siemens, Erlangen, Germany). Anatomical images were acquired using a T1-weighted MPRage protocol (256 x 256 matrix, 176 sagittal slices of 1 mm thickness). Functional images were acquired using T2*-weighted echoplanar images (EPI) (TR = 2.0 s, TE = 30 ms, 64x64 matrix, flip angle = 70°, field of view = 192 mm, interslice gap = 0.6 mm). A total of 37 axial slices (3 x 3 x 3 mm voxels) parallel to the anterior and posterior commissure were collected per volume. A total of 545 volumes were collected per run. The total scanning time amounted to 54 min 54 s.

fMRI data analysis

Data were preprocessed and analyzed using FMRIB's Software Library, version 5.0.6 (FSL; Smith et al 2004). Preprocessing included: brain extraction (BET; Smith 2002), motion correction (MCFLIRT; Jenkinson et al 2002), spatial smoothing with a Gaussian kernel of 6-mm full-width at half-maximum (FWHM), and high-pass temporal filtering with a cutoff of 80 s. EPI images were registered to each participant's structural image using boundary-based registration (Greve and Fischl 2009), and then normalized into Montreal Neurological Institute (MNI) space using linear registration with 12 degrees of freedom (Jenkinson et al 2002; Jenkinson and Smith 2001).

We estimated a general linear model (GLM) of the blood oxygen level dependent (BOLD) signal using the following 13 regressors (R):

- R1-R3 were indicator functions for the period of the trial when participants discovered the proposer (i.e., PARTNER, UNKNOWN, or COMPUTER)
- R4-R6 were indicator functions denoting a fair offer in each condition
- R7-R9 were indicator functions denoting an unfair offer in each condition
- R10-12 were indicator functions denoting the outcome for each condition
- R13 was an indicator function denoting missed trials where no decision was made within the 4 s time window

All regressors of interest were convolved with a double-gamma hemodynamic response function (HRF). Individual contrast images were calculated and then submitted to a higher-level mixed-effects analysis using voxel-wise one-sample t tests. To identify brain regions showing greater activation for unfair offers from humans (vs. fair offers from humans), we computed the contrast [unfair PARTNER + unfair UNKNOWN > fair PARTNER + fair UNKNOWN]. To identify brain regions showing greater activation for unfair offers from one's partner (vs. unfair from the unknown person), we computed the contrast [unfair PARTNER > unfair UNKNOWN]. For all whole-brain analyses, Z -statistic images were thresholded at $z = 2.3$ and cluster corrected to $P < 0.05$.

Psychophysiological interaction (PPI) analysis

We performed a PPI analysis to identify brain regions showing context-specific changes in the relationship with the MPFC as a function of receiving unfair offers from one's partner versus unfair offers from the unknown person. The model was estimated in three steps (O'Reilly et al 2012). First, we identified the group peak response of the [unfair PARTNER > unfair

UNKNOWN] contrast [MNI peak coordinates (x, y, z) : -2, 66, 12] and created a sphere (5 mm radius) at this location. Second, we extracted individual average time-series of BOLD signal within the seed region. Third, for each participant, we estimated a GLM of the BOLD responses with the following three regressors:

- R1 was a psychological regressor denoting the main effect of task, convolved with a double-gamma HRF
- R2 was a physiological regressor denoting the activation time course of the MPFC seed region
- R3 was a PPI regressor denoting the element-by-element product of the previous two (i.e., the PPI term)

Individual contrast images were calculated and then submitted to a higher-level mixed-effects analysis using voxel-wise one-sample t tests. For all whole-brain analyses, Z-statistic images were thresholded at $z = 2.3$ and cluster corrected to $P < 0.05$.

Results

Behavioral/questionnaire results did not differ between participants inside the fMRI scanner and those outside ($P > 0.129$, for all effects involving group inside vs. group outside scanner). Therefore, in the interest of statistical power, behavioral/questionnaire analyses included participants from both groups (inside and outside the scanner, $N = 46$).

Closeness to the romantic partner

Participants perceived their level of closeness to the romantic partner to be significantly higher than the midpoint (i.e., 4) of the scale [$M = 5.48 \pm 0.14$ SEM, one-sample t -test: $t(45) = 10.44$, $P < 0.001$ (two-tailed); Cohen's $d = 3.11$].

Behavior

Behavioral results revealed that participants were significantly less likely to accept unfair than fair offers [Fig. 1B; repeated-measures analysis of variance: $F(1, 45) = 183.30, P < 0.001$; partial ratio of variance accounted for (η_p^2) = 0.80]. Importantly, participants were significantly more likely to accept offers from their partner than offers from the unknown person [$F(1, 45) = 46.40, P < 0.001; \eta_p^2 = 0.51$]. Furthermore, relationship closeness interacted significantly with fairness [$F(1, 45) = 44.77, P < 0.001; \eta_p^2 = 0.50$]; participants were significantly more likely to accept unfair offers from their partner than unfair offers from the unknown person.

Control analysis

We carried out an additional analysis to rule out an alternative interpretation of our findings. A potential concern is that individual differences in resource sharing could confound participants' acceptance rates in the UG. Greater resource sharing among couples may be reflected in greater acceptance of monetary offers made by the romantic partner. To rule out this interpretation, we repeated the behavioral analysis including resource sharing as a covariate. We replicated our main results, finding significant main effects of fairness ($P < 0.001$) and closeness ($P < 0.001$) as well as a significant fairness x closeness interaction ($P < 0.001$). All possible interactions with resource sharing (closeness x resource sharing, fairness x resource sharing, and closeness x fairness x resource sharing) were not significant (P 's were 0.574, 0.712, and 0.574, respectively). This control analysis indicates that the effects we observe in the UG task relate to differences in interpersonal closeness rather than resource sharing among couples.

Ratings

Mean fairness, attractiveness, and trustworthiness ratings were analyzed using repeated-measures analyses of variance (ANOVA). The within-subject factor was the person rated, with

three levels, the romantic partner (PARTNER), the unknown person (UNKNOWN), and the person from the previous experimental session that participants rated as well (OTHER). We found significant main effects of person for fairness [$F(2,44) = 6.64, P = 0.003$; partial ratio of variance accounted for (η_p^2) = 0.23], attractiveness [$F(2,44) = 85.01, P < 0.001$; $\eta_p^2 = 0.79$], and trustworthiness [$F(2,44) = 43.06, P < 0.001$; $\eta_p^2 = 0.66$]. Planned comparisons revealed the PARTNER ratings to differ significantly from the UNKNOWN and OTHER ratings. However, there was no significant difference between UNKNOWN and OTHER ratings (Fig. 2). Thus, although participants' ratings of their romantic partner were positively biased, our sample was objectively of comparable fairness, attractiveness, and trustworthiness.

We next examined the extent to which relationship closeness also modulates individuals' internal feeling states in response to unfairness. The order of the rated proposer (PARTNER vs. UNKNOWN) had no effect on participants' subjective affective state for unfair offers [repeated-measures ANOVA: $F(1, 44) = 0.18, P = 0.67$; $\eta_p^2 = 0.00$]. Differences in mean emotion ratings for unfair offers made by the romantic partner and the unknown person were analyzed using paired-samples *t* tests. Most notably, emotion ratings revealed greater forgiveness and benevolence in response to unfair offers from the partner than unfair offers from the unknown person (see Table 1 for all results).

Neuroimaging

We analyzed the neuroimaging data to reveal brain regions showing greater activation for unfair than fair offers from human proposers. Our results included the bilateral anterior insula (AI), dorsal anterior cingulate cortex (DACC), and right dorsolateral prefrontal cortex (DLPFC) (Fig. 3, A and B; Table 2).

We then compared the neural correlates of unfairness separately for each proposer, and found that unfair offers from the partner and the unknown person activated an overlapping subset of brain areas, including right insular cortex, middle occipital gyrus, and DACC (Fig. 4; Tables 3 and 4).

We next performed a contrast to reveal brain regions showing greater activation for unfair offers from one's partner than from an unknown person. This analysis revealed only one significant cluster in the MPFC (Fig. 5, A and B) [Montreal Neurological Institute (MNI) coordinates (x, y, z) : $-2, 66, 12$]. The MPFC has been implicated in meta-cognitive processes of reasoning about intentions (Amodio and Frith 2006) and, hence, activation here might reflect increased efforts to understand the partner's motives for his/her unfair behavior (Campanhã et al 2011; Meyer et al 2013). This consideration motivated us to examine whether the region of MPFC identified overlaps with regions that have been preferentially associated with tasks relevant to mentalizing in the neuroimaging literature. To demonstrate that MPFC activity in the contrast [unfair PARTNER > unfair UNKNOWN] maps directly onto the mentalizing network, we used Neurosynth (Yarkoni et al 2011). Neurosynth is a platform for large-scale, automated meta-analysis of fMRI data (<http://neurosynth.org/>). In particular, it allows a term-based search of published fMRI studies (numbering 9721 at the time of analysis) for common activation patterns. For the term "mentalizing", we isolated 98 studies with 4526 activation foci. The activation patterns retrieved from Neurosynth are depicted in Figure 5C (light blue). Results from the contrast [unfair PARTNER > unfair UNKNOWN] are depicted as well (orange), and the overlap with the Neurosynth map (purple) indicates that a significant number of voxels within the MPFC (660 voxels, 1.98 cm³), which were recruited in the unfair PARTNER

condition, are also reliably recruited during mentalizing. That is, as expected, the observed task-related change in MPFC activation overlaps to a large degree with the mentalizing network.

We had hypothesized that forgiveness is related to the interaction between brain regions involved in mentalizing and regions that process cognitive conflict. Therefore, we predicted different functional connectivity between these regions in the context of unfair behavior from one's partner relative to unfair behavior from the unknown person. To this end, we performed a psychophysiological interaction (PPI; O'Reilly et al 2012) analysis to identify brain regions exhibiting a change in functional connectivity with the MPFC depending on whether the partner or the unknown person made an unfair offer. This analysis revealed heightened functional connectivity between the MPFC and anterior portions of the DACC (Fig. 6A; Table 5).

We next tested to which degree MPFC-DACC functional connectivity was related to participants' forgiveness of unfair partner offers (vs. unfair unknown offers). Therefore, we calculated the absolute difference between accepted unfair offers from the partner and the unknown person [$\Delta\text{forgiveness} = (|\% \text{ of accepted unfair offers}_{\text{PARTNER}} - \% \text{ of accepted unfair offers}_{\text{UNKNOWN}}|)$]. We refer to this measure as the forgiveness index. A forgiveness index greater than zero indicates that an individual accepted unfair offers from his/her partner more frequently than from the unknown person and, hence, showed a higher tendency to forgive his/her partner (vs. the unknown person). We found that MPFC-DACC functional connectivity was inversely associated with participants' forgiveness index (Fig. 6B); individuals who showed lower levels of functional connectivity between the MPFC and DACC were better able to forgive their partners following unfairness [$\beta = -0.58$, $t = -3.29$, $P = 0.003$ (two-tailed); 95% confidence interval (CI) for $\beta = -0.80/-0.22$].

Finally, we determined whether the task-related change in MPFC-DACC functional connectivity predicted individuals' forgiveness level in two independent measures of forgiveness assessed outside the MRI scanner: one was a self-report questionnaire [Tendency to Forgive scale (TTF; Brown 2003)] and one was a behavioral test [forgiveness IAT (Fatfouta et al 2014)]. As Figure 6C illustrates, the task-related change in MPFC-DACC functional connectivity inversely predicted individuals' IAT scores [left panel; $\beta = -0.59$, $t = -3.34$, $P = 0.003$ (two-tailed); 95% CI for $\beta = -0.81/-0.23$], but not their TTF scores [right panel; $\beta = -0.01$, $t = -0.05$, $P = 0.961$ (two-tailed); 95% CI for $\beta = -0.42/0.40$]. In other words, those individuals who showed lower levels of functional connectivity between the MPFC and DACC showed stronger automatic associations between the self and forgiveness, which were not readily accessible in the self-report.

We additionally performed a leave-one-subject out (LOSO) cross-validation (Esterman et al 2010) to establish the generalizability of regressions estimated with participants' PPI term (MPFC-DACC functional connectivity). In the LOSO cross-validation, a linear regression to predict an individual's forgiveness level (i.e., implicit and explicit forgiveness) from his/her PPI term, was repeated n times, whereby each time a different subject was left out from the analysis. The robustness of the result is demonstrated through significance of all LOSO iterations, which was the case here ($0.015 > P's > 0.002$).

Discussion

This study sought to identify how facing transgressions (in our case, unfairness in an iterated social-exchange game) relates to subsequent forgiveness and its neural mechanism. We were specifically interested in determining how relationship closeness modulates reactions to unfairness, with the source of unfairness — and therefore the object of forgiveness — being

one's own partner, compared to an unknown person. Our study yielded three main results: First, participants were more likely to accept unfair offers from their partner compared with the unknown person, and this effect remained significant after adjustment for resource sharing among couples. Second, when participants received unfair partner offers (vs. unfair unknown offers), a significant change in brain activation was observed in the MPFC; this region also demonstrated heightened functional connectivity with the DACC when the partner rather than the unknown person made an unfair offer. Third, lower levels of the task-related change in MPFC-DACC functional connectivity predicted participants' tendency to forgive one's partner following unfairness as well as behavioral but not self-reported expressions of forgiveness.

Our findings support the valuable-relationships account of forgiveness (McCullough 2008), with relationship closeness influencing economically relevant social decision-making: when a partner rather than an unknown person made an unfair offer, individuals were more likely to accept this offer. The observed effects of relationship closeness on participants' acceptance rates are unlikely to simply reflect differences in how resources were actually being shared with one's partner, as resource sharing did not affect participants' responses. Rather, this behavioral tendency to accept unfair partner offers reflects a forgiveness component, as mirrored in participants' emotion ratings. In this respect, our findings fit with the notion that enforcing fairness norms through retaliation would be mitigated when interacting with a close ally, but prevail when interacting with a stranger (Campanhã et al 2011). Consistent with this notion, related research has demonstrated that individuals accept more unfair offers when playing with an in-group member (and by extension, close ally) than with an out-group member (Diekhof et al 2014).

When facing unfairness, participants activated a network of areas comprising bilateral AI, DACC, and DLPFC. These areas are consistent with those found in previous fMRI studies of the UG, and related tasks involving social decision-making (Gabay et al 2014; Sanfey et al 2003). Extending these studies with regard to the relationship one has with the proposer, we found activation overlap in the right insular cortex and DACC for unfair offers from both proposers, that is, one's partner and the unknown person. This finding lends additional support for the notion that the insula and DACC play a critical role in detecting social norm violations (fairness norms, in particular) induced by inequity (Güroğlu et al 2011).

The fact that the MPFC was the only region activated when participants received unfair offers from their partner, as compared to unfair offers from the unknown person, fits with the proposed role of the MPFC in mentalizing-related computations during social exchange (Hampton et al 2008), and also with the recent literature implicating MPFC in mentalizing about close others (Krienen et al 2010). Being treated unfairly by one's partner likely instigates a thought process, by which one tries to make sense of the reasons and motivations underlying his/her behavior as well as to anticipate his/her future behavior. The overlap between our fMRI results and the reverse-inference meta-analysis results confirmed this rationale. Concordance between the results of this study and fMRI studies of mentalizing indicates that processing unfair offers from one's partner versus an unknown person activates the same region as mental-state reasoning (i.e., MPFC).

Our results point to a functional relation between MPFC and DACC in situations in which participants decide whether to forgive (i.e., refrain from retaliating) unfair partner behavior. Supporting our hypothesis, we found that the changes in brain activity in MPFC covaried significantly more with the DACC when one's partner behaved unfairly than when an

unknown person behaved unfairly. Previous studies have implicated the DACC in the evaluation of conflict (Botvinick et al 1999) and, more specifically, the conflict between cognitive (“accept”) and emotional (“reject”) motivations in the UG (Gabay et al 2014; Sanfey et al 2003). In the present study, unfair partner offers may face individuals with such a conflictual situation: acting upon fairness norms might promote better future offers on the one hand, but also signal retaliatory cost impositions on the other hand and, through this, jeopardize one’s relationship with a loved one. In line with the regulatory role of prefrontal cortex in modulating ACC activation when competing response tendencies are at hand (Botvinick et al 2001) and with the MPFC’s connectivity with the DACC during decisional conflicts (Nakao et al 2010), such conflict encoded in DACC may then be resolved through mentalizing-related MPFC activation.

Participants with reduced MPFC-DACC connectivity showed greater forgiveness, such that unfair offers from one’s partner were more frequently accepted than from an unknown person. As discussed above, DACC activation appears to represent conflict (Botvinick et al 1999) and MPFC likely represents mentalizing (Fig. 5C; Hampton et al 2008; Krienen et al 2010), therefore, we suggest that the less information about conflict that is online and readily available to the MPFC during a transgression, the more likely an individual is to adopt an accepting stance towards that transgression. Importantly, this pattern was also obtained with an independent behavioral measure of forgiveness outside the MRI scanner; reduced MPFC-DACC connectivity predicted individuals’ automatic (implicit) mental associations regarding forgiveness. In line with related research on the neural basis of implicit social cognition (Stanley et al 2008), only the implicit (IAT-measured) but not the explicit (self-reported) expression of forgiveness was predicted by brain activation. We interpret these results in terms of an enhanced tendency toward forgiveness in the context of close relationships rather than an enhancement of

rational self-interested behavior. To achieve an optimal exchange in the context of close relationships, one is better advised to forfeit even small gains as a way of preventing potential loss in the long run (e.g., fitness or other benefits such as social support).

Taken together, our results provide evidence for a neural mechanism of forgiveness. We have demonstrated that feeling socially close to someone enhanced the tendency towards forgiveness (i.e., reduced the drive for rejecting unfair offers), and this tendency was associated with a network comprising the MPFC and DACC. Differences in connectivity between these regions were inversely related to forgiveness across participants. In the future, MPFC-DACC functional connectivity may serve as a neurocognitive marker of forgiveness, which opens up exciting avenues for investigating problems with expressing this ability — in particular, psychopathological disorders following trauma that are accompanied by embitterment and revenge (Gäbler and Maercker 2011). More generally, these findings have important implications for human decision-making because they demonstrate how relationship closeness modulates prefrontal brain function during strategic interactions, thereby advancing our knowledge of the context-dependent nature of social behavior.

References

- Amodio DM, Frith CD. 2006. Meeting of minds: the medial frontal cortex and social cognition. *Nature Reviews Neuroscience* 7(4):268-277.
- Aron A, Aron EN, Smollan D. 1992. Inclusion of Other in the Self Scale and the structure of interpersonal closeness. *Journal of personality and social psychology* 63(4):596.
- Botvinick MM, Braver TS, Barch DM, Carter CS, Cohen JD. 2001. Conflict monitoring and cognitive control. *Psychol Rev* 108(3):624-52.
- Botvinick MM, Nystrom LE, Fissell K, Carter CS, Cohen JD. 1999. Conflict monitoring versus selection-for-action in anterior cingulate cortex. *Nature* 402(6758):179-181.
- Brown RP. 2003. Measuring individual differences in the tendency to forgive: construct validity and links with depression. *Pers Soc Psychol Bull* 29(6):759-71.
- Camerer CF. 2003. *Behavioral game theory: Experiments in strategic interaction*. New York, NY, US: Russell Sage Foundation.
- Campanhã C, Minati L, Fregni F, Boggio PS. 2011. Responding to unfair offers made by a friend: Neuroelectrical activity changes in the anterior medial prefrontal cortex. *The Journal of Neuroscience* 31(43):15569-15574.
- Crockett MJ, Apergis-Schoute A, Herrmann B, Lieberman M, Müller U, Robbins TW, Clark L. 2013. Serotonin modulates striatal responses to fairness and retaliation in humans. *The Journal of Neuroscience* 33(8):3505-3513.
- De Waal FB, Pokorny JJ. 2005. Primate conflict resolution and its relation to human forgiveness. In: Worthington E, editor. *Handbook of forgiveness*. New York: Routledge. p. 17-32.

- Diekhof EK, Wittmer S, Reimers L. 2014. Does competition really bring out the worst? Testosterone, social distance and inter-male competition shape parochial altruism in human males. *Plos One* 9(7):e98977-e98977.
- Ebner NC, Riediger M, Lindenberger U. 2010. FACES – a database of facial expressions in young, middle-aged, and older women and men: development and validation. *Behav Res Methods* 42(1):351-62.
- Enright RD, Fitzgibbons RP. 2015. *Forgiveness therapy: An empirical guide for resolving anger and restoring hope*. Washington, DC, US: American Psychological Association.
- Esterman M, Tamber-Rosenau BJ, Chiu YC, Yantis S. 2010. Avoiding non-independence in fMRI data analysis: leave one subject out. *Neuroimage* 50(2):572-6.
- Farrow TF, Hunter MD, Wilkinson ID, Gouneea C, Fawbert D, Smith R, Lee K-H, Mason S, Spence SA, Woodruff PWR. 2005. Quantifiable change in functional brain response to empathic and forgivability judgments with resolution of posttraumatic stress disorder. *Psychiatry Research: Neuroimaging* 140(1):45-53.
- Farrow TF, Zheng Y, Wilkinson ID, Spence SA, Deakin JFW, Tarrrier N, Griffiths PD, Woodruff PWR. 2001. Investigating the functional anatomy of empathy and forgiveness. *NeuroReport: For Rapid Communication of Neuroscience Research* 12(11):2433-2438.
- Fatfouta R, Schröder-Abé M, Merkl A. 2014. Forgiving, fast and slow: validity of the implicit association test for predicting differential response latencies in a transgression-recall paradigm. *Front Psychol* 5:728.
- Gabay AS, Radua J, Kempton MJ, Mehta MA. 2014. The Ultimatum Game and the brain: A meta-analysis of neuroimaging studies. *Neuroscience & Biobehavioral Reviews* 47:549-558.

- Gäbler I, Maercker A. 2011. Revenge after trauma: Theoretical outline. In: Linden M, Maercker A, editors. *Embitterment. Societal, psychological, and clinical perspectives*. Wien: Springer. p. 42-69.
- Greenwald AG, McGhee DE, Schwartz JL. 1998. Measuring individual differences in implicit cognition: the implicit association test. *J Pers Soc Psychol* 74(6):1464-80.
- Greenwald AG, Nosek BA, Banaji MR. 2003. Understanding and using the implicit association test: I. An improved scoring algorithm. *J Pers Soc Psychol* 85(2):197-216.
- Greve DN, Fischl B. 2009. Accurate and robust brain image alignment using boundary-based registration. *Neuroimage* 48(1):63-72.
- Güroğlu B, van den Bos W, van Dijk E, Rombouts SARB, Crone EA. 2011. Dissociable brain networks involved in development of fairness considerations: Understanding intentionality behind unfairness. *NeuroImage* 57(2):634-641.
- Güth W, Schmittberger R, Schwarze B. 1982. An experimental analysis of ultimatum bargaining. *Journal of economic behavior & organization* 3(4):367-388.
- Hampton AN, Bossaerts P, O'Doherty JP. 2008. Neural correlates of mentalizing-related computations during strategic interactions in humans. *PNAS Proceedings of the National Academy of Sciences of the United States of America* 105(18):6741-6746.
- Hayashi A, Abe N, Ueno A, Shigemune Y, Mori E, Tashiro M, Fujii T. 2010. Neural correlates of forgiveness for moral transgressions involving deception. *Brain Research* 1332:90-99.
- Ho MY, Fung HH. 2011. A dynamic process model of forgiveness: A cross-cultural perspective. *Review of General Psychology* 15(1):77-84.

- Jenkinson M, Bannister P, Brady M, Smith S. 2002. Improved optimization for the robust and accurate linear registration and motion correction of brain images. *Neuroimage* 17(2):825-41.
- Jenkinson M, Smith S. 2001. A global optimisation method for robust affine registration of brain images. *Med Image Anal* 5(2):143-56.
- Karremans JC, Aarts H. 2007. The role of automaticity in determining the inclination to forgive close others. *Journal of Experimental Social Psychology* 43(6):902-917.
- Karremans JC, Regalia C, Paleari FG, Fincham FD, Cui M, Takada N, Ohbuchi K-I, Terzino K, Cross SE, Uskul AK. 2011. Maintaining harmony across the globe: The cross-cultural association between closeness and interpersonal forgiveness. *Social Psychological and Personality Science* 2(5):443-451.
- Koenigs M, Tranel D. 2007. Irrational economic decision-making after ventromedial prefrontal damage: Evidence from the Ultimatum Game. *The Journal of Neuroscience* 27(4):951-956.
- Koski SE, Koops K, Sterck EHM. 2007. Reconciliation, relationship quality, and postconflict anxiety: Testing the integrated hypothesis in captive chimpanzees. *American Journal of Primatology* 69(2):158-172.
- Krienen FM, Tu P-C, Buckner RL. 2010. Clan mentality: Evidence that the medial prefrontal cortex responds to close others. *The Journal of Neuroscience* 30(41):13906-13915.
- McCullough ME. 2008. *Beyond revenge: The evolution of the forgiveness instinct*. San Francisco: Jossey-Bass.
- McCullough ME, Kurzban R, Tabak BA. 2013. Cognitive systems for revenge and forgiveness. *Behav Brain Sci* 36(1):1-15.

- McCullough ME, Worthington E. 1999. Religion and the Forgiving Personality. *Journal of Personality* 67(6):1141-1164.
- McCullough ME, Worthington E, Rachal KC. 1997. Interpersonal forgiving in close relationships. *Journal of Personality and Social Psychology* 73(2):321-336.
- Mehta PH, Beer J. 2010. Neural Mechanisms of the Testosterone-Aggression Relation: The Role of Orbitofrontal Cortex. *Journal of Cognitive Neuroscience* 22(10):2357-2368.
- Meyer ML, Masten CL, Ma Y, Wang C, Shi Z, Eisenberger NI, Han SH. 2013. Empathy for the social suffering of friends and strangers recruits distinct patterns of brain activation. *Social Cognitive and Affective Neuroscience* 8(4):446-454.
- Murray SL, Holmes JG. 1997. A leap of faith? Positive illusions in romantic relationships. *Personality and Social Psychology Bulletin* 23(6):586-604.
- Nakao T, Osumi T, Ohira H, Kasuya Y, Shinoda J, Yamada J, Northoff G. 2010. Medial prefrontal cortex-dorsal anterior cingulate cortex connectivity during behavior selection without an objective correct answer. *Neuroscience Letters* 482(3):220-224.
- O'Reilly JX, Woolrich MW, Behrens TE, Smith SM, Johansen-Berg H. 2012. Tools of the trade: psychophysiological interactions and functional connectivity. *Soc Cogn Affect Neurosci* 7(5):604-9.
- Ricciardi E, Rota G, Sanil L, Gentili C, Gaglianese A, Guazzelli M, Pietrini P. 2013. How the brain heals emotional wounds: The functional neuroanatomy of forgiveness. *Frontiers in Human Neuroscience* 7.
- Rye MS, Pargament KI, Ali MA, Beck GL, Dorff EN, Hallisey C, Narayanan V, Williams JG. 2000. Religious perspectives on forgiveness. In: McCullough M, Pargament KI,

Thoresen CE, editors. *Forgiveness: Theory, research, and practice*. New York: Guilford Press. p. 17-40.

Sanfey AG, Rilling JK, Aronson JA, Nystrom LE, Cohen JD. 2003. The neural basis of economic decision-making in the Ultimatum Game. *Science* 300(5626):1755-8.

Smith SM. 2002. Fast robust automated brain extraction. *Hum Brain Mapp* 17(3):143-55.

Smith SM, Jenkinson M, Woolrich MW, Beckmann CF, Behrens TE, Johansen-Berg H, Bannister PR, De Luca M, Drobnjak I, Flitney DE et al. . 2004. Advances in functional and structural MR image analysis and implementation as FSL. *Neuroimage* 23 Suppl 1:S208-19.

Stanley D, Phelps E, Banaji M. 2008. The Neural Basis of Implicit Attitudes. *Current Directions in Psychological Science* 17(2):164-170.

Worthington EL. 2006. *Forgiveness and reconciliation: Theory and application*. New York, NY, US: Routledge/Taylor & Francis Group.

Yarkoni T, Poldrack RA, Nichols TE, Van Essen DC, Wager TD. 2011. Large-scale automated synthesis of human functional neuroimaging data. *Nat Methods* 8(8):665-70.

Young L, Saxe R. 2009. Innocent intentions: A correlation between forgiveness for accidental harm and neural activity. *Neuropsychologia* 47(10):2065-2072.

Tables

Table 1. Affective responses to unfairness. Mean self-reported emotion ratings (scale 1–7) in response to unfair offers in the UG, by proposer. *SEM* indicates one standard error of the mean and *d* indicates effect size estimates (Cohen’s *d*).

Emotion	$M_{\text{PARTNER}} \pm SEM$	$M_{\text{UNKNOWN}} \pm SEM$	Paired-samples $t(45)$	P (two-tailed)	d
Surprise	4.35 ± 0.28	2.50 ± 0.18	6.38	<0.001	1.16
Anger	2.89 ± 0.26	3.13 ± 0.26	−0.72	0.475	−0.14
Happiness	2.09 ± 0.20	1.52 ± 0.13	2.42	0.02	0.50
Disgust	1.09 ± 0.04	1.57 ± 0.20	−2.41	0.02	−0.50
Revenge	1.61 ± 0.16	1.83 ± 0.20	−0.89	0.379	−0.18
Pain	1.91 ± 0.18	1.52 ± 0.16	1.96	0.057	0.33
Confusion	4.15 ± 0.29	2.67 ± 0.23	5.43	<0.001	0.84
Forgiveness	4.48 ± 0.26	3.26 ± 0.25	4.11	<0.001	0.71
Benevolence	3.35 ± 0.29	2.02 ± 0.21	4.60	<0.001	0.77
Disappointment	3.41 ± 0.27	2.52 ± 0.25	2.64	0.011	0.51

Table 2. Brain regions showing significantly greater activation for unfair than fair offers from a human partner (unfair PARTNER + unfair UNKNOWN > fair PARTNER + fair UNKNOWN).

Brain region	MNI coordinates			Cluster size (voxels)	Peak <i>z</i> -value
	<i>x</i>	<i>y</i>	<i>z</i>		
L middle occipital gyrus	-38	-84	12	5203	5.39
R dorsal anterior cingulate cortex (DACC)	14	12	60	2755	4.17
R insula	34	20	-10	1542	4.54
R superior parietal lobule	22	-66	50	1344	4.55
R dorsolateral prefrontal cortex (DLPFC)	34	50	16	1145	4.16
R cerebellum	48	-52	-34	1053	4.11
L insula	-44	18	-12	889	4.68

MNI = Montreal Neurological Institute. $z > 2.3$, $P < 0.05$, cluster-corrected; L, left; R, right.

Table 3. Brain regions showing significantly greater activation for unfair than fair offers from a romantic partner (unfair PARTNER > fair PARTNER)

Brain region	MNI coordinates			Cluster size (voxels)	Peak <i>z</i> -value
	<i>x</i>	<i>y</i>	<i>z</i>		
R dorsal anterior cingulate cortex (DACC)	8	32	30	903	3.96
L middle occipital gyrus	-44	-78	8	872	3.64
R insula	32	18	-10	498	3.76

MNI = Montreal Neurological Institute. $z > 2.3$, $P < 0.05$, cluster-corrected; L, left; R, right.

Table 4. Brain regions showing significantly greater activation for unfair than fair offers from an unknown person (unfair UNKNOWN > fair UNKNOWN).

Brain region	MNI coordinates			Cluster size (voxels)	Peak <i>z</i> -value
	<i>x</i>	<i>y</i>	<i>z</i>		
L middle occipital gyrus	-36	-84	10	4305	5.11
R Cerebellum	50	-52	-34	2958	4.19
R dorsal anterior cingulate cortex (DACC)	-6	10	46	2023	4.02
R insula/ inferior frontal operculum	42	24	-14	1674	3.8
L temporal pole	-46	16	-12	1261	4.31
R middle occipital gyrus	40	-80	24	1182	3.92
L middle frontal gyrus (MFG)	-38	44	32	688	3.98
L middle temporal gyrus (MTG)	-54	-50	20	517	4.0

MNI = Montreal Neurological Institute. $z > 2.3$, $P < 0.05$, cluster-corrected; L, left; R, right.

Table 5. Brain regions showing increased coupling with the MPFC as a function of receiving unfair offers from one's partner relative to the unknown person.

Brain region	MNI coordinates			Cluster size (voxels)	Peak <i>z</i> -value
	<i>x</i>	<i>y</i>	<i>z</i>		
L lingual gyrus	-28	-84	-14	27026	5.59
L dorsal anterior cingulate cortex (DACC)	0	22	38	1192	3.75
R precentral gyrus	44	6	42	1142	4.38

MNI = Montreal Neurological Institute. $z > 2.3$, $P < 0.05$, cluster-corrected; L, left; R, right.

Captions to figures

Figure 1. Experimental design and behavioral results. (A) Timeline of a single trial in the Ultimatum Game. In each trial, participants first viewed a fixation cross, then a photograph of the proposer, the offer, and then the outcome. While the offer was on screen, participants responded whether they accepted or rejected it by button press. (B) Mean acceptance rates (% of accepted offers) for fair and unfair offers from the partner and the unknown person. Error bars indicate one standard error of the mean, calculated within-subjects. *** $P < 0.001$ (two-tailed), * $P < 0.05$ (two-tailed).

Figure 2. Fairness, attractiveness, and trustworthiness ratings. Mean rating scores for perceived fairness, trustworthiness, and attractiveness, by person. Persons included the romantic partner (PARTNER), the unknown person (UNKNOWN), and the person from the previous experimental session (OTHER). Error bars indicate one standard error of the mean, calculated within-subjects. *** $P < 0.001$ (two-tailed), ** $P < 0.005$ (two-tailed), * $P < 0.05$ (two-tailed).

Figure 3. BOLD signal change related to unfairness. (A) Significant changes in BOLD signal for the contrast [unfair PARTNER + unfair UNKNOWN > fair PARTNER + fair UNKNOWN] were found in bilateral AI, DACC, right cerebellum, left middle occipital gyrus, and right DLPFC. For complete list of results, see Table 2. Results were thresholded at $z > 2.3$, $P < 0.05$, cluster corrected. L = left hemisphere. (B) Mean parameter estimates in arbitrary units (a.u.) within bilateral AI and DACC for fair and unfair offers from the partner and the unknown person. Error bars indicate one standard error of the mean, calculated within-subjects.

Figure 4. Overlap of BOLD signal change related to unfairness. Comparison of z statistic images of unfair offers from each proposer [unfair PARTNER > fair PARTNER in red-yellow; unfair UNKNOWN > fair UNKNOWN in blue-light blue]. Overlapping activations are displayed in green. The overlap included areas in right insular cortex, left middle occipital gyrus, and DACC. Results were thresholded at $z > 2.3$, $P < 0.05$, cluster corrected.

Figure 5. Brain activation associated with unfairness by the partner. (A) Significant changes in BOLD signal for the contrast [unfair PARTNER > unfair UNKNOWN] were found in MPFC. Results were thresholded at $z > 2.3$, $P < 0.05$, cluster corrected. **(B)** Mean parameter estimates in arbitrary units (a.u.) within MPFC for unfair offers from the partner and the unknown person. Error bars indicate one standard error of the mean, calculated within-subjects. **(C)** Comparison of Neurosynth meta-analysis search for the term “mentalizing” (light blue) to study results for the contrast [unfair PARTNER > unfair UNKNOWN] (orange, same as in A). Overlapping activations are displayed in purple (660 voxels, 1.98 cm³). Neurosynth results were corrected for multiple comparisons using a whole-brain false discovery rate (FDR) threshold of $P < 0.05$.

Figure 6. Reduced connectivity between MPFC and DACC predicts forgiveness. (A) PPI analysis identifying brain regions showing connectivity changes with the MPFC as a function of receiving unfair offers from the partner versus the unknown person. The upper panel depicts our functionally defined seed region (i.e., MPFC). The lower panel depicts brain regions whose connectivity with MPFC increases for unfair offers from the partner (vs. the unknown person). Results were thresholded at $z > 2.3$, $P < 0.05$, cluster corrected. **(B)** Regression line depicting the relationship between in MPFC-DACC functional connectivity and forgiveness index (Δ) across

participants. The x-axis represents the mean parameter estimates in arbitrary units (a.u.) for task-related change in MPFC-DACC functional connectivity. The y-axis represents the forgiveness index, defined as the absolute difference between accepted unfair offers from the partner versus the unknown person [$\Delta\text{forgiveness} = (|\% \text{ of accepted unfair offers}_{\text{PARTNER}} - \% \text{ of accepted unfair offers}_{\text{UNKNOWN}}|)$]. (C) Regression lines depicting the relationship between MPFC-DACC functional connectivity and forgiveness levels across participants. The x-axis in both plots represents the mean parameter estimates in arbitrary units (a.u.) for task-related change in MPFC-DACC functional connectivity. The y-axis represents individual differences in forgiveness as measured by implicit (left panel; IAT score) and explicit (right panel; TTF score) measures.

Figures

Figure 1

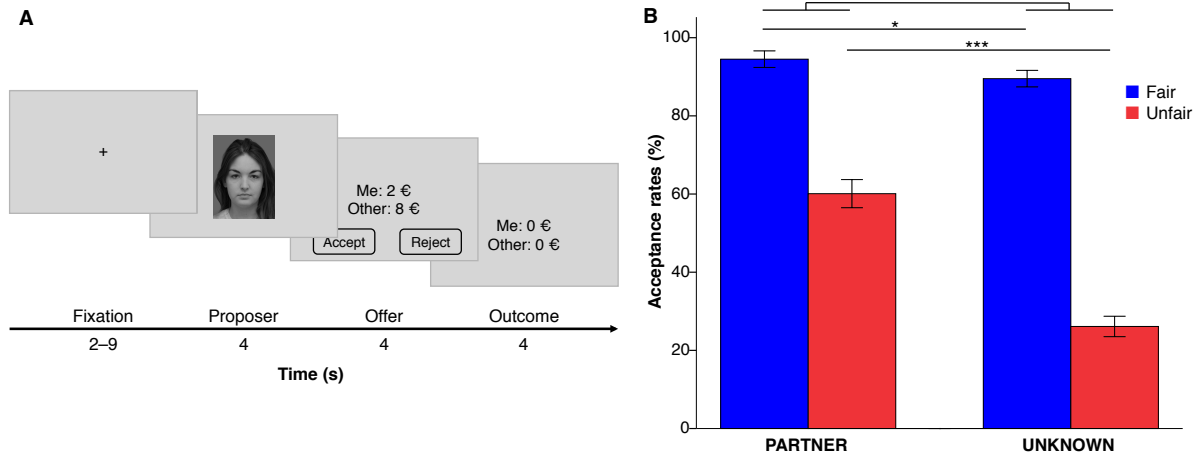


Figure 2

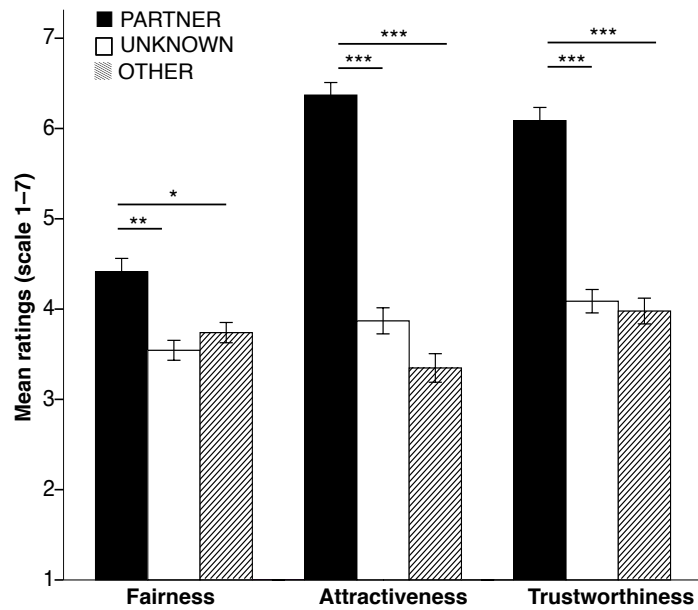


Figure 3

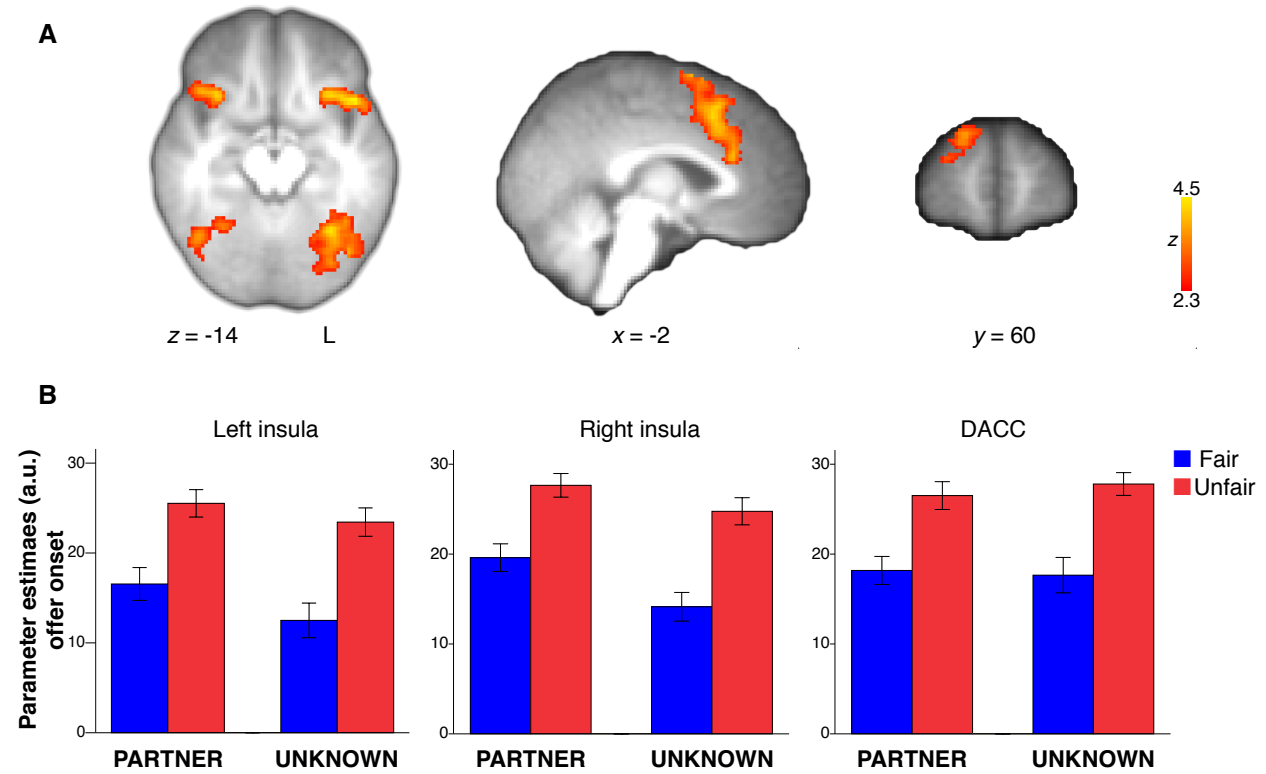


Figure 4

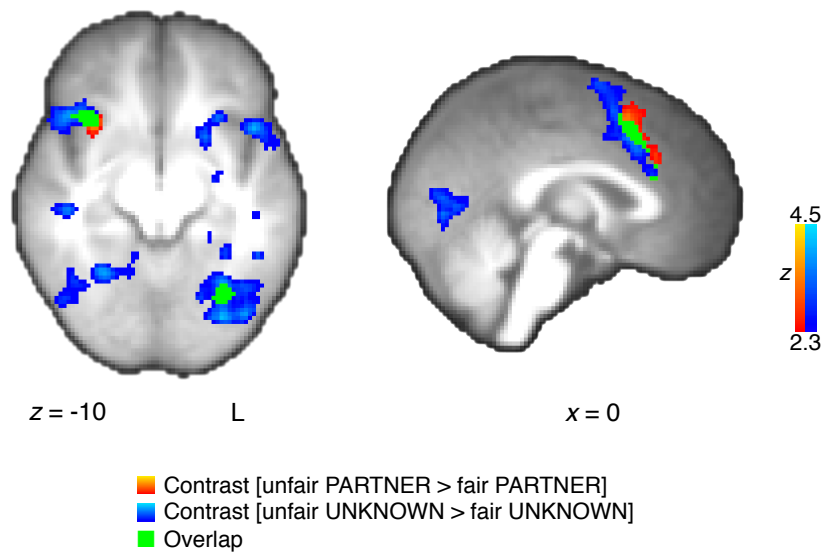


Figure 5

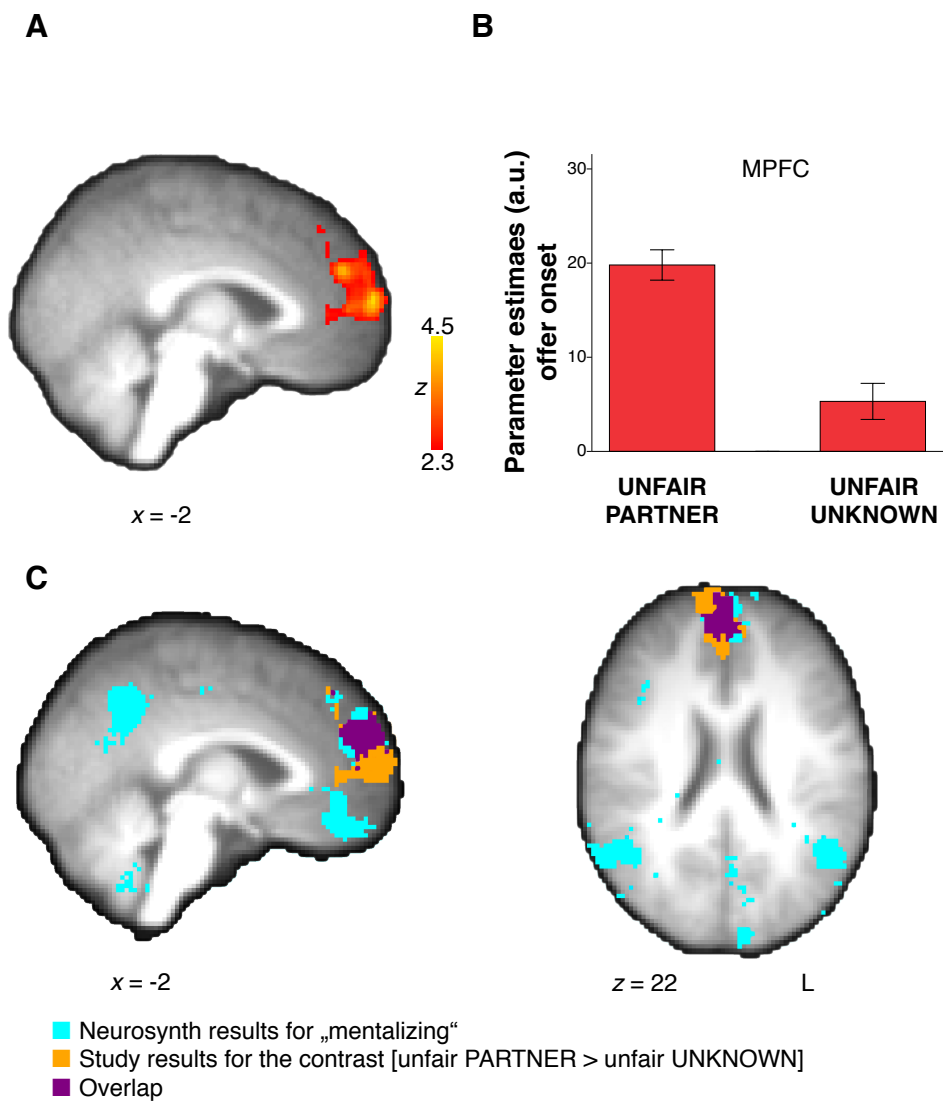


Figure 6

