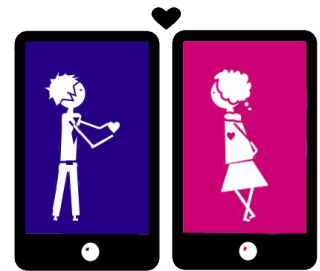

**Taking a Closer Look at the Ups and
Downs in Couple Relationships: Using the
Experience Sampling Method to Study
Motivation, Behavior, and Feelings in Dyads**



**Caroline
Zygar-Hoffmann**

München, 2020



**Taking a Closer Look at the Ups and Downs in Couple Relationships:
Using the Experience Sampling Method to Study Motivation,
Behavior, and Feelings in Dyads**

Inauguraldissertation
zur Erlangung des Doktorgrades der Philosophie
an der Ludwig-Maximilians-Universität München



vorgelegt von
Caroline Zygar-Hoffmann
aus München

2020

Erstprüfer: PD Dr. Felix Schönbrodt

Zweitprüfer: Prof. Dr. Markus Bühner

Datum der mündlichen Prüfung: 28.05.2020

Acknowledgements

Ich hatte das Privileg meine Promotion in einer unglaublich fördernden Arbeitsumgebung mit einem riesigen sozialen Rückhalt durchzuführen.

An erster Stelle möchte ich Dir danken, Felix. Deine Rolle als Doktorvater hättest Du nicht besser machen können: Du hast mich mit Deiner fachlichen und sozialen Kompetenz in jedem Schritt begleitet, so dass ich unglaublich viel von Dir lernen konnte. Ich konnte offen mit Dir reden, Du hast meine Fragen ernst genommen, und mit mir auf Augenhöhe diskutiert. Du hast mich gefordert und gefördert, mich dazu bewegt über meine fachliche Komfortzone hinwegzublicken und dabei Vertrauen in meine Fähigkeiten gesteckt. Du hast mir den Freiraum gelassen, wo ich ihn gebraucht habe, und warst immer erreichbar, wenn es für mich notwendig war. Egal ob bei inhaltlichen, konzeptionellen, datenschutzrechtlichen, technischen oder zwischenmenschlichen Herausforderungen, Du hast die richtigen Fragen gestellt, hast Dich mit mir zusammen reingehängt und bist den Problemen dabei mit dem notwendigen Humor begegnet. Ich danke Dir aufrichtig dafür, dass Du meine Promotion auf diese menschliche und bereichernde Art mitgestaltet hast.

Ich durfte während meiner Promotion an einem Lehrstuhl arbeiten, an dem ich nicht nur Kollegen, sondern Freunde gewonnen habe. Markus, Du führst diesen Lehrstuhl so, dass Deine Mitarbeiter sich bestmöglich entfalten können. Danke, dass ich ein Teil davon sein darf, und dass Du mir die Möglichkeit geschaffen hast, hier für die nächsten Jahre zu bleiben. Ich möchte allen meinen ehemaligen und jetzigen Kollegen für die tolle Zeit danken: Für die Gespräche und Diskussionen, die gemeinsam überstandenen Krisen und den Spaß – beim Mittagessen, beim Kaffee, in Parkrunden, bei Spiele-, Kubb-, Völkerball-, und Karaoke-Abenden, im Hot-Office, im Cold-Office, im Open-Door-Office und im Broffice. Das Gemeinschaftsgefühl am Lehrstuhl trägt maßgeblich dazu bei, dass ich gerne in die Arbeit gehe. Ein besonderer Dank gilt Clemens, Flo, Felix (N.), Johannes, Philipp,

Stella und Sven, die mich von Anfang an mit offenen Armen empfangen haben. Außerdem möchte ich David, Kathryn, Lena, Laura, Ramona, Samsad, Sebastian (aus Jena) und Ricarda besonders danken: Ihr wart den Großteil meiner Promotion meine Doktoranden-Mitstreiter, habt die Höhen und Tiefen mitgemacht, und wart dabei großartige Begleiter.

Bei meinen Freunden, insbesondere bei Addy, Marco, Markus, Michi, Stefan, und Ulrike, möchte ich mich für die offenen Ohren bei Sorgen und für den verlässlichen Freizeit-Ausgleich bedanken. Eure Perspektiven auf meine Probleme, Euer Rat, Eure Hilfsbereitschaft und Euer Zuspruch haben mir bei der Lösung von Problemen geholfen und mir immerzu Zuversicht geschenkt. Die entspannende, unterhaltsame gemeinsame Zeit nach der Arbeit und während unserer Urlaube hat mir den für die Promotion nötigen Gedankenfreiraum geschenkt.

Meinen Eltern danke ich für das Vertrauen in mich und meine Entscheidungen. Euer Rückhalt und Eure Obhut bedeuten mir sehr viel. Ich möchte Euch auch ganz pragmatisch für das Dach über dem Kopf während der Promotion danken. Vincent und Olly: Ihr helft mir immer wieder dabei meine Prioritäten zu kalibrieren. Es ist schön, zwei coole Brüder zu haben.

Bei der Durchführung der zwei Studien hatte ich viel Unterstützung. Ich möchte mich bei meinen Studienteilnehmern bedanken, die sich von den technischen Problemen nicht abschrecken lassen haben, dabei Geduld und Verständnis gezeigt haben, und mir durch das Berichten aus Ihrem Leben das Schreiben dieser Doktorarbeit ermöglicht haben. Auf technischer Seite haben mich zusätzlich zu Lukas und Ludwig auch noch besonders Addy und Raoul mit Bug-Suche, Bug-Fixes und eigenen Software-Lösungen unterstützt: Vielen lieben Dank für Euren selbstlosen und hilfreichen Einsatz. Eine Vielzahl an Personen, darunter auch einige Therapeuten, haben meine Flyer verteilt und ihr soziales Netzwerk eingebunden, um damit Werbung für die Studien zu machen – vielen Dank.

Ich möchte mich bei Prof. Dr. Diana Rieger dafür bedanken, dass sie die Rolle der Drittprüferin für meine Promotion einnimmt.

Für das hilfreiche Feedback zum Mantel der Doktorarbeit danke ich ganz herzlich Flo, Lena, Laura, Marco, Philipp und Raoul. Ihr habt mich im Endspurt angefeuert, und Euch in der turbulenten Zeit vor und nach Weihnachten die Zeit genommen, meine Arbeit zu lesen und zu kommentieren. Damit habt Ihr mir einen großen Gefallen getan.

Zuletzt gebührt mein Dank meinem engsten Begleiter. Raoul, Du hast mir den nötigen Halt gegeben, um mich entfalten zu können. Du hast an mich geglaubt, mir Mut gemacht und mich motiviert. Du hast meine Bedürfnisse noch vor mir erkannt, und warst für mich da, wenn ich Dich brauchte. Du hast mir die Ruhe gegeben, wenn der Stress mich überwältigt hat. Du hast mich verstanden, wenn kein anderer es konnte. Und dabei warst Du immer so unglaublich geduldig mit mir. Danke, dass Du an meiner Seite bist.

List of Abbreviations

APIM	Actor-Partner Interdependence Model
BH	Benjamini-Hochberg
CI	Confidence Interval
CSI	Couple Satisfaction Index (global relationship satisfaction)
CWCM	Centered Within Context with reintroduction of the subtracted Means at level 2
DRM	Day Reconstruction Method
DynaMoS	Dynamics of Motive Satisfaction (theoretical process model)
ESM	Experience Sampling Method
FDR	False-Discovery Rate
HDI	Highest Density Interval
IPC	Interpersonal Circumplex
M	Mean
MLM	Multi-Level regression Model
NRQ	Negative Relationship Quality (global relationship satisfaction)
PACT	Partner-related Agency and Communion Test

pnAgency	partner-related need for Agency (implicit motive)
pnCommunion	partner-related need for Communion (implicit motive)
pnIndependence ...	partner-related need for Independence (implicit motive)
pnPower	partner-related need for Power (implicit motive)
PNRQ	Positive-Negative Relationship Quality (global relationship satisfaction)
PRQ	Positive Relationship Quality (global relationship satisfaction)
RQ	Research Question
SD	Standard Deviation
SE	Standard Error
SEM	Structural Equation Modeling
SSL	Secure Sockets Layer
WEIRD	Western, Educated, Industrialized, Rich, and Democratic

Contents

	Page
Acknowledgements	i
List of Abbreviations	iv
Chapter	
1 Deutsche Zusammenfassung	5
1.1 Einleitung	5
1.2 Paper 1	6
1.3 Paper 2	7
1.4 Paper 3	8
1.5 Diskussion	10
1.6 Eigener Beitrag zu Inhalt und Umfang	12
1.6.1 Paper	12
1.6.2 Datensatzveröffentlichungen	14
2 General Abstract	15

3	General Introduction	17
3.1	Research on the Determinants of Global Relationship Satisfaction	18
3.2	Challenges in the Field	20
3.2.1	Challenge A: Selected Drawbacks of Common Study Designs	20
3.2.1.1	Drawbacks of classical observational methods: feasibility and ecological validity	21
3.2.1.2	Drawback of cross-sectional studies: Simpson's Paradox	21
3.2.1.3	Drawback of long-term longitudinal studies: the time scale of processes	22
3.2.2	Challenge B: The Relation Between Dispositions and States	22
3.2.3	Challenge C: Reliability and Validity of Global Self-Reports	23
3.3	Solutions Provided by Dyadic ESM Studies and Their Application in the Current Dissertation	24
3.3.1	Solution A: Large-Scale Study of Short-Term Intra- and Interpersonal Processes in Everyday Life	24
3.3.1.1	Using smartphones to study processes in a natural context	24
3.3.1.2	Using multilevel models to separate between- and within-person effects	26
3.3.1.3	Consideration of short-term fluctuations in within-person variables	27
3.3.2	Solution B: Comparing Dispositions and Aggregated States	28
3.3.2.1	Motives and motivation	28
3.3.2.2	Relationship satisfaction	29
3.3.3	Solution C: Reducing Retrospective Errors and Obtaining Partner-Reports	29
3.4	Overview of the Current Research: Testing Theory and Examining Measures Transparently	30

4	Paper 1: From Motive Dispositions to States to Outcomes: An Intensive Experience Sampling Study on Communal Moti- vational Dynamics in Couples	33
	Abstract	36
	Introduction	36
	Method	47
	Results	53
	Discussion	60
	References	69
5	Paper 2: Motivated Behavior in Intimate Relationships: Com- paring the Predictive Value of Motivational Variables	79
	Abstract	82
	Introduction	82
	Method	94
	Results	101
	Discussion	111
	References	118
6	Paper 3: Recalling Experiences: Looking at Momentary, Ret- rospective and Global Assessments of Relationship Satisfaction	125
	Abstract	128
	Introduction	128
	Study 1: Method	136
	Study 2: Method	143
	Results of Both Studies	148
	Discussion	160
	References	174

7	General Discussion	185
7.1	Summary of Results	186
7.1.1	The DynaMoS Model (Paper 1 & Paper 2)	186
7.1.2	Comparison of Dispositions and States (Paper 2 & Paper 3)	187
7.2	General Limitations	189
7.2.1	Challenges of the Experience Sampling Method	189
7.2.2	Constraints on Generalizability	190
7.2.3	Replicability of Effects	193
7.3	Contributions and Future Directions	194
7.3.1	Motivational Psychology	194
7.3.2	Relationship Research and Psychological Methods	197
7.3.3	Practitioner Settings	199
7.4	Conclusion	201
8	References	203

Kapitel 1

Deutsche Zusammenfassung

1.1 Einleitung

Die vorliegende Dissertation beschäftigt sich mit der Frage, welche Erkenntnisse über Partnerschaften gewonnen werden können, wenn intraindividuell variierende Erlebnis- und Situationszustände betrachtet werden, anstelle von oder zusätzlich zu in der Psychologie üblicheren, globaleren Maßen. Dafür wurden zwei dyadische Studien mit der sogenannten „Experience Sampling Methode“ (ESM) durchgeführt: Bei dieser Methode werden Personen in ihrem Alltag mehrmals zu ihren aktuellen Erfahrungen, Verhaltensweisen und Gefühlen befragt (Studie 1: $N_{ESM} = 130$ Teilnehmer, 2 Wochen, 5 Befragungen pro Tag, insgesamt 7573 beantwortete ESM Messungen, Studie 2: $N_{ESM} = 510$ Teilnehmer, 4 Wochen, 5 Befragungen pro Tag, insgesamt 60942 beantwortete ESM Messungen).

Ein solches Studiendesign bietet viele Vorteile im Vergleich zu anderen in der Partnerschaftsforschung häufig vorkommenden Designs: Die Durchführung von ESM-Studien mit großen Stichproben ist dank der weiten Verbreitung von Smartphones ökonomisch umsetzbar, und erlaubt es Schlüsse über den tatsächlichen Alltag der Studienteilnehmer zu ziehen. Mit Hilfe von statistischen Multilevel-Modellen können Effekte, die auf Unterschiede zwischen Personen zurückzuführen sind, getrennt betrachtet werden von Effekten die auf Schwankungen der Erlebnisse einer Person basieren. Außerdem können zeitlich kurzfristig ablaufende Dynamiken und deren Zusammenhänge untersucht werden, die sich beispielsweise aus dem Zusammenspiel von psychologischen Erlebniszuständen und variierenden Situationseinflüssen ergeben. Die resultierenden Daten erlauben darüber hinaus, konzeptuelle Fragen

zum Zusammenhang von Dispositionsmaßen und Zustandsmaßen zu beantworten. Letztendlich ermöglicht dieses Studiendesign auch, den Einfluss von Erinnerungsfehlern bei der Messung zu verringern, sowie Partnerberichte als in mancher Hinsicht weniger fehlerbehaftete, zusätzliche Informationsquelle zu nutzen.

Die Dissertation umfasst drei Paper, die sich diese Vorteile zu Nutze machen, um ein Erklärungsmodell für den Zusammenhang zwischen Motiven und Beziehungszufriedenheit zu untersuchen, und um diese beiden zentralen Konstrukte in der Motivationspsychologie und Beziehungsforschung besser zu verstehen. Zum einen werden motivationale Prozesse untersucht, die sich innerhalb einer Person abspielen, und in Kombination mit den interpersonellen Erfahrungen mit dem Partner die eigene Beziehungszufriedenheit beeinflussen können (Paper 1 und Paper 2). Zum anderen werden die Ergebnisse von ESM-Messungen von motivationalen Zuständen und momentaner Beziehungszufriedenheit mit jenen von rückblickenden und/oder globaleren Maßen verglichen (Paper 2 und Paper 3). Alle drei Veröffentlichungen folgen dabei den Prinzipien transparenter Wissenschaft.

1.2 Paper 1

From motive dispositions to states to outcomes: An intensive experience sampling study on communal motivational dynamics in couples (Von Motivdispositionen zu Zuständen zur Wirkung: Eine umfangreiche Experience-Sampling-Studie zu kommunalen Motivationsdynamiken in Paaren)

In diesem Paper wird ein theoretisch fundiertes, präregistriertes Prozessmodell vorgestellt und empirisch anhand der Daten der ersten Studie für Nähe als Motivdomäne geprüft. Das Modell hat den Anspruch, eine Erklärung für in der Forschungsliteratur beschriebene interindividuelle Zusammenhänge zwischen Motivdispositionen und der Beziehungsqualität zu liefern. Es beschreibt, a) den funktionalen Zusammenhang zwischen Motivdispositionen und motivationalen Zuständen; b) die Rolle von motivationalen Zuständen und Situationseinflüssen für motiv-kongruentes, zielführendes Verhalten; c) eine angenommene affektverstärkende Wirkung von motivationalen Zuständen, sowie d) einen Zusammenhang zwischen momentaner Beziehungszufriedenheit und globaler Beziehungszufriedenheit. Die Affektverstärkung ist zentraler Bestandteil des Modells, und postuliert

das Zusammenspiel von Motivation und bestimmten Beziehungserfahrungen für das Erleben von momentaner Beziehungszufriedenheit und positivem Affekt: Es wird davon ausgegangen, dass bestimmte Beziehungserfahrungen besonders positiv erlebt werden, wenn sie zu der Motivation passen, die Personen zuvor erlebt haben.

Das Paper konnte zunächst den interindividuellen Effekt aus der Literatur replizieren, dass ein hohes implizites Nähemotiv mit einer höheren Beziehungszufriedenheit zusammenhängt. Weiterhin sind die Ergebnisse von Multilevel-Modellen größtenteils im Einklang mit dem postulierten Prozessmodell: Je stärker das implizite Nähemotiv bei Personen ausgeprägt ist, desto häufiger erleben sie die akute Motivation Nähe zu ihrem Partner aufzubauen. Auf intraindividuelle Ebene ist das Erleben von Nähe-Motivation nachfolgend mit dazu passendem Verhalten assoziiert: Nachdem Personen einen stärkeren Wunsch nach Nähe verspüren als für sie üblich ist, berichten sie mehr eigenes Verhalten, das potenziell zu dem Erleben von Nähe zum Partner führt (z.B. Zuneigung). Auch für die affektverstärkende Wirkung von Motivation findet sich Evidenz in Hinblick auf Beziehungszufriedenheit als abhängige Variable: Die momentane Zufriedenheit mit der Beziehung ist dann besonders hoch, wenn Personen sich Nähe zum Partner gewünscht haben und daraufhin Erlebnisse mit dem Partner hatten, die durch hohe Nähe gekennzeichnet waren. Für eine verstärkende Wirkung von aktuell erlebtem positiven Affekt als abhängige Variable gab es jedoch keine Evidenz. Zuletzt wurde auch der postulierte Zusammenhang zwischen durchschnittlich angegebener momentaner Beziehungszufriedenheit und globaler Beziehungszufriedenheit gefunden. Die Bedeutung und Grenzen des Modells für die Erklärung von Unterschieden in der Beziehungszufriedenheit von Personen werden diskutiert.

1.3 Paper 2

Motivated behavior in intimate relationships: Comparing the predictive value of motivational variables (Motiviertes Verhalten in engen Beziehungen: Ein Vergleich des Vorhersagewerts von motivationalen Variablen)

Der Zusammenhang zwischen motivationalen Variablen und Verhaltensberichten wurde in diesem Paper anhand der Daten der zweiten Studie aus mehreren Perspektiven betrachtet. Zusätzlich zu aktueller Motivation wurden als Prädiktoren

implizite und explizite Motivdispositionen untersucht, sowie auch innerhalb von Personen aggregierte motivationale Zustände. Die Effekte wurden diesmal nicht nur für Nähe, sondern auch für Macht und Unabhängigkeit als weitere relevante Motivdomänen in Partnerschaften geprüft. Außerdem wurde nicht nur selbstberichtetes Verhalten als Kriterium herangezogen, sondern auch die Verhaltensberichte durch den Partner. In der Veröffentlichung wird verglichen, welche motivationalen Variablen einen signifikanten Haupteffekt für Verhaltensberichte zeigen, und wie hoch deren absolute und inkrementelle Effektstärken sind. Die Ergebnisse zeigen zunächst eine direkte Replikation des Pfads von Nähe-Motivation zu zielführendem Verhalten, als auch eine konzeptuelle Replikation für andere Operationalisierungen von Motivation und Verhalten, für Partner-Berichte, sowie für Macht- und Unabhängigkeits-Motivation und entsprechendes Verhalten. Alle untersuchten motivationalen Variablen waren für mindestens ein Kriterium signifikante Prädiktoren, und zeigten in mindestens einer Motivdomäne einen inkrementellen Beitrag über die anderen motivationalen Variablen hinaus. Dennoch stechen die Effekte von aktueller Motivation und aggregierter Motivation heraus, da sie konsistent in allen Motivdomänen für fast alle Verhaltens-Kriterien, unabhängig davon ob selbst- oder fremdberichtet, signifikante und inkrementelle Haupteffekte zeigen. Es wird diskutiert, welche Vorteile die Erhebung von motivationalen Zuständen für die Forschung mit sich bringt.

1.4 Paper 3

Recalling experiences: Looking at momentary, retrospective and global assessments of relationship satisfaction (Die Erinnerung an Erfahrungen: Ein Blick auf momentane, retrospektive und globale Einschätzungen von Beziehungszufriedenheit)

In diesem Paper wird Beziehungszufriedenheit als zentrales Konstrukt in der Partnerschaftsforschung anhand der Daten beider Studien untersucht. Beziehungszufriedenheit liegt in den Daten auf dreifache Weise vor: Als globale, generelle Bewertung der Beziehung, als momentane ESM-Messung, und als rückblickende Einschätzung der Beziehungszufriedenheit während des Studienzeitraums. Es gibt verschiedene Möglichkeiten alle momentanen ESM-Messungen einer Person für ein Konstrukt zu einem einzelnen Wert zusammenzufassen. In dem Paper

werden der Mittelwert (über den ganzen Zeitraum, während der ersten und letzten Woche, und während des letzten Tages), der Median, das 0.1-Quantil, sowie das 0.9-Quantil betrachtet, und die Zusammenhänge von den auf verschiedene Weise aggregierten Variablen mit der retrospektiven und globalen Einschätzung berechnet. Dabei zeigt sich, dass der Mittelwert über alle momentanen Einschätzungen zur Beziehungszufriedenheit deskriptiv der beste Prädiktor für globale und retrospektive Einschätzungen ist, andere Maße aber teilweise nur unbedeutend schlechter sind. In anderen Worten ausgedrückt zeigen die Ergebnisse, dass sowohl generelle als auch retrospektive Einschätzungen der eigenen Beziehungszufriedenheit durch das typische Erleben von Beziehungszufriedenheit im Alltag gekennzeichnet sind.

Weiterhin wird in dem Paper untersucht, inwiefern sich die retrospektive Einschätzung im Durchschnitt vom Mittelwert der momentanen Einschätzungen unterscheidet: Die Ergebnisse zeigen, dass im Rückblick durchschnittlich ein bestimmter Aspekt von Beziehungszufriedenheit überschätzt wird, nämlich die Intensität mit der die Personen von Ihrem Partner genervt waren. Dies führt in der Folge zu einer für die Stichprobe durchschnittlich geringeren Beziehungszufriedenheit im Rückblick im Vergleich zum Mittelwert der momentanen Beziehungszufriedenheitsmessungen. Eine Analyse von potentiellen moderierenden Einflussfaktoren zeigt, dass die globale Beziehungszufriedenheit zum Zeitpunkt des Rückblicks eine wesentliche Rolle für den Unterschied spielt: Personen, die generell eine geringere Beziehungszufriedenheit angeben, berichten im Rückblick eine stärkere Genervtheit von ihrem Partner als sie im Mittel in den momentanen ESM-Messungen angegeben haben. Außerdem gibt es eine Reihe weiterer Faktoren, die einen moderierenden Einfluss haben (u.a. Lebenszufriedenheit, Bindungsstile, Persönlichkeitsaspekte). Diese leisten jedoch zum Teil keinen inkrementellen Beitrag mehr, wenn eine Variable im Modell aufgenommen wird, die durch generelle Positivität oder soziale Erwünschtheit beschrieben werden könnte. Diese Ergebnisse erweitern bisherige Befunde in der Literatur welche zeigen, dass Personen im Rückblick im Vergleich zu gemittelten momentanen Messungen einen stärkeren negativen Affekt berichten.

Zuletzt wird in dem Paper noch geprüft, wie stark der Zusammenhang zwischen aggregierten ESM-Messungen und globaler Beziehungszufriedenheit über die Dauer der Studie ansteigt und welchen Einfluss eine unterschiedliche Anzahl an Messungen pro Tag hat. Hierzu konnte festgehalten werden, dass nach etwa zwei Wochen ein Deckeneffekt eintritt: Nach diesem Zeitraum steigt der Zusammen-

hang mit globaler Beziehungszufriedenheit nicht mehr bedeutsam an, was darauf hinweist, dass nach etwa zwei Wochen die maximal erreichbare Repräsentativität der ESM-Messungen für die globale Messung erreicht wurde. Es zeigt sich auch, dass eine unterschiedliche Anzahl an täglichen Messungen nur für die ersten Studientage relevant ist. Es wird diskutiert wie diese Ergebnisse für eine ökonomische Planung von Studien genutzt werden können.

1.5 Diskussion

Alle Veröffentlichungen sind von generelle Limitationen betroffen, die aus dem angewandten ESM Studiendesign resultieren. Zudem ergeben sich Grenzen in Bezug auf die Generalisierbarkeit der Ergebnisse, und hinsichtlich erster weiterführender Untersuchungen zur Replizierbarkeit. Eine generelle Einschränkung der Methode begründet sich darin, dass das das mehrmalige Befragen von Personen zu Reaktivität (d.h., zu einer Änderung der zu beobachtenden Variablen) und Ermüdungseffekten führen kann. Diese hätten in der Folge möglicherweise Einfluss auf die Reliabilität, Validität und Repräsentativität der Messungen, sowie auf die Interpretation der Ergebnisse.

Hinsichtlich der Generalisierbarkeit der Ergebnisse ist festzuhalten, dass beide Studien sich in der Zusammensetzung der Stichprobe in einigen Aspekten stark unterscheiden. Replizierte Effekte in beiden Stichproben legen daher zunächst nahe, dass eine Generalisierbarkeit dieser Ergebnisse auf die Population möglich ist, die durch beide Stichproben beschrieben wird. Eine repräsentative Zufallsstichprobe wäre jedoch nötig, um dies zu untermauern. Aus theoretischer Sicht sollte das aufgestellte Prozessmodell für Paarbeziehungen im Allgemeinen, sowie auch für verschiedene Motivdomänen gelten. Die Ergebnisse zur Beziehungszufriedenheit sollten größtenteils ebenfalls verallgemeinerbar auf Erwachsene in Beziehungen sein. Andere Studienergebnisse legen jedoch nahe, dass für Rückschaueffekte das Alter der Personen und in mancher Hinsicht auch die Beziehungsdauer eine Rolle spielen, was die Generalisierbarkeit über die untersuchten Stichproben hinaus einschränken würde. Zudem könnte die Operationalisierung von Beziehungszufriedenheit für die aktuellen Ergebnisse eine große Rolle spielen, und die gefundenen Rückschaueffekte daher nicht verallgemeinerbar sein für Beziehungszufriedenheitsmaße, die sich anders zusammensetzen (z.B. solche, die keine negativen Aspekte von Beziehungen abfragen).

Erste weiterführende Analysen zur Replizierbarkeit der in der Dissertation gefundenen Effekte zeigen bestimmte Einschränkungen, vor allem hinsichtlich der Affektverstärkung von Motivation und der Generalisierbarkeit aller Pfade des Prozessmodells auf Unabhängigkeits- und Machtmotivation. Die Robustheit der dargelegten Effekte sollte daher kritisch betrachtet werden bis weitere Replikationen überzeugende Evidenz dafür bereitstellen.

Die vorliegende Dissertation trägt zu verschiedenen Forschungsbereichen bei: Zur Grundlagenforschung in der Motivationspsychologie, zu angewandter Beziehungsforschung, und zur Forschung zu psychologischen Methoden. Für die Motivationspsychologie wird zunächst aufgezeigt, dass der Zusammenhang von Motiven und Beziehungszufriedenheit robust replizierbar ist, und somit ein Forschungsergebnis darstellt, auf dem in Zeiten der Replikationskrise aufgebaut werden kann. Außerdem wird die Rolle von bewussten motivationalen Zuständen auch für implizite, nicht bewusst zugängliche Motive deutlich. Die Arbeit prüft grundlegende theoretische Annahmen der Motivationspsychologie auf dem intraindividuellen Analyselevel, und trägt damit zu einem besseren Verständnis der postulierten Mechanismen bei. Der Vergleich verschiedener motivationaler Variablen für die Vorhersage von Verhaltensberichten stellt die Relevanz der unterschiedlichen, untersuchten Motiv- und Motivations-Maße dar.

Die Beziehungsforschung profitiert von einem motivationalen Blick auf die intra- und interpersonellen Prozesse, die zur Erklärung von Beziehungszufriedenheit herangezogen werden. Zudem liefert die ausführliche Analyse zu Beziehungszufriedenheit auf verschiedenen Messebenen eine empirische Grundlage, um konzeptionelle Fragen zu dem Konstrukt zu klären. Auch konkrete Forschungsdesigns können basierend auf den vorliegenden Erkenntnissen besser geplant werden, hinsichtlich der Bedeutung verschiedener Maße, und der nötigen Dauer von ESM-Studien.

Ein Beitrag zur Forschung zu psychologischen Methoden wird durch die Analyse von Unterschieden zwischen den verschiedenen Maßen der Beziehungszufriedenheit geleistet. Wenn im Rückblick bestimmte Aspekte anders bewertet werden als im Moment des Erlebens, dann kann dies auf einen Messfehler hindeuten, der in weiterführenden Analysen berücksichtigt werden kann. Da die ESM jedoch nicht alle Erlebnisse erfasst, und eine bestimmte Aggregation der einzelnen Messungen gewählt werden muss (die von der von den Versuchspersonen implizit durchgeführten Aggregation im Rückblick abweichen kann), sind diese Unterschiede zwischen den Maßen nicht zwingend als Messfehler zu betrachten. Die Konsequenz solcher

Unterschiede muss dann in Hinblick auf die Validität der Maße anerkannt werden. Denkbare praktische Implikationen sollten sowohl eine Replikation, eine kausale Prüfung der Effekte und dazu passende gezielte Interventionsstudien abwarten. Dann könnten diverse therapeutische Maßnahmen aus den Ergebnissen abgeleitet werden, die sich durch eine Integration einer motivationalen Perspektive auf Partnerschaftsprozesse auszeichnen und Erinnerungsverzerrungen berücksichtigen.

Die Transparenz der vorliegenden Dissertation trägt zu einer kumulativen Forschungslinie bei, indem sie die Präregistrierungen, reproduzierbaren Analyseskripte, Studienmaterialien und wissenschaftlich zugängliche Studiendaten bereitstellt und damit ermöglicht auf diesen Ergebnissen aufzubauen, sie zu re-analysieren und konzeptuelle sowie direkte Replikationsstudien durchzuführen. Damit leistet die Dissertation einen Beitrag zum Fortschritt der Psychologie als offene Wissenschaft und als ernstzunehmende Disziplin in Zeiten der Replikationskrise.

1.6 Eigener Beitrag zu Inhalt und Umfang

1.6.1 Paper

Paper 1: Zygar, C., Hagemeyer, B., Pusch, S., & Schönbrodt, F. D. (2018). From motive dispositions to states to outcomes: An intensive experience sampling study on communal motivational dynamics in couples. European Journal of Personality, 32, 306–324. <https://doi.org/10.1002/per.2145>

Autoren: Caroline Zygar (CZ), Birk Hagemeyer (BH), Sebastian Pusch (SP), Felix D. Schönbrodt (FS)

Die Autorin dieser Dissertation ist Erstautorin dieses Papers. Alle Autoren haben zu Konzeption und Design beigetragen. CZ hat die Daten erhoben und gemeinsam mit FS analysiert und interpretiert. CZ hat den Entwurf für diesen Artikel geschrieben, sowie Kommentare und Überarbeitungen von allen Autoren eingearbeitet.

Paper 2: Zygar-Hoffmann, C., Pusch, S., Hagemeyer, B., & Schönbrodt, F. D. (major revision). Motivated behavior in intimate relationships: Comparing the predictive value of motivational variables. Social Psychological Bulletin.

Autoren: Caroline Zygar-Hoffmann (CZ-H), Sebastian Pusch (SP), Birk Hagemeyer (BH), Felix D. Schönbrodt (FS)

Die Autorin dieser Dissertation ist Erstautorin dieses Papers. Alle Autoren haben zu Konzeption und Design beigetragen. CZ-H hat die Daten erhoben und gemeinsam mit FS analysiert und interpretiert. CZ-H und FS haben den Entwurf für diesen Artikel geschrieben, sowie Kommentare und Überarbeitungen von allen Autoren eingearbeitet.

Paper 3: Zygar-Hoffmann, C., & Schönbrodt, F. D. (2020). Recalling experiences: Looking at momentary, retrospective and global assessments of relationship satisfaction. Collabra: Psychology, 6(1), Article 7. <https://doi.org/10.1525/collabra.278>

Autoren: Caroline Zygar-Hoffmann (CZ-H), Felix D. Schönbrodt (FS)

Die Autorin dieser Dissertation ist Erstautorin dieses Papers. Beide Autoren haben zu Konzeption und Design beigetragen. CZ-H hat die Daten erhoben und gemeinsam mit FS analysiert und interpretiert. CZ-H hat den Entwurf für diesen Artikel geschrieben, sowie Kommentare und Überarbeitungen von FS eingearbeitet.

1.6.2 Datensatzveröffentlichungen

Studie 1: Zygar, C., Hagemeyer, B., Pusch, S., & Schönbrodt, F. D. (2018). From motive dispositions to states to outcomes: Research data of an intensive experience sampling study on communal motivational dynamics in couples [Translated title] (Version 2.0.0) [data and documentation]. Trier, Germany: Center for Research Data in Psychology PsychData of the Leibniz Institute for Psychology Information ZPID. https://doi.org/10.5160/psychdata.zrce16dy99_v20000

Studie 2: Zygar-Hoffmann, C., Hagemeyer, B., Pusch, S., & Schönbrodt, F. D. (2020). A large longitudinal study on motivation, behavior and satisfaction in couples: Research data from a four-week experience sampling study with a pre-, post-, and one-year follow-up-assessment [Translated title] (version 1.0.0) [data and documentation]. Trier, Germany: Center for Research Data in Psychology PsychData of the Leibniz Institute for Psychology Information ZPID. <https://doi.org/10.5160/psychdata.zrce18mo99>

Autoren: Caroline Zygar/-Hoffmann (CZ/CZ-H), Birk Hagemeyer (BH), Sebastian Pusch (SP), Felix D. Schönbrodt (FS)

Die Autorin dieser Dissertation ist Erstautorin bei beiden Datensatzveröffentlichungen. CZ/CZ-H war für die Aufbereitung und Dokumentation der Daten verantwortlich. SP hat Informationen zur Kodierung der impliziten Motive bereitgestellt. Alle Autoren haben die Daten zur Veröffentlichung freigegeben.

Chapter 2

General Abstract

Most people consider a satisfying couple relationship an important goal in their life. Past research reveals that relationship satisfaction is a multi-determined phenomenon which is related to various desirable outcomes. It differs as a global, generalized relationship evaluation between individuals, but also on a momentary basis as a state between situations for a single individual. This dissertation investigates motivational processes that precede the experience of momentary instances of relationship satisfaction and illustrates how such state relationship evaluations relate to global and retrospective relationship satisfaction. Thereby, this research examines and demonstrates the added value of assessing states beyond more traditional measures in psychology for the study of couple relationships. Two studies applying the Experience Sampling Method were conducted, in which both members of a committed, heterosexual relationship reported on their motivational, behavioral, affective and satisfaction states five times a day for two and four weeks. Additionally, person and relationship characteristics as well as retrospective reports on the study period were obtained. The data of these studies was analyzed in three papers that constitute the central content of this dissertation. The results shed light on the behaviors, functioning and evaluations in couple relationship through the investigation of states and processes in individuals' everyday lives.

The papers are introduced by an overview of the covered topics, and by a description on how the presented research follows from general approaches to explain relationship satisfaction with interpersonal processes and between-person differences. It is elaborated how challenges in the field are tackled with the employed dyadic experience sampling study design and how this relates to the presented

papers.

Paper 1 focuses on the description and empirical examination of a motivational process model, which is proposed as potential explanation for how individual's motive dispositions influence their global relationship satisfaction. The results provide evidence for the model in the domain of communion motives.

Paper 2 scrutinizes one part of the process model by conducting a direct replication, and many conceptual replications of the association between motivational states and behavior, revealing strong supporting evidence for this within-person effect. For the prediction of behavior levels between persons, this paper additionally compares the predictive value of aggregated motivational states, explicit and implicit motives. The results demonstrate the relevance of assessing motivational states not only for within-person analyses, but also for between-person analyses of behavior occurring in couple relationships.

Paper 3 focuses on relationship satisfaction as an important outcome variable for the current research and for a large body of relationship research in general. Aggregated state assessments of relationship satisfaction are compared with retrospective and global assessments. The results indicate that the average state relationship satisfaction is related to global relationship satisfaction, but that in retrospection a negative mean-level bias emerges. Different relationship and person characteristics are identified as moderators of this bias.

Finally, the goals and findings from the presented papers are summarized in a holistic discussion. Some general limitations of this dissertation are illustrated: reactivity and fatigue as methodological side effects of the Experience Sampling Method, constraints on the generalizability of the findings, and preliminary results on the non-replicability of some of the presented effects. The contributions of this dissertation for motivational psychology, relationship research, psychological methods and practical settings are highlighted, along with a discussion of implications for future research.

Adhering to open science principles, the current work provides two large dyadic experience sampling data sets – both published as scientific use-files – as well as transparent, reproducible, partly preregistered analyses on important topics in several psychological research domains.

Chapter 3

General Introduction

Repeatedly asking individuals about their experiences, behavior, and feelings on momentary occasions in their daily life instead of relying on a single snapshot: This is the central idea of the so-called Experience Sampling Method (ESM). Since its first introduction in the 1970s, the ESM gained large popularity in several research disciplines, with psychology leading the way (Hektner, Schmidt, & Csikszentmihalyi, 2007). The promised advantages compared to traditional research methods include the opportunity to study short-term processes on a large scale and in an ecologically valid manner in everyday life, a better understanding of dispositions through the examinations of states, and a reduction of retrospective biases during measurement (see e.g., Bolger, Davis, & Rafaeli, 2003).

This dissertation covers three empirical papers, which rely on data from two dyadic studies that applied the ESM. The papers draw upon and examine the mentioned benefits of the ESM, by investigating processes derived from theories in motivational psychology (Paper 1 and Paper 2), and by critically contrasting the results obtained from ESM measures with retrospective and/or global measures (Paper 2 and Paper 3). All papers are concerned with the study of heterosexual couple relationships varying in length and commitment. An overarching research question guides the topics of the presented papers: *What can state measures tell us about different aspects of couple relationships (e.g., behaviors, processes, evaluations), compared with and in addition to traditional, more global measures used in psychology (e.g., dispositional, generalized, retrospective measures)?*

Although there is a declining trend in marriages in Europe across the last decades (Eurostat, 2019), having intimate social relationships is still among the most

important life goals across ages and generations (Bühler, Weidmann, Nikitin, & Grob, 2019; Twenge, Campbell, & Freeman, 2012). Relationships can be the source of positive, but also negative experiences, leading to momentary instances of relationship satisfaction and dissatisfaction. The overall quality of individuals' relationships is related to various important outcomes, such as life satisfaction, mental health and well-being (Gómez-López, Viejo, & Ortega-Ruiz, 2019; Proulx, Helms, & Buehler, 2007), physical health (Robles, Slatcher, Trombello, & McGinn, 2014; Slatcher, 2010), or child behavior problems (e.g., Linville et al., 2010; see Halford, Rhoades, & Morris, 2018; Vaez, Indran, Abdollahi, Juhari, & Mansor, 2015 for reviews).

Some relationships break up after some time, while others show high stability. The overall, global level of relationship satisfaction is considered as one of the most important predictors for the stability of relationships (Bescheid & Lopes, 1997; Karney & Bradbury, 1995). A study by Anderson, Van Ryzin, & Doherty (2010) investigated trajectories of marital satisfaction assessed across 50 years and shows that most marriages are characterized by a stable, high relationship satisfaction across time; three other distinct trajectories emerged, which can be categorized by being constantly low, declining or curvilinear (see also Eastwick, Finkel, & Simpson, 2019; Lavner, Bradbury, & Karney, 2012). Hence, the question arises which factors determine whether individuals experience, maintain, or restore high global relationship satisfaction.

3.1 Research on the Determinants of Global Relationship Satisfaction

Scholars have approached the question on the determinants of globally satisfying relationships from various perspectives (see Bradbury, Fincham, & Beach, 2000; Finkel, Simpson, & Eastwick, 2017 for reviews): A large body of research addresses the interpersonal processes that occur in relationships as the underlying basis of relationship (dis-)satisfaction (e.g., behaviors and interaction patterns, cognitive processes, or affect dynamics). For instance, a pattern of one partner showing demanding behavior (e.g., complaints) while the other shows withdrawal behavior (e.g., avoiding the topic) is associated with detrimental relationship outcomes (“demand-withdraw pattern”, Baucom et al., 2015). When trying to understand the causal mechanisms that influence individuals' behaviors, experiences and feel-

ings, the study of such processes is central. As Back & Vazire (2015) put it in a nutshell: “[...] more process knowledge is needed because it tells us how humans actually navigate their real-life social environments – and how some people do so more successfully than others. It gives us insight into the more proximal causes of social outcomes” (p. 300). Processes are hard to study using between-person analyses as they are inherently tied to the psychological sequences in the unit of consideration: For intrapersonal processes the variation on the within-person level provides the necessary information, for interpersonal processes in relationships the variation on the within-couple level is crucial. In other words, intrapersonal processes aim to describe how variations of individuals’ own internal states can explain their subsequent thoughts, behaviors or feelings. Interpersonal processes aim to describe how variation in the interactions between the thoughts, behaviors, and feeling of different individuals can explain certain outcomes (such as the relationship satisfaction of both individuals). In consequence, when looking at couples, the (intrapersonal) variation of each partner’s states lay the foundation of different interpersonal processes. Hence, in a first step, the longitudinal measurement of variables varying within persons and couples on a short-time scale is necessary to study processes. In a next step, to warrant causal claims about short-term longitudinal effects found in such data, experimental designs on the within-person or within-couple level are needed (e.g., Schmiedek & Neubauer, 2019). Revealing the causal processes that promote relationship satisfaction provides crucial information for interventions, but it does not clarify the role of individual or couple characteristics for differences in the occurrence of states and corresponding processes (see e.g., Caughlin & Vangelisti, 2000 for an analyses of various individual differences predicting the demand-withdraw pattern).

A diverse range of between-person or between-couple differences has been considered in the literature as potential influences on such processes in general, but also as independent determinants of relationship satisfaction (Bradbury et al., 2000; Finkel et al., 2017). These include individual characteristics (e.g., personality, attachment styles, personal histories, or motives) and differences elicited by situational and/or contextual influences (e.g., external demands or stressors, the attractiveness of alternative relationship forms or partners, or cultural, social and economic factors). For instance, individuals’ motive dispositions have been identified in the literature as important determinants for relationship satisfaction (e.g., Hagemeyer & Neyer, 2012): Whereas communal motives for closeness and unity are positively associated with relationship satisfaction, there is a negative asso-

ciation for agentic motives, such as power and independence. The vulnerability-stress-adaptation model by Karney & Bradbury (1995) attributes a central role to such between-person differences (in terms of “enduring vulnerabilities”) for the emergence of and coping with stressful events in marriages (such as conflicts), which in turn is supposed to affect marital quality. The authors of the model emphasize the need for longitudinal next to cross-sectional research to investigate how changes in between-person or between-couple factors affect changes in relationship quality (i.e., long-term within-person/-couple processes). This line of research demonstrates the relevance of considering the broader idiographic and contextual system couples are situated in to obtain a holistic understanding of the emergence of relationship (dis-)satisfaction.

The current work draws on the perspective that interpersonal processes are determinants of relationship satisfaction, which are influenced by certain dispositions of the partners. Of particular interest are the predispositions individuals bring to the relationship in the form of their motives and the expected influence of these motives on within-person processes.

3.2 Challenges in the Field

Psychological research in general, and in extension also the study of relationships, faces some challenges which are associated with (A) the study designs that are most commonly applied, (B) the conceptual nature of dispositions in relation to states, and (C) problems surrounding the reliability and validity of self-reports. Before illustrating how these challenges can be tackled in part by using dyadic ESM study designs, the problems associated with each challenge will first be presented in the following sections.

3.2.1 Challenge A: Selected Drawbacks of Common Study Designs

Different study designs have different benefits and drawbacks that should be considered when interpreting results from such studies.

3.2.1.1 Drawbacks of classical observational methods: feasibility and ecological validity

Interpersonal, short-term processes are frequently studied using classical behavioral observation methods, for instance by inviting couples to the laboratory and recording their interactions (see Gottman, 1998; Gottman & Notarius, 2000 for reviews). However, they are costly to conduct, and create artificial settings that in many cases do not resemble real-life situations (Laurenceau & Bolger, 2005). For example, several studies observed married couples and their interactions during the discussion of a high conflict issue in the lab (e.g., Gottman & Krokoff, 1989). This study design is time-consuming both for the participating couples, who have to appear in person, and for the researchers, who have to be present to record behavioral cues during the interaction. As a result of this effort, the researcher is rewarded with arguably more objective behavior ratings compared to self-reports from the individuals about their behavior. However, only a very specific interaction is observed, leaving merely a narrow snapshot of everyday interactions. Further, social desirability effects caused by being observed and the artificial nature of engaging in a conflict discussion without particular reason might limit the generalizability to situations outside of the laboratory.

3.2.1.2 Drawback of cross-sectional studies: Simpson's Paradox

Cross-sectional study designs are common to investigate the role of between-person factors, for example to investigate how individuals' motives are related to relationship quality (Hagemeyer & Neyer, 2012). Such cross-sectional studies are limited to the analyses of between-person differences - still, researchers and practitioners may be inclined to transfer the results of the between-person to the within-person level, for instance, to assume that variation in individuals' state motivation is similarly related to their state relationship quality (see Fisher, Medaglia, & Jeronimus, 2018 for examples). However, the results of corresponding between- and within-person analyses can show discrepancies: Commonly known as the Simpson's Paradox, it is possible to observe a different association between variables when examined on the between-person level compared to the within-person level (see Kievit, Frankenhuis, Waldorp, & Borsboom, 2013 for a review). Investigations of short- or long-term within-person effects are needed to determine whether results differ between the two levels of analysis, hence avoiding incorrect conclusions (Hoffman & Stawski, 2009).

3.2.1.3 Drawback of long-term longitudinal studies: the time scale of processes

Long-term longitudinal studies (e.g., with panels) allow to investigate within-person or within-couple processes that unfold across larger time spans. However, a central challenge is to identify the time scale in which the processes of interest operate. When individuals answer surveys only once every year, then mid- and short-term processes are not visible, for example those that unfold within a few hours of a single day. Therefore, it is crucial for studies to investigate different time scales; otherwise certain processes might be missed, such as interactions between intrapersonal states and variable situations (see e.g., Kritzler, Krasko, & Luhmann, 2019 for the relevance of fit between Big Five personality states and situations for momentary evaluations of happiness; in contrast to negligible effects for the interaction of Big Five personality traits and situations, see Mueller et al., 2019; Wilt & Revelle, 2019).

3.2.2 Challenge B: The Relation Between Dispositions and States

A central question when studying within-person effects compared to between-person effects is the relationship between dispositions (or traits), and their corresponding states. Not all between-person variables have corresponding equivalents on the within-person level: For example, when looking at generational effects, the birth year might be an important variable, but this variable only varies between persons, not within a person. However, most of the between-person variables that are thought to be relevant for relationship satisfaction are associated with corresponding experiences that may vary for a single person (see Geiser, Götz, Preckel, & Freund, 2017; Steyer, Ferring, & Schmitt, 1992, for general discussions of trait and state components of psychological constructs). For example, even though most personality traits are considered as relatively stable, research suggests the presence of corresponding, momentarily varying personality states (Fleeson, 2001; Fleeson & Gallagher, 2009). Similarly, individuals who have an anxious attachment style may still feel more or less anxious about their relationship in certain situations (Gillath, Hart, Nofle, & Stockdale, 2009; Haak, Keller, & DeWall, 2017). This is also true for many outcomes in relationship research, like global relationship satisfaction (Hofmann, Finkel, & Fitzsimons, 2015). Theoretical con-

siderations and empirical work are needed to establish their correspondence: Is there a functional relationship between dispositions and states? Can dispositions be described by a summary of states? Is there more to a disposition than what is captured by the average state? Do dispositions differ in their predictive validities compared to aggregated states? Depending on the construct of interest, these questions are more or less tackled from theoretical perspectives, and empirical investigations are difficult to realize with the study designs reviewed so far.

3.2.3 Challenge C: Reliability and Validity of Global Self-Reports

Psychological research often relies on global self-report measures, with studies on relationships being no exception. For instance, due to its central role for several research questions about relationships, various self-report measures exist for the assessment of global relationship satisfaction (see Fincham, Rogge, & Beach, 2018 for a review). Like for other self-reports, when assessing relationship satisfaction, researchers rely on individuals' introspection to accurately measure the construct of interest. These self-reports might be subject to several biases, for example social desirable responding or response styles like acquiescence (see Paunonen & O'Neill, 2010; Rammstedt & Farmer, 2013 for reviews).

Compared to self-reports about specific instances (e.g., "How satisfied are you feeling right now?"), global evaluations often demand to make a generalized judgment about multiple instances (e.g., "How satisfied do you generally feel?"). Such global judgments might additionally introduce so-called retrospective biases. For emotion reports, Robinson & Clore (2002) differentiate between memory-based and belief-consistent distortions. Memory-based biases describe influences on the answers caused by recalling certain experiences better or weighing them more than other experiences. Belief-consistent biases describe influences on the answers caused by situation-related or identity-related beliefs that are dissociated from actual experiences. When asking individuals about their retrospective evaluation of experiences during a certain time, memory-based biases are proposed to be especially pronounced, while belief-consistent biases are supposed to take effect for more generalized evaluations or for retrospection taking place a long time after the experience. Despite these potential biases, there is a widespread use of global self-reports, probably because they are easy to apply and prove to

be valid in many cases (Haeffel & Howard, 2010). However, to better understand the measures we apply, it is crucial to investigate and acknowledge the biases that might be in place when individuals answer certain questions.

3.3 Solutions Provided by Dyadic ESM Studies and Their Application in the Current Dissertation

The methodological problems reviewed so far demand solutions that can be to some extent tackled by relationship studies that apply the ESM. Studying dyads instead of single individuals additionally provides the benefit to obtain two perspectives on the phenomenon of interest, and thus to examine processes beyond the perception provided by only one partner. Hence, combining the ESM with the study of dyads extends its strengths by the benefits of dyadic studies. In the following, it will be outlined how the general advantages provided by dyadic ESM study designs were specifically implemented in this dissertation.

3.3.1 Solution A: Large-Scale Study of Short-Term Intra- and Interpersonal Processes in Everyday Life

First, the reviewed drawbacks of observational methods (low feasibility and ecological validity), cross-sectional studies (potential fallacy due to Simpson's Paradox) and long-term longitudinal studies (neglect of short-term processes) can be overcome with the ESM.

3.3.1.1 Using smartphones to study processes in a natural context

Technological advances and the growing number of individuals who possess smartphones (see Miller, 2012) made it increasingly feasible to apply the ESM in several psychological research areas. It provides an alternative to observational methods for investigating processes in individuals and couples as a source of relationship satisfaction in an economic and ecologically more valid manner, and is therefore often used for this purpose (Laurenceau & Bolger, 2005). For instance, studies applied the ESM to examine the so-called interpersonal process model of intimacy

(Laurenceau, Barrett, & Pietromonaco, 1998; Laurenceau et al., 2005b; Reis & Shaver, 1988), which proposes that self-disclosing behavior and responsiveness by the partner foster the experience of intimacy (an important predictor of positive affect in relationships, Laurenceau et al., 2005a). In Paper 1, a new process model termed “Dynamics of Motive Satisfaction” (DynaMoS) is introduced and empirically tested in individuals’ everyday lives. In contrast to the intimacy process model, the DynaMoS model does not only illustrate how certain aspects of an individual’s own and their partner’s behavior result in positive or negative interactions and experiences, thereby shaping in part their satisfaction with the relationship; it also specifies when these experiences are especially beneficial, namely depending on an individual’s motivational states.

For this purpose, the development of the smartphone application “Tellmi” was co-supervised by the author in the course of this dissertation. Tellmi is a flexible, SSL encrypted ESM app for Android and iOS devices, that can be customized for different study purposes. It supports all common questionnaire types, conditional branching of questions, a flexible implementation of survey schedules and push notifications. As a crucial feature it does not require an internet connection for submitting answers, enabling to theoretically capture the full breadth of everyday situations (e.g., even when participants are in (German) trains, or in the subway on the way to work). It is intended to be an open source software, and has already been used beyond the scope of the current work for other studies as well (e.g., Rosenkranz, Takano, Watkins, & Ehring, 2019).

By trying to ensure a random sample of momentary reports, researchers hope to obtain a representative picture of “life as it is lived” (Bolger et al., 2003, p. 579). Yet, some situations do not allow to answer questions on the smartphone, so it cannot be assumed that missed surveys are missing at random (Graham & Donaldson, 1993; McLean, Nakamura, & Csikszentmihalyi, 2017), which restricts the assumed representativeness of the ESM answers (and thus generalizability to everyday life; but see Silvia, Kwapil, Eddington, & Brown, 2013; Sun, Rhemtulla, & Vazire, 2019 showing at least only negligible systematic associations of ESM missings with momentary emotions and the presence of others). In the current studies, couples were further allowed to exclude certain time spans from the survey schedule to avoid scheduling surveys during periods in which it was impossible for one or both partners to provide answers. This option was given to ensure a high response rate, but it naturally affects the representativeness as well. Despite these constraints, ESM answers are expected to reflect individuals’ experiences in their

natural environments better than observational assessments in the lab.

3.3.1.2 Using multilevel models to separate between- and within-person effects

When asking individuals repeatedly about their momentary behaviors, experiences or feelings, their answers can be examined on different levels. The provided information at each assessment can be aggregated within each person, and can then be treated as person-level variable. Usually, an average is calculated across all assessments of a person, and the resulting variable is then interpreted as the typical state – for example, the typical behavior, relationship experience or relationship satisfaction of that person for the examined period. Bolger et al. (2003) suggest the variability of the state distribution as another promising way to aggregate states to answer certain research questions (see e.g., Cooper, Totenhagen, McDaniel, & Curran, 2018; Geukes, Nestler, Hutteman, Küfner, & Back, 2017). Within-person measures that are aggregated in such ways can be treated and interpreted as regular between-person measures: As predictor variables, they indicate how certain person-level outcome variables differ compared to other persons who have a different aggregated state in the variable of interest. In other words, such variables can only explain variance on the between-person level. For example, Paper 1 and Paper 2 investigate the relationship between the average motivation of a person and the average of that person’s corresponding behavior. In Paper 3, different summary statistics of the state distribution of relationship satisfaction are examined in their ability to predict global and retrospective relationship satisfaction measures.

Alternatively, responses can be examined on the state-level on which they were originally assessed and then be interpreted as within-person measures: As predictor variables, they indicate how certain state-level outcomes differ compared to other moments in which the individual had a different state in the variable of interest. Such variables can only explain variance on the within-person level. Further, these state variables can be brought into perspective, by relativizing them by the person’s typical state (“person-mean centering”): Treated this way, the variables allow to interpret how variations from an individual’s typical state predict certain state-level outcomes. For example, Paper 1 and Paper 2 investigate whether individuals, when they are more motivated compared to their typical motivation, subsequently show more corresponding behavior.

Multilevel models allow to examine both person- and state-level variables simultaneously, thereby separating the effects that are due to between-person differences and those that are due to within-person variation (at least when person-mean centering is applied, see Zhang, Zyphur, & Preacher, 2009; Wang & Maxwell, 2015). This circumvents the problem of drawing false conclusions from between-person effects about within-person effects, for instance due to the potential presence of a Simpson's Paradox. Moreover, within-person effects in multilevel models can be computed as fixed and random effects: Whereas fixed effects reflect the average within-person effects across individuals taking into account a potential different number of assessments per person, the random effects illustrate the variation of fixed effects across individuals (Bolger & Laurenceau, 2013, p. 32f). The papers in this dissertation focused on fixed effects, but random effects are reported as well, as these provide valuable information about the heterogeneity of the investigated effects (see also Bolger, Zee, Rossignac-Milon, & Hassin, 2019).

3.3.1.3 Consideration of short-term fluctuations in within-person variables

Within-person measures obtained with the ESM are meaningful for state-level variables that differ between moments and situations. Different situations are characterized by different contextual influences. In the case of couple relationships, the behaviors from and experiences with the partner in certain situations are such a contextual influence. Thereby, when individuals describe the interactions that happen on a momentary basis in a relationship, these variables assessed within-person become within-couple measures that can reveal the effects of short-term interpersonal processes. This approach was used in Paper 1, which examined relationship experiences as combination of own and partner behavior, and the effects of this interpersonal interaction on an individual's feelings.

Further, by looking at associations between states at different time points (e.g., using lagged variables), a short-term longitudinal effect can be examined so that experiences at a certain time point predict measures assessed at subsequent time points (Bolger & Laurenceau, 2013). In this way, studies that apply the ESM can uncover if the data are in line with short-term processes that propose to elicit different responses on the intra-individually varying constructs of interest. This logic was applied in two papers of this dissertation, by a) investigating how motivation shapes subsequent behavior (Paper 1 and Paper 2) and b) looking at the

interaction of relationship experiences and motivation at a prior time for predicting state relationship satisfaction and affect at a subsequent time point (“affect amplification of motivation”, Paper 1). Several time-spans can be investigated with ESM data, such as processes within days, between days, or between weeks. For the investigation of within-person effects in this dissertation, the smallest possible time span that was possible due to the study design was predominantly used, namely the time span between two answered surveys (translating into an average of three hours in both studies).

3.3.2 Solution B: Comparing Dispositions and Aggregated States

By providing a large amount of state assessments, the ESM provides the possibility to test theoretical claims and generate new hypotheses about the relationship between dispositions and everyday states. Two constructs are particularly relevant for the current work: motive dispositions and their relation to motivational states, as well as global relationship satisfaction and its relation to momentary state relationship satisfaction. In both cases, data provided by the ESM can be used for confirmatory and/or exploratory empirical investigations on how these different measures are related (see Paper 1 and Paper 3) and whether they have different predictive validities (see Paper 2 for motives/motivation).

3.3.2.1 Motives and motivation

Winter, John, Stewart, Klohnen, & Duncan (1998) define motive dispositions as “people’s wishes and desires – states of affairs that they would like to bring about (consciously or unconsciously) or, in the case of avoidance motives, states of affairs they would like to prevent” (p.231). Motives can be differentiated by their level of consciousness (implicit and explicit motives, see McClelland, Koestner, & Weinberger, 1989), or by the classes of goals they pertain to (e.g., communion or agency, Hagemeyer & Neyer, 2010). Further, many theories differentiate dispositional motives and motivational states, sometimes postulating a functional relationship between both constructs (e.g., Atkinson, 1957; Bischof, 1975, 1995; Deci & Ryan, 2000). Whereas the strength of a motive disposition is supposed to vary between persons, motivation is described as a state experience of desire that can vary for a single person from moment to moment (Rheinberg, 2002;

Schultheiss, Strasser, Rösch, Kordik, & Graham, 2012). The functional relationship between motives and motivation is thoroughly discussed in Paper 1 and Paper 2, and the important theoretical assumption that they are related is empirically tested in Paper 1. Further, the predictive validity of motive dispositions compared to motivational states and aggregated motivation is investigated in Paper 2.

3.3.2.2 Relationship satisfaction

Compared to the elaborated theoretical views on motives, there is no clear consensus in the field about the conceptualization of inter-individually varying global relationship satisfaction (Fincham et al., 2018). Some measures add together behavioral and emotional components, while conceptualizing it as a unidimensional construct (e.g., in the Couple Satisfaction Index the typical amount of disagreements with the partner is grouped with feelings about the partner or the relationship, Funk & Rogge, 2007). Other measures have a narrower focus on the assessment of positive and negative qualities of the relationship, conceptualizing the construct as two-dimensional (e.g., the Positive and Negative Relationship Quality Scale, Rogge, Fincham, Crasta, & Maniaci, 2017). Theory is largely lacking on which varying states within persons are supposed to be related to global measures; or when the aggregation of momentary assessments of relationship satisfaction gets representative for a global measure. Paper 3 investigates some of these questions by additionally looking at retrospective assessments of relationship satisfaction and their association with states and global evaluations to bridge theoretical gaps with exploratory, empirical work.

3.3.3 Solution C: Reducing Retrospective Errors and Obtaining Partner-Reports

In contrast to a potential reliance on memory and semantic knowledge that is associated with global self-reports, a central assumption for questions about momentary experiences is that individuals can access their feelings directly in the moment, thereby relying on what is termed experiential knowledge (Robinson & Clore, 2002). Consequently, such self-reports are not supposed to be influenced by retrospective biases (although this does not preclude that other response biases might be present, such as acquiescence; see e.g., Baird, Lucas, & Donnellan, 2017). The ESM allows assessing multiple instances of momentary evaluations,

aggregating them in certain ways, and comparing them to retrospective evaluations. Such comparisons can uncover what kind of retrospective biases occur for the construct of interest. In Paper 3, the assumption about the accessibility of feelings is transferred to relationship satisfaction states and the bias in retrospection of relationship satisfaction is investigated.

Compared to reports about momentary feelings, behavior reports assessed with the ESM in our studies might still be subject to retrospective biases: Specifically, participants were asked how they behaved since the last survey, which is a short-term retrospective assessment. Yet, in our dyadic ESM studies, partner-reports of behavior are also available, which are potentially less influenced by social desirability effects than self-reports (see e.g., Cui, Lorenz, Conger, Melby, & Bryant, 2005). Paper 2 looks at differences between partner- and self-reported behavior as outcome variables.

3.4 Overview of the Current Research: Testing Theory and Examining Measures Transparently

Table 1 presents an overview of the three papers included in this dissertation that draw on data from two conducted ESM studies (Study 1: $N_{ESM} = 130$ participants, in total 7,573 answered ESM measurement points, Study 2: $N_{ESM} = 510$ participants, in total 60,942 answered ESM measurement points). The analyses of Paper 1 and Paper 2 were conducted within a confirmatory framework, testing mainly preregistered hypotheses about the interplay of motivational constructs, relationship behaviors or experiences, and relationship satisfaction. In contrast to the first two papers, the analyses of Paper 3 were mainly exploratory.

Paper 1 presents the preregistered process model DynaMoS, which proposes a motivational dynamic occurring in intimate relationships' everyday life. The model could explain which processes might produce existing findings in the literature on between-person associations of motive dispositions and global relationship satisfaction. The DynaMoS model is tested on the data of Study 1 for the domain of communion motives.

Paper 2 builds upon this model, describing the results of an exact replication of a within-person path of the DynaMoS model with the data of Study 2, namely

the relationship between communal motivational states and behavior self-reports. This analysis is further extended to partner-reported behavior, to the agency motive domain, and compares the between-person associations of motivational variables and behavior reports.

Paper 3 focuses on the comparison of different assessment modalities of relationship satisfaction. Different summaries of relationship satisfaction states and different numbers of assessments are used to predict retrospective and/or global relationship satisfaction. Further, mean-level bias occurring in retrospection is examined, together with potential moderators of this bias.

The research in this dissertation is located at the intersection of basic motivational psychology, applied research on relationships and psychological methods. It advances the field of motivational psychology by translating classical theoretical accounts of the functionality of motives into corresponding within-person analyses. For the field of relationship science, the process perspective provides insights about the circumstances in which interpersonal behaviors are effective in translating into momentary experiences of relationship satisfaction. Finally, the presented work broadens the understanding of both classical measurements as well as newly implemented ESM measures, by illustrating their congruence and differences.

The whole dissertation adheres to all principles of transparent research: First, the data of both studies were made available as scientific use files (Zygar et al., 2018a; Zygar-Hoffmann, Hagemeyer, Pusch, & Schönbrodt, 2020), allowing other researchers to use the data for their own research questions, while preserving the anonymity of the participants. Second, all results presented in the papers are accompanied with analysis scripts written in the open source statistical language R (R Core Team, 2018). This ensures the reproducibility of the presented results, and allows verifying their robustness under different model specifications and analysis plans (e.g., adding covariates, using other exclusion criteria). Third, the materials of both studies (i.e., items, codebooks) are publicly available (osf.io/b8pu6/, osf.io/psqx8/), enabling direct replications of the studies, and providing the opportunity for a more consistent use of ESM items across study sites. Currently, the fast rise in studies applying the ESM is missing an according rise in validated state measures. The ESM item repository (www.esmitemrepository.com) is an attempt to countervail this problem by providing an overview of the application, the psychometric properties, and the evidence for the validity of ESM items. All items of both studies were submitted to this repository.

Table 1
Overview of Papers

Paper: Short Title	Research Questions	Focal Variables	Data	Status	Citation
<p>Paper 1: From motive dispositions to states to outcomes</p>	<p><i>For the domain of communion:</i> a) Can the association between motive dispositions and global relationship satisfaction be replicated? b) How do individuals with different motive strength vary in their experience of everyday motivation? c) What processes can predict their state and global relationship satisfaction?</p>	<ul style="list-style-type: none"> - Implicit motives - State and aggregated motivation - Behavior self-reports - Interpersonal relationship experiences - Momentary and aggregated relationship satisfaction - Global relationship satisfaction 	Study 1	published	<p>Zygar, C., Hagemeyer, B., Pusch, S., & Schönbrodt, F.D. (2018). From motive dispositions to states to outcomes: An intensive experience sampling study on communal motivational dynamics in couples. <i>European Journal of Personality</i> 32, 306-324. https://doi.org/10.1002/per.2145</p>
<p>Paper 2: Motivated behavior in intimate relationships</p>	<p><i>For the domains of communion and agency:</i> a) Can the path between motivation and behavior of the DynaMoS model be directly and conceptually replicated? b) How do motivational states and their aggregates compare to motive dispositions for the prediction of behavior reports?</p>	<ul style="list-style-type: none"> - Implicit motives - Explicit motives - State and aggregated motivation - Behavior self- and partner -reports 	Study 1	revision required (meanwhile published in 2020)	<p>Zygar-Hoffmann, C., Pusch, S., Hagemeyer, B., & Schönbrodt, F.D. (major revision). Motivated behavior in intimate relationships: Comparing the predictive value of motivational variables. <i>Social Psychological Bulletin</i>.</p>
<p>Paper 3: Recalling experiences</p>	<p><i>For relationship satisfaction:</i> a) What summary statistic of states corresponds best to retrospection and global assessments? b) What bias occurs in retrospection? c) What moderates mean-level bias? d) What level of aggregation is sufficient to approach a reliable measurement of the global index?</p>	<ul style="list-style-type: none"> - Momentary and aggregated relationship satisfaction - Retrospective relationship satisfaction - Global relationship satisfaction 	Study 1 + 2	published	<p>Zygar-Hoffmann, C., & Schönbrodt, F.D. (2020). Recalling experiences: Looking at momentary, retrospective and global assessments of relationship satisfaction. <i>Collabra: Psychology</i>, 6(1), Article 7. https://doi.org/10.1525/collabra.278</p>

Note. Study 1: $N_{ESM} = 130$ participants, in total 7,573 answered ESM measurement points, Study 2: $N_{ESM} = 510$ participants, in total 60,942 answered ESM measurement points. In both studies participants answered questions five times a day, but the ESM study period lasted two weeks in Study 1 and four weeks in Study 2.

Chapter 4

Paper 1: From Motive Dispositions to States to Outcomes: An Intensive Experience Sampling Study on Communal Motivational Dynamics in Couples



From Motive Dispositions to States to Outcomes: An Intensive Experience Sampling Study on Communal Motivational Dynamics in Couples

Caroline Zygar

Ludwig-Maximilians-Universität München

Birk Hagemeyer

Friedrich-Schiller-Universität Jena

Sebastian Pusch

Friedrich-Schiller-Universität Jena

Felix D. Schönbrodt

Ludwig-Maximilians-Universität München

Corrected version of the published article in the *European Journal of Personality*.

Zygar, C., Hagemeyer, B., Pusch, S., & Schönbrodt, F. D. (2018). From motive dispositions to states to outcomes: An intensive experience sampling study on communal motivational dynamics in couples. *European Journal of Personality*, *32*, 306–324. <https://doi.org/10.1002/per.2145>

Zygar, C., Hagemeyer, B., Pusch, S., Schönbrodt, F. D. (2019). Corrigendum. *European Journal of Personality*, *33*(2), 214. <https://doi.org/10.1002/per.2199>

Abstract

Embedded in a theoretically founded process model (termed Dynamics of Motive Satisfaction, “DynaMoS”), the present study examined the links between the implicit dispositional communion motive, everyday motivational dynamics, and relationship outcomes in couples. Within-subject processes are proposed to explain between-subject associations of dispositional motives and relationship satisfaction. For an empirical test of the model, data on the dispositional partner-related need for communion and global relationship satisfaction was obtained from 152 individuals in heterosexual relationships. In an extensive experience sampling spanning two weeks, a subsample of 130 individuals answered questions about their current motivational states, mood, state relationship satisfaction and experiences with their partner five times a day. The results were largely consistent with the DynaMoS model: 1) Individuals with a strong dispositional implicit communion motive reported more often to be in a communal motivational state. 2) Communally motivated individuals were more likely to engage in subsequent instrumental behavior. 3) Relationship experiences that potentially satisfy communion motivation led to more positive relationship outcomes when individuals were motivated before compared with when they were not. It is discussed how these results and the experience sampling method can foster our understanding of how dispositional characteristics translate into everyday processes and shape relationship outcomes.

Keywords: implicit communion motives, traits, states, relationship satisfaction, experience sampling

When looking into the psychological literature about relationship functioning, a variety of research can be found on between-person and between-

Caroline Zygar, Department of Psychology, Ludwig-Maximilians-Universität München, Germany. Birk Hagemeyer, Institute of Psychology, Friedrich-Schiller-Universität Jena, Germany. Sebastian Pusch, Institute of Psychology, Friedrich-Schiller-Universität Jena, Germany. Felix D. Schönbrodt, Department of Psychology, Ludwig-Maximilians-Universität München, Germany.

We embrace the values of openness and transparency in science (<http://www.researchtransparency.org/>). We provide all data necessary to reproduce the reported results as a scientific use file (Zygar, Hagemeyer, Pusch, & Schönbrodt, 2018) and provide reproducible scripts for all data analyses reported in this paper at <https://osf.io/b8pu6/>. Our preregistration can be found at <https://osf.io/hafsx/>.

This research was funded by grants from the German Research Foundation to Felix Schönbrodt (SCHO 1334/5-1) and Birk Hagemeyer (HA 6884/2-1).

The authors thank Paula Fehrmann, Nicole Horn, and Lara Lietge for their support in PACT scoring, Helen Baumann for assistance during data collection, and Tobias Kächele for the app development.

Correspondence concerning this article should be addressed to Caroline Zygar, Leopoldstr. 13, 80802 München, Germany. Email: caroline.zygar@psy.lmu.de. Phone: +49 89 2180 5079.

couple factors that are associated with better or worse relationship quality. These include, for example, attachment styles (Nofhle & Shaver, 2006; Shaver & Brennan, 1992), personality traits (Heller, Watson, & Ilies, 2004; Malouff, Thorsteinsson, Schutte, Bhullar, & Rooke, 2010; Wilson, Harris, & Vazire, 2015), and relationship commitment (Le & Agnew, 2003), to name just a few. Research on such between-subject factors may serve as a starting point when addressing the within-subject processes that drive these associations (see, e.g. Gable & Poore, 2008; Howell, Ksendzova, Nestingen, Yerahian, & Iyer, 2017; Kanat-Maymon, Argaman, & Roth, 2017; Nezlek, Newman, & Thrash, 2017; Sadikaj, Moskowitz, & Zuroff, 2016; Sadikaj, Moskowitz, & Zuroff, 2015; Vater & Schröder-Abé, 2015, for such approaches). That being said, it is important to note that results of inter-individual analyses cannot be transferred to the intra-individual level of analysis, at least not under reasonable assumptions of non-ergodic psychological processes (Molenaar, 2004; Molenaar & Campbell, 2009). Doing so is known as an *ecological fallacy* (Curran & Bauer, 2011): Variation between individuals has to be distinguished from variation within individuals. Causal processes, however, typically operate on the within-person level (Hamaker, 2012). Furthermore, insights about mediating causal processes are necessary for interventions that aim to improve relationship quality (Back & Vazire, 2015).

The starting point for our research question was the finding that the dispositional need for closeness and communion is positively related to global relationship functioning in couples (Hagemeyer, Neberich, Asendorpf, & Neyer, 2013; Hagemeyer & Neyer, 2012). This between-subject result, however, is at best suggestive for the underlying causal processes. For example, one cannot conclude from this result that individuals' momentary relationship satisfaction will change when experiencing a momentary motivation for closeness (which is a within-subject effect). By changing the level of analysis to the within-person level, the current paper takes a closer look at which intra-individual process chain can explain the between-subject finding (Back et al., 2011): How do individuals with different motive strength vary in their everyday experiences and what processes can predict their state and global relationship satisfaction? We propose a theoretically founded process model illustrating the within-person **Dynamics of Motive Satisfaction** ("DynaMoS") occurring in close relationships (Figure 1).

The DynaMoS model includes components from all phases of social interactions (Back & Vazire, 2015; Heckhausen & Heckhausen, 2008): motivation (pre-action phase), instrumental behavior (action phase), and relationship perception (post-action phase). It is not a novel theoretical model, but derived

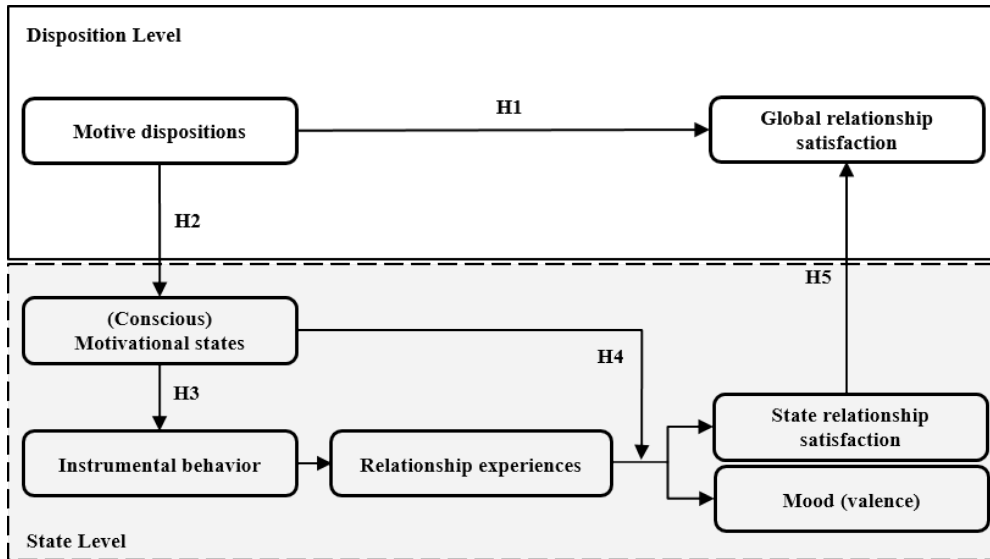


Figure 1. Overview of the DynaMoS model and associated hypotheses. Figure available at <https://osf.io/b8pu6/>, under a CC-BY4.0 license.

from assumptions of existing theories and is meant to be generic for motivational processes in all kinds of relationships that furthermore should *mutatis mutandis* apply to all motivational domains, such as power or achievement. In the following descriptions we will apply it to the target domain of the current study, which is communal motivation. We used an intensive longitudinal method (Experience Sampling Method ESM, Csikszentmihalyi & Larson, 1987; Wrzus & Mehl, 2015) to test our predictions in the everyday lives of couples. In the following sections we will elaborate on our understanding of motives, motivational states, and the assumed processes.

Motive Dispositions

Individuals differ in their preferences for certain classes of goal states. McClelland (1987) conceptualizes motives as such dispositional preferences, energizing behavior in a desired direction, orienting attention to relevant incentives, and selecting behavior by facilitating learning in that domain. Desirable classes of goals in interpersonal contexts concern for example power, independence, affiliation, and intimacy. Based on a classification first introduced by Bakan (1966), these motives can be combined to the broad categories of agency (for power and independence) and communion (for affiliation and intimacy; Brunstein, Schultheiss, & Grassmann, 1998; Hagemeyer & Neyer, 2012). These categories are also being used in the literature on interpersonal psychology, describing the two orthogonal behavioral dimensions of the interpersonal circumplex (Horowitz et al., 2006).

According to dual motives theory (McClelland, Koestner, & Weinberger, 1989; Schultheiss, 2001) two distinct motive systems can be distinguished that have functionally different underpinnings: The explicit motive system refers to self-attributed goals and values which can be deliberately retrieved by introspection and are supposed to guide decisions in highly structured situations. The system of implicit motives on the other hand corresponds to more spontaneous, inherently rewarding behavior, for instance becoming apparent in unexpected situations that require prompt action. The current paper focuses on implicit motives, which are typically assessed by indirect measures, as they are not expected to be consciously accessible.

More specifically, we are interested in the implicit communion motive, that comprises the motives of affiliation and intimacy. While those two motives can theoretically be distinguished in more detail, they both entail the need for closeness in positive, warm relationships (see Hofer & Hagemeyer, in press; Schönbrodt & Gerstenberg, 2012; Weinberger, Cotler, & Fishman, 2010, for a discussion of the difference between affiliation and intimacy) and show a high correlation (Hagemeyer, Dufner, & Denissen, 2016). In the context of couple relationships, the partner-related need for communion (*pnCommunion*) is defined as “a recurrent concern for closeness to one’s partner and for experiences of the self as part of a dyad” (Hagemeyer & Neyer, 2012, p. 115). While physical closeness facilitates the implementation of *pnCommunion*, emotional closeness is at the heart of it: deriving pleasure from a sense of unity, by both members of a couple sharing thoughts and emotions, involving the partner in one’s experiences, and showing compassion and affection.

From Motive Dispositions to Global Relationship Satisfaction

We will now turn to highlighting different parts of the DynaMoS model presented in Figure 1. While implicit motives have been studied since the late 1940s in various settings (see Schultheiss & Brunstein, 2010, for an overview), research in the context of couple relationships is still rather sparse. Previous studies mainly highlighted the result of implicit motives being related to relationship outcomes (*H1* in Figure 1). The implicit power and more broadly agency motives were found to be negatively related to relationship outcomes (Hagemeyer & Neyer, 2012; Hagemeyer, Schönbrodt, Neyer, Neberich, & Asendorpf, 2015; Mason & Blankenship, 1987; Stewart & Rubin, 1976; Zurbriggen, 2000).

Of specific interest for the current study though, is research examining the influence of implicit communion motives: McAdams and Vaillant (1982) showed initial evidence for a positive association of the implicit intimacy mo-

tive with marital quality in a longitudinal study of 17 years investigating a male sample. The relationship-specific communion was also shown to be positively related to one's own – as well as the partner's – relationship satisfaction (Hagemeyer & Neyer, 2012). Two other findings indirectly suggest a worse relationship quality for those with high communion motives: Mason and Blankenship (1987) reported that female undergraduates characterized by a strong affiliation motive and activity inhibition exerted more physical and psychological abuse in their relationship when experiencing a stressful year. Similarly, Zurbriggen (2000) found that women with a high implicit affiliation-intimacy motive showed increased levels of aggression in their relationship. Weinberger et al. (2010) discuss these findings in the light of a “dark side” of the affiliation motive (pp.73, 81), driving individuals to maladaptive behavior when fearing their need for closeness to be frustrated by rejection or dissolution of a relationship (see also Hofer & Hagemeyer, in press). As relationships should on average satisfy the need for closeness more often than frustrate it, these results might not generalize to a worse relationship satisfaction per se. They might instead be indicative of more aroused reactions during episodes of frustration.

Taken together, some research points to communion motives being positively correlated with one's own overall relationship satisfaction, although person and situation variables might moderate and even reverse this association (e.g. during frustrating episodes in a relationship).

From Motive Dispositions to Motivational States

The process we assume behind the findings on associations between motives and global relationship satisfaction starts with the translation of motive dispositions into motivational states (*H2* in Figure 1). Additional to assessing motive dispositions, the current study thus works with intensive longitudinal data on individuals' self-reported motivational states throughout the day. It is therefore central to establish how motivation investigated in this way is assumed to relate to and differ from motive dispositions.

Generally, dispositions and traits depict stable differences in certain kinds of behaviors and experiences between persons.¹ States, however, also differ within a person and can vary on small time scales. Dispositions and states

¹For the distinction between the term of (motive) dispositions and (other) traits (like the Big Five) we base our reasoning on Winter, John, Stewart, Klohnen, and Duncan (1998). They argue that both motives and traits are important influences on the behavior of a person, but motives describe what a person wants, and *why* the person is acting in a certain way, while traits describe *how* a person acts. To avoid confusion between these two concepts, we use the term “disposition” for motives.

can be closely related, for example, when behavior and experiences captured in state measures reflect the tendencies described in the definition of according disposition measures.

Traits as Density Distribution of States

Looking into research on the Big Five personality traits, Fleeson (2001) formulated and tested the assumption that inter-individual differences in personality are reflected in different density distributions of respective personality states. This was corroborated by studies reporting that average personality states show robust correlations with global personality trait measures (Augustine & Larsen, 2012; Fleeson, 2001; Fleeson & Gallagher, 2009). However, it was also pointed out that aggregated states are not to be considered equivalent to global trait reports, having different predictive validities (Augustine & Larsen, 2012; Finnigan & Vazire, 2017).

In a similar vein we consider the density distributions of motivational states to reflect the correspondent motive dispositions: Individuals with a strong motive have a strong preference for certain classes of goals. Being motivated represents the momentary need to adjust the behavior in the direction of such a preferred goal (e.g. increasing closeness to the partner for the communion motive). The frequency and strength of experiencing such motivational states should therefore be influenced by the individual preference described in motive dispositions. Additionally though, motivational states are determined by situational influences, as will be described in the following.

A Systems Theory of Motivation

The Zurich Model of Social Motivation (Bischof, 1975, 1995) provides a functional perspective on the relationship between motive dispositions and motivational states. In terms of systems theory, motivation reflects a discrepancy between the current level and the reference level for a specific need. The reference level is supposed to be influenced by the motive disposition, being higher for individuals with a stronger motive. The current level, however, is shaped by the situation: When situations frustrate a specific motive, the current level can drop below the reference level, and motivational *appetence* should emerge. Appetence describes the momentary activation to reduce the discrepancy between current level and reference level (i.e. to have need satisfying experiences). For communion, this could translate to the motivation to seek out closeness to the partner. In contrast, when the current level exceeds the reference level the term motivational *aversion* is used, indicating a momentary activation to move away from a certain end state. This could for

instance be the motivation to reduce the amount of closeness to the partner (see also Hofer & Hagemeyer, in press; Schneider, 2001).²

When motives would be permanently being fulfilled by the situation on a perfect level, even individuals with a strong dispositional motive should rarely experience motivational states. Nevertheless, a higher amount of satisfactory events is needed for these individuals to reach their reference level, compared to individuals with a weak motive. Therefore, averaged across typical situations, individuals with a strong motive disposition should experience the respective motivational appetence more often and more strongly (see Hagemeyer, Neyer, Neberich, & Asendorpf, 2013 for an application of this principle to scale development).

Conscious Motivational States

According to Bischof (2008), implicit motives are considered to influence the (visible) behavior and (conscious) affective experiences of individuals, but the associated motivational state that guides the attention to relevant incentives is not necessarily accessible to introspection. Nonetheless, individuals seem to be able to report on their motivational states, and these reports seem to be related to the implicit motive: McAdams and Constantian (1983) for instance showed that the implicit intimacy motive predicted a weaker self-reported (state) desire to be alone when participants interacted with someone. In contrast, when participants were alone, the motive predicted a stronger desire to interact with someone. This raises the question under which conditions motivational states are accessible to introspection. Bischof (2008) argues that these conditions include situations that constitute barriers impeding the satisfaction of the respective implicit motive (e.g. unavailability of the partner, or frustrating circumstances). As such barriers prevent an automatic regulation of the motivational activation and behavior, the explicit system is activated to provide additional cognitive resources to overcome the barriers, and motivational states tend to become conscious (see also Strack & Deutsch, 2004).

From Motivational States to Instrumental Behavior

Following from the functional properties of motives, individuals who report feeling motivated should show corresponding instrumental behavior to attain the desired end state of the respective motive (*H3* in Figure 1). This

²In H4 of the current study, we preregistered to focus on appetence in contrast to aversion. From a theoretical view predictions should be expendable, but aversion did not occur as often as appetence in prior pilot data (and the current data as well) and might not be as well understood by participants. The interpretation of the results do not change, when including aversion in the analyses.

process should serve to align the current level for a specific need with the reference level, for example adjust an individual's behavior to achieve a satisfactory level of emotional closeness. As already outlined earlier, this is not always possible, as situational circumstances might prohibit desired behavior (e.g. the partner, norms, obligations). Moreover, individuals might have competing motives or other interests, interfering with the implementation of a specific behavior (see Hofmann, Baumeister, Förster, & Vohs, 2012, Riediger & Freund, 2004 for studies on competing desires and goals). Despite these constraints, we would on average expect a positive relationship between motivational states and behavior that aligns with these motivational states.

The Satisfaction of Motivational States in Relationships

An actor's instrumental behavior can facilitate the satisfaction of his or her motives, but due to the interpersonal nature of the communal motive, the partner's behavior is equally important.

Relationship Experiences as a Source of Satisfaction

To foster closeness in a relationship it is typically not sufficient for the actor to seek and show affection – the partner has to reciprocate and be willing to give the desired affection. In this sense, the partner's behavior is a situational factor for the satisfaction of an actor's motives. Accordingly, the intimacy process model of Reis and Shaver (1988) proposes the perceived responsiveness of the partner as a central influence for the experience of intimacy. In this context, event-contingent and daily diary studies showed that not only self-disclosure, but also partner-disclosure predicted the experience of intimacy in interpersonal exchanges, with partner responsiveness partially mediating this association (Laurenceau, Feldman Barrett, & Pietromonaco, 1998, Laurenceau, Feldman Barrett, & Rovine, 2005; see also Debrot, Schoebi, Perrez, & Horn, 2013 for a study on actor and partner effects of everyday physical signs of affection).

However, the partner's motivated behavior does not always align with an actor's motivation and instrumental behavior, potentially leading to difficulties in the regulation of dyadic closeness. Still, given that most couples report at least slightly positive relationship satisfaction (Funk & Rogge, 2007; Heyman, Sayers, & Bellack, 1994), and communal behavior tends to elicit communal behavior (Markey, Funder, & Ozer, 2003), we expect that communal experiences generally are an inherent feature of a couple relationship (i.e. couples typically show affection and provide feelings of unity, especially if one partner initiates such behavior), thereby satisfying communal motivation.

In this regard, our process model shares important aspects with a revision of the circumplex model of interpersonal behavior (Horowitz et al., 2006). Likewise, this model (a) emphasizes that motives are the underlying force of interpersonal behavior and (b) explicates that motivated behavior requires and invites complementary, that is, matching reactions, to satisfy the motive behind it. As such, the model can be used to derive motive-specific partner reactions that are potentially involved in the satisfaction of a motivational state. Whereas the circumplex model describes the interpersonal reaction certain behavior invites in other persons (i.e., one that satisfies the motive behind the behavior; see also Markey et al., 2003), our model focuses on the intra-individual processes.

Motives and Motivational States as Affect-Amplifiers

Behavior that leads to rewarding experiences should elicit positive affect: Studies showed on the intra-individual level of analysis, that satisfaction of different desires is generally positively related to daily well-being (Neubauer & Voss, 2016; Sheldon, Ryan, & Reis, 1996), and positive emotional experiences (Le & Agnew, 2001). The increase of perceived intimacy was found to predict positive affect in an ESM study with couples (Laurenceau, Troy, & Carver, 2005). These studies, however, did not consider the strength of the motive or momentary motivational state.

Motive theory (McClelland, 1987) posits that the attainment of a goal – such as emotional closeness – is not equally satisfactory for everybody, and the non-attainment of a goal not equally frustrating. Rather, the motive disposition should modulate the strength of the affective reaction: For persons with a weak intimacy motive, for example, a close interaction should be less satisfying than for a person with a strong intimacy motive. Likewise, being separated from significant others should be more frustrating for a person with a strong intimacy motive (“motives as affect-amplifiers”, see Schultheiss, 2008 for a review). A study from Dufner, Arslan, Hagemeyer, Schönbrodt, and Denissen (2015) indeed showed such affective contingencies, namely individuals with a strong affiliation motive having a stronger tendency to experience and display physical indicators of joy in response to affiliative incentives. Further empirical studies showed that motives moderate the effect of motive-relevant experiences and goal progress on various measures of affect, well-being and satisfaction (Brunstein et al., 1998; Job, Bernecker, & Dweck, 2012; McAdams & Constantian, 1983; McAdams, Jackson, & Kirshnit, 1984; Schultheiss, Jones, Davis, & Kley, 2008). Specifically for the communion domain, Hofer and Busch (2011a) could show in an inter-cultural study that individuals with a strong

implicit affiliation motive had a higher relationship satisfaction when reporting high levels of relatedness experiences compared to those with a weak implicit motive (see Hofer & Busch, 2011b for similar results regarding feelings of envy and aggression after frustration). While these existing studies on the affect-amplifying effect of dispositional motives are on the inter-individual level, the current study focuses on intra-individually varying motivational states, behavior, satisfaction and affect instead.

In this respect, we extend the hypothesis of “motives as affect-amplifiers” to the motivational *state* level: Just as food tastes better when you are hungry, we expect that individuals who are momentarily highly motivated to experience closeness to their partner should be more happy when this desire is satisfied, compared to moments when they did not care as much. Generally speaking, the “motivation as affect-amplifiers” hypothesis states that affective reactions are stronger – both in the positive and the negative direction – when the current motivation to fulfill a motive is strong, compared to when motivation is weak. As couple relationships typically provide such rewarding experiences, individuals with a strong p_nCommunion should experience the affect-amplifying nature of motivation more often in a positive, satisfying way, leading to a higher relationship satisfaction on average compared to those with a less pronounced motive. We assume this motivational dynamic to be one of the processes not only influencing mood, but also momentary relationship satisfaction (*H4* in Figure 1). The within-subjects process of affect-amplification would then contribute to the explanation of the between-subjects association between the implicit communion motive and global relationship satisfaction.

From State to Global Relationship Satisfaction

Finally, we assume an association between the mean of relationship satisfaction states and global assessments (*H5* in Figure 1). Analogous to the domain of motives (mentioned earlier), and personality (see Fleeson, 2001), the global evaluation should be a reflection of the density distribution of states: Individuals who frequently experience momentary feelings of satisfaction with their relationship should also assess their relationship globally positive. This was conceptually already shown, for instance, by Hofmann, Finkel, and Fitzsimons (2015) although with different operationalizations of relationship satisfaction on the state and global level. The association of stable constructs and mean states on the one hand for motives and on the other hand for relationship satisfaction, relates the start and end of the within-subject process to the between-subject result.

Hypotheses

The aim of the present study is to test the proposed DynaMoS model in Figure 1 exemplarily for the domain of communion motives. We suggest that the illustrated processes are key to understanding why implicit motives are related to relationship outcomes.

In a first step, we aimed to replicate the previously reported inter-individual association between implicit communion motives and individuals' global relationship satisfaction. A priori we did not see a theoretical justification for gender differences; therefore we specified our hypothesis irrespective of gender, namely:

H1: Individuals' implicit partner-related need for communion (disposition) is positively related to their global relationship satisfaction.

As the corresponding partner effect has not been as extensively studied in previous research, we did not make a prediction about it. Still, we included the partner effect of pnCommunion on global relationship satisfaction in our analyses. In a second step, we formulated hypotheses pertaining to the different parts of the proposed process in the DynaMoS model:

H2: Individuals' implicit partner-related need for communion (disposition) is positively related to their mean state of communion motivation.

H3: Individuals' communion motivation (state) is positively related to their subsequent instrumental communal behavior (state).

H4A: Individuals' communion motivation (appetence state) interacts with communal relationship experiences (state) to predict individuals' subsequent mood (valence, state).

H4B: Individuals' communion motivation (appetence state) interacts with communal relationship experiences (state) to predict individuals' subsequent relationship satisfaction (state).

Finally we assumed that the mean experience of state relationship satisfaction represents more stable, global assessments of relationship satisfaction. All hypotheses were preregistered.³

H5: Individuals' mean state of relationship satisfaction is positively related to their global relationship satisfaction.

³We preregistered additional hypotheses, that are not mentioned in this paper (such as the influence of barriers on the emergence of motivational states or the extension to implicit agency motives). These analyses will be part of future work.

Method

We report how we determined our sample size, all data exclusions, and all measures in the study (Simmons, Nelson, & Simonsohn, 2012).

Sample

Formal power analyses require a guess about several (co)variances and effect size components in a complex multi-level data structure. Given the largely unexplored nature of our research design, hypotheses and measurement instruments (particularly at the state level), we determined sample size by practical constraints: Data collection was scheduled between November and December 23rd. One couple started late, with two days of the ESM taking place during the Christmas holidays. In order to eliminate potential bias due to the special nature of the holidays, we excluded the last two study days of this couple from our analyses.

During participant registration we excluded three couples from the entire study for not having compatible smartphones. During the study we learned that one “couple” participated without actually being in a relationship, thus we excluded their data. In the end, we managed to collect data from 152 persons pertaining to 77 couples for the preliminary questionnaire. For two couples only one partner participated, but the data from these individuals were still included. For one couple, both partners gave identical answers in the measure of implicit motives, therefore these answers were treated as missings. Most of the participants were students (77%), mean age was 22.74 years ($SD = 4.54$, Range = 18–40 years), average relationship duration was 2.49 years ($SD = 2.01$, Range = 2 weeks to 8 years), and only five individuals had children.

After finishing the preliminary questionnaire, six couples opted out of the ESM part of the study. Another two couples and six individuals answered less than one third of all surveys, which was below the preregistered minimum for inclusion in the analyses. This resulted in a final sample of 130 individuals (from 68 couples) for the ESM part. A post-study feedback questionnaire was completed by 117 of these individuals.

Procedure

Couples living in a heterosexual relationship were recruited via social networks, newsletters, and at a German university to participate in a study on social desires. When registering, each couple chose a time span of 13.5 hours in which they were usually awake and able to answer five surveys on

a daily basis for two weeks.⁴ Each individual received a personal identifier that served to link their data across datasets and to match partners. Subsequently participants were instructed to individually answer an online preliminary questionnaire on their personal computers (programmed with *formr*, Arslan & Tata, 2016). Participants further received instructions to install an experience sampling application on their smartphones, which was developed at LMU Munich for Android devices. Upon logging into the app, the questions and survey modalities were introduced. Starting with the day after the login, five daily surveys were scheduled at semi-random time-points in the chosen time span for the following two weeks. The surveys were scheduled to be approximately evenly distributed throughout the day. Participants had 45 minutes to answer the questions before the survey became inactive. As the timing schedule was the same for both partners, the surveys were available at the same time, but participants were instructed to answer the questions separately from their partners and not to talk about their answers. Fifty-three entries were excluded from data analysis, because participants indicated that they had discussed their answers.

The questions were identical in each survey and median duration for answering was 3.28 minutes (interquartile range = 2.5). After one week, participants were encouraged via email to keep answering as many surveys as possible. Participants could receive a report on their own answers, were eligible for course credit, and had the opportunity to win a voucher when completing at least 80% of the 70 surveys. Actual compliance was on average 84% ($SD = 14\%$), leading to a total of 7742 completed measurement points. After finishing the two weeks of ESM, participants were invited to give feedback about the study and to answer a few additional questions.

Measures of the Preliminary Questionnaire

The complete codebook of our measures can be found at <https://osf.io/d5jp2>. It includes all variables of this study, also those not included in the current paper.

Implicit partner-related need for communion. We used the Partner-Related Agency and Communion Test (Hagemeyer & Neyer, 2012) consisting of eight ambiguous pictures to assess pnCommunion. Participants were instructed to write a story based on three questions about the relationship(s) of the person(s) depicted on each of the pictures. The stories were coded for the appearance of communal themes without knowledge of the rest of the

⁴Starting times ranged from 08:00 am to 10:30 am, end times from 9:30 pm to midnight. The preregistration contains erroneous timeframes on this matter.

data by 5 trained coders. Each case was coded independently by two coders, who were randomly assigned to cases. Ambiguous codings were resolved by discussion, and inter-coder consistency was high ($ICC(1,2) = .95$). The sums of communion codings across the eight pictures from the two coders were averaged. The covariance between these raw motive scores and word count ($r = .37$) was partialled out in a linear regression to control for confounding of motive scores with verbal fluency (see Hagemeyer & Neyer, 2012).

Global relationship satisfaction. We used two measures to assess global relationship satisfaction with sixteen items each: The Couple Satisfaction Index (CSI(16), Funk & Rogge, 2007) and the Positive–Negative Relationship Quality (PN-RQ) Scale (Rogge, Fincham, Crasta, & Maniaci, 2017). The CSI is meant to assess global evaluations of the relationship as a unidimensional construct. Participants were asked to rate statements such as “Please indicate the degree of happiness, all things considered, of your relationship.” on 6- and 7-point Likert scales and to evaluate their relationship on bipolar adjective scales (see codebook for details). Ratings were summed, with higher scores indicating higher satisfaction. The PN-RQ on the other hand assesses specifically the positive and negative qualities of the relationship as two distinct constructs. Participants rated their relationship regarding positive adjectives (e.g. enjoyable) and negative adjectives (e.g. miserable) on 7-point Likert scales ranging from 1 = *Not at all* to 7 = *Extremely*.

Measures During Experience Sampling

Communion motivation. Communion motivation was measured by two items, asking participants whether they wished for a certain relationship experience right now (“Share experiences, thoughts or feelings with your partner” and “Receive emotional affection from your partner”). The instruction changed for all items when participants indicated that they did not actively spend time with their partners at the moment of the survey. They were then asked to imagine they had 30 minutes of free time at that moment, which they could spend with their partner – and whether they wished for the mentioned behavior in that time. We adjusted the instruction because of the potential distortions on self-reported motivational states while being busy (e.g. working, studying). In such situations individuals might not report on their actual desires, but instead on the restricted opportunities in the situation. Answers were given on 7-point Likert scales, with four appetence answers (from *yes, very strongly* to *yes, but only weakly*), one middle category (*no, I don’t need this right now*), and two aversion answers (*no, that would rather bother me a little bit* and *no, that would bother me quite a lot*). We calculated a scale by

taking the mean of the two items (item level reliability was .66 with aversion answers, and .63 without⁵).

Instrumental communal behavior. Instrumental behavior for communion motivation was measured with a multiple choice item. Participants indicated whether they had displayed any of a number of different behaviors since the last survey (for example *interest, acceptance, affection, appreciation, understanding or emotional empathy* - all with regard to their partner; see codebook). The different behaviors were assigned weights regarding the degree to which they are apt to foster or hinder the fulfillment of communal needs. The weighting for all indices in the study was preregistered and derived from discussion among the four authors. An index was calculated by summing up the weighted answers. Positive scores reflect suitable instrumental behavior, and negative scores reflect adverse behavior.

Communal relationship experiences. Participants did not only provide information about the way they behaved as actors, but they also answered multiple choice items on (a) reciprocal behavior (e.g. *Stronger fight or conflict*, see codebook) and (b) behavior their partner showed since the last survey. The list of partner behaviors was identical to the list of their own behaviors (wording accordingly adjusted), but had different weights regarding their potential for communal satisfaction or frustration. For example the option *praise, admiration or recognition* had a small positive weight for communal satisfaction when it was received from the partner, but a zero weight when it was marked as one's own behavior. Communal satisfaction was then calculated by summing up the weighted answers from own, partner and reciprocal behavior within individuals.

Mood and state relationship satisfaction. Mood was measured by an affect grid (Russell, Weiss, & Mendelsohn, 1989), asking participants how they felt right at the time of the survey. The x-axis reflected the valence dimension of mood, ranging continuously from *unpleasant feeling* (= 0) over *neutral* (= 0.5) to *pleasant feeling* (= 1). The y-axis reflected arousal, ranging from *inactive* (= 0) over *neutral* (= 0.5) to *activated* (= 1). Examples for mood states were displayed in the edges of the grid. In the current paper, we only focus on the valence dimension, but use the arousal dimension as a control variable in a robustness analysis.

State relationship satisfaction was initially measured with two items. We

⁵This reliability reflects a measure of internal consistency. It was calculated by computing an unconditional four-level mixed model, with items on level 1, surveys on level 2, days on level 3, and individuals on level 4. To obtain a reliability coefficient, the variance of items between surveys (level 2 variance) has to be divided by the total variance on level 1 and level 2 (adjusted for the number of items), i.e. by the variance between surveys and the variance between items during one survey (Nezlek, 2017).

planned to compute a scale if the item level reliability exceeded .40. The scale did not reach this threshold, we therefore only used the single item “How do you feel about your relationship at the moment?” with answers on a continuous slider ranging from *bad* (= 1) to *exceptionally good* (= 7). In the reverse-coded discarded item, participants were asked “How annoyed are you about your partner at the moment?” with answers on a continuous slider ranging from *not at all* (= 1) to *strongly* (= 7). We conducted robustness checks with this item alone, as well as both items as a scale. All results replicated (see Footnotes 9 and 11).

Control Variables

Explicit desire for closeness. The explicit desire for closeness to one’s partner was assessed in the preliminary questionnaire with the ABC scales of social desires (Hagemeyer, Neyer, et al., 2013). Our hypotheses focus on implicit motives, but the corresponding explicit motive was applied as a covariate in a robustness analysis. On eight items, participants rated the frequency of appetitive (e.g. “I like being very close to my partner”) and aversive experiences related to closeness (e.g. “I avoid being very close to my partner”; reversed) on 7-point scales (1 = *Never*, 4 = *Sometimes*, 7 = *Always*).

Amount of time spent with partner. During experience sampling, participants answered the question “How much time did you actively spend with your partner since the last survey (technically mediated as well)?” on a continuous slider from *none at all* (= 1) over *half of the time* (= 3.5) to *all of the time* (= 7). The variable was used as a covariate in the analysis for H3.

Analysis Plan

As we preregistered the direction of the effects for our hypotheses, we used one-tailed tests and p -values for these.⁶ All other reported p -values are two-tailed, and we tagged all one-tailed p -values in the tables.

Descriptive statistics for within-subject measures were calculated on the basis of item answers aggregated within persons. To test H1 about the replication of the inter-individual association between pnCommunion and global relationship satisfaction we conducted an actor-partner interdependence model (APIM, Kenny, Kashy, & Cook, 2006) with structural equation modeling (SEM) using the *lavaan* package (Rosseel, 2012) in the *R* statistical computing

⁶An exception to this approach are the analyses for H1 and H5: Although we preregistered directional hypotheses here as well, we tested these hypotheses for three different, but correlated operationalizations of global relationship satisfaction. In the preregistration we didn’t specify any corrections for multiple testing. Therefore, at least, we chose to be conservative for these analyses by reporting two-sided p -values instead of one-sided p -values.

environment (R Core Team, 2016). APIMs account for the nonindependence of dyadic data, while estimating the effect from one’s own motive on one’s own relationship satisfaction (actor effects), and the effect of one’s own motive on the partner’s relationship satisfaction (partner effects). As we did not expect gender effects, we a priori constrained the paths to be equal for the two genders.

All other hypotheses concerned data repeatedly measured at the individual level. For these analyses we used multilevel regression models (MLMs) using the *lme4* and *lmerTest* package (Bates, Mächler, Bolker, & Walker, 2015; Kuznetsova, Brockhoff, & Christensen, 2016), to account for the nonindependence of the data, with item answers on level 1 nested within individuals on level 2. Individuals are further nested in couples on a third level, but as this level consists only of the two data points of the dyad, no within-level slope variability can be calculated. We therefore used *double-intercept-models* (Bolger & Laurenceau, 2013), creating a dummy variable for each member of the dyad based on their gender, and including these dummies in the fixed and random parts of the model.⁷ This formally results in two-level models, with separate fixed and random intercepts for each gender. Further, we *z*-standardized all continuous measures using the grand-mean and standard deviation across both genders. For analyses with predictor variables on the within-subject level, we additionally centered these variables at the individual mean, so that zero reflects a typical state for that individual.⁸ In these analyses, we controlled for the person-mean of the states at level 2 (“centered within context with reintroduction of the subtracted means at Level-2” method; Zhang, Zyphur, & Preacher, 2009, p.709). We also controlled for linear trends over time and potentially confounding variables correlated with the passage of time (Bolger & Laurenceau, 2013) by entering the index of the survey (0 = first survey). Finally, we accounted for potential differences between weekdays and weekends by entering the type of day as a dummy variable (0 = weekday, 1 = weekend).

When estimating a fixed slope for a within-subject variable that is focal to our hypothesis, we added the corresponding random slope as well (Barr, Levy, Scheepers, & Tily, 2013). We report the *marginal R*² as an effect size ($R_{\text{GLMM}(m)}^2$), representing the explained variance by the fixed effects (Johnson, 2014; Nakagawa & Schielzeth, 2013), calculated with the *MuMIn* package (Barton, 2016). For all outcomes on level 1 (H3, H4), we followed our pre-registration and excluded data for motivational states from the last survey of

⁷For H2, that is, the relationship between motives and average motivation, all variables were on level 2. Therefore we used double intercepts only for the fixed part of the models (i.e. separate means for women and men), but no double random intercepts could be included.

⁸For covariates on the within-subject level we only *z*-standardized the variables.

each day, as we did not expect the proposed process to persist overnight until the next day. When indicating the temporal sequence of surveys, we refer to any given measurement occasion as t_1 and to the next measurement occasion after t_1 as t_2 .

Results

Due to the dyadic nature of our data and the accompanying problem with anonymity, our data is available as a scientific use file which restricts access to academic users (Zygar, Hagemeyer, Pusch, & Schönbrodt, 2018). We performed all analyses in *R*, and reproducible analysis scripts can be found in the associated OSF repository (<https://osf.io/b8pu6/>). A complete description of the parameter estimates, confidence intervals and effect sizes for all following MLMs can be found in the Supplemental Materials (Tables S1-S10).

Descriptive Statistics

Tables 1 and 2 show means and standard deviations of trait and state measures, respectively. Furthermore, we computed Intra-Class-Correlations (ICCs) with an unconditional random intercept model to separate between-person and within-person variance in the state measures. The trait measures (Table 1) had high reliability estimates, and there was a low correlation between implicit and explicit motives. All ESM measures had nominally a higher amount of within-person variance compared to between-person variance (Table 2), confirming the conceptual nature of these measures as states varying over time within individuals.⁹

Table 1
Descriptive Statistics and Correlations for Trait Measures

Variables	<i>M</i> (<i>SD</i>)	Range	ω_t	1	2	3	4	5
1. Implicit pnCommunion	5.34 (2.12)	1 to 12		.23				
2. Explicit Desire for Closeness	6.16 (0.69)	3.5 to 7	.86	.20*	.10			
3. Couple Satisfaction Index	66.3 (10.23)	32 to 81	.92	.27***	.61***	.44***		
4. Positive Relationship Quality	5.86 (0.94)	1.5 to 7	.92	.21**	.48***	.58***	.34***	
5. Negative Relationship Quality	1.82 (0.99)	1 to 5.9	.91	-.23**	-.45***	-.66***	-.44***	.15

Note. $N = 152$ individuals from 77 couples. pnCommunion = partner-related need for Communion. The reliability coefficient ω_t refers to McDonald's omega total, calculated with the *MBESS* package (Kelley, 2016). Cronbach's α was equal to ω_t for all measures, except for the explicit need for closeness, α was .87 (calculated with the *psych* package, Revelle, 2016). Correlations below the diagonal refer to associations between individuals. Correlations on the diagonal refer to dyadic associations. *M* (*SD*) of pnCommunion refer to raw motive scores (number of motive categories). Correlations of pnCommunion were calculated with motive scores corrected for word count.

* $p < .05$, ** $p < .01$, *** $p < .001$

⁹Note that the within-person variance component also contains measurement error.

Table 2
Descriptive Statistics and Intra-Class Correlations for State Measures

Variable	Range	<i>Grand-Mean</i> (<i>SD_{Means}</i>)	<i>Grand-SD</i> (<i>SD_{SDs}</i>)	Women		Men	
				Between-Person Variance (ICC in %)	Within-Person + Error Variance (1-ICC in %)	Between-Person Variance (ICC in %)	Within-Person + Error Variance (1-ICC in %)
State Communion Motivation	-2 to 4	1.99 (0.83)	0.93 (0.35)	36.57	63.43	42.14	57.86
Instrumental Communal Behavior	-2 to 6.5	1.59 (0.85)	1.33 (0.33)	26.61	73.39	26.92	73.08
Communal Experiences	-3 to 11.5	2.65 (1.44)	2.14 (0.58)	28.96	71.04	28.86	71.14
Mood (Valence)	0 to 1	0.62 (0.11)	0.21 (0.05)	19.84	80.16	21.53	78.47
State Relationship Satisfaction	1 to 7	5.19 (0.69)	0.80 (0.28)	44.18	55.82	34.69	65.31

Note. $N = 130$ individuals. The Grand-Mean is the mean of the intra-individual (person) means, with the standard deviation of these (person) means from the Grand-Mean in parentheses. The Grand-SD is the mean of the intra-individual (person) standard deviations, with the standard deviation of these (person) SDs from the Grand-SD in parentheses. Intra-Class-Correlations (ICCs) were calculated with an unconditional random intercept model, with one fixed and random intercept for each gender.

H1: From Motive Dispositions to Global Relationship Satisfaction

We calculated APIMs using SEM, regressing both partners' relationship satisfaction on both partners' pnCommunion. First, we performed analyses comparing gender-constrained models with the corresponding unconstrained models with χ^2 likelihood ratio tests. As the tests indicated that the constrained models were not significantly worse, $\Delta \chi^2(2) \leq 2.79$, $ps > .247$, $\Delta AICs \leq 3.55$, we only report the results of the constrained models. As expected, we found significant actor effects, indicating that high pnCommunion was associated with one's own high global assessment of relationship satisfaction (CSI and PNRQ, see Figure 2). Exploratorily, we found one significant partner effect of pnCommunion: Individuals' own pnCommunion was positively related to their partners' global relationship satisfaction (CSI), but not to their partners' more specific measures of relationship quality (PNRQ). Detailed results can be found in Table S1.

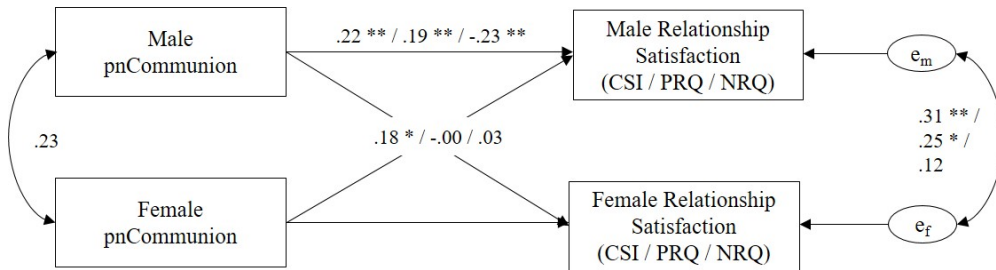


Figure 2. Path diagram of the fitted APIM depicting actor and partner effects of pnCommunion (partner-related need for communion) on three measures of global relationship satisfaction (Couple Satisfaction Index (CSI) / Positive Relationship Quality (PRQ) / Negative Relationship Quality (NRQ)). Variables were z -standardized a priori. Coefficients for actor and partner effects were constrained to be equal for both genders. $N = 74$ couples. Figure available at <https://osf.io/b8spu6/>, under a CC-BY4.0 license. * $p < .05$, ** $p < .01$

H2: From Motive Dispositions to Motivational States

Consistent with our hypothesis, implicit pnCommunion predicted mean communion motivation in the MLM, $b = 0.11$, $SE_b = 0.05$, $p_{\text{one-tailed}} = .023$. We conducted exploratory analyses to check for the incremental contribution of the implicit motive to the average occurrence of motivational states over and above explicit motives. When controlling for the explicit desire for closeness to one's partner, pnCommunion was no longer significantly related to the mean of communion motivation, $b_{\text{pnCommunion}} = 0.08$, $SE_b = 0.05$, $p_{\text{one-tailed}} = .067$. Table S2 contains the full results of these analyses.

H3: From Motivational States to Instrumental Behavior

H3 concerned the prediction of an actor’s instrumental behavior by his or her motivational state. As the variety of behaviors that can be shown depends heavily on the amount of time that was spent with the partner, we added this variable as a covariate in these analyses. The interpretation of our results does not change when omitting this variable. The results can be found in Table 3 (complete results in Table S3). Communion motivation at $t1$ significantly predicted instrumental communal behavior shown between $t1$ and $t2$. There was also an inter-individual effect, that is, individuals experiencing on average strong communion motivation showed on average also more instrumental communal behavior.

To examine the proposed direction of the intra-individual effect, we further explored whether behavior between $t1$ and $t2$ predicted motivational states at $t2$. We did find such a reversed relationship, $b = 0.18$, $SE_b = 0.02$, $p < .001$ (see Table S4). This points to possible bidirectional influences between motivational states and behavior.

Table 3
Multilevel Analyses (Fixed Effects) Predicting Instrumental Communal Behavior (z) Between Surveys by Communion Motivation (H3)

Variable	Estimate	SE	df	t value	p
Female Intercept	-0.022	0.059	83.543	-0.372	.711
Male Intercept	-0.047	0.063	74.764	-0.745	.458
Weekend Dummy ^a	-0.028	0.023	5037.613	-1.235	.217
Survey Index ^a	0.001	0.000	5003.490	1.880	.060
$t1$ - $t2$ Amount of Time Spent with Partner ^a (z)	0.511	0.011	5082.490	46.316	< .001
Mean Communion Motivation ^b (z)	0.253	0.062	129.645	4.059	< .001
$t1$ Communion Motivation^a (z)	0.073	0.021	59.343	3.531	< .001^c

Note. $N = 5153$ observations in 68 couples. $z = z$ -standardized (level 1 variables are additionally person-mean centered). The effect focal to our hypothesis is printed in bold-face. A full report including random effect variances can be found in Table S3.

^a Level 1 variable, ^b Level 2 variable, ^c This p -value is one-tailed.

H4: The Satisfaction of Motivational States in Relationships

In final confirmatory analyses, we turned to the prediction of state mood (H4A) and relationship satisfaction (H4B) by the interaction of motivational appetite and relationship experiences. The results are presented in Table 4 (Models 1A and 1B). For both the prediction of mood and relationship satisfaction, there was a significant positive main effect of communal experiences. A significant positive main effect of communal motivational states, however,

was only found in the model for relationship satisfaction. Further, communal motivational states at $t1$ significantly interacted with communal relationship experiences between $t1$ and $t2$ to predict state relationship satisfaction at $t2$ (see Figure 3):¹⁰ Individuals were most satisfied with their relationship when their motivation to be involved in communal activities had been strong and this motivation was fulfilled afterwards. Similarly, they were most dissatisfied when their motivation had been strong, but they made little communal relationship experiences. The slope for communal experiences was therefore more positive for those being highly motivated. Both simple slopes were significant, $b = 0.30$, $SE_b = 0.03$, $p < .001$ for communal motivational states 1 SD below the mean, and $b = 0.38$, $SE_b = 0.03$, $p < .001$ for communal motivational states 1 SD above the mean.

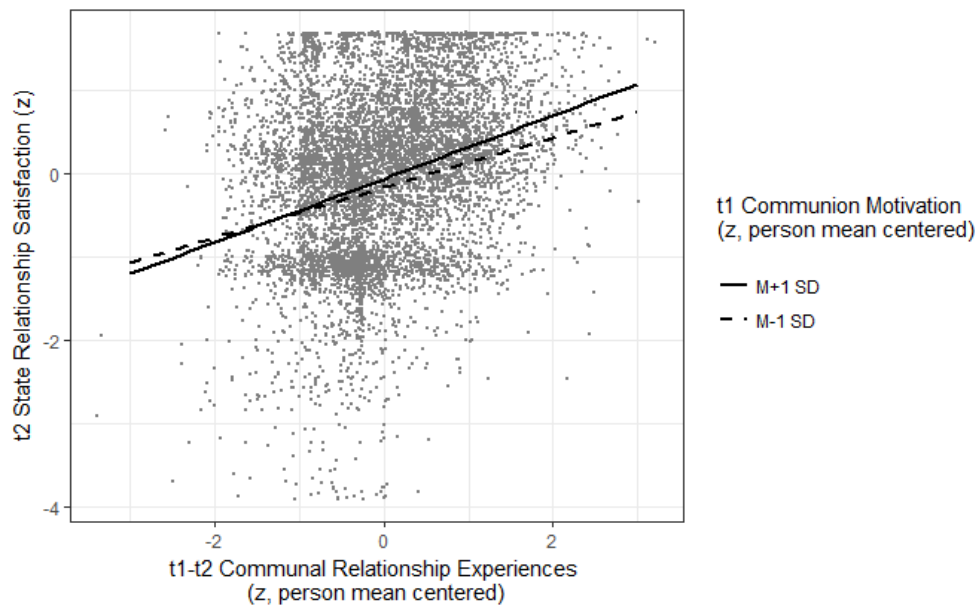


Figure 3. Prediction of state relationship satisfaction at $t2$ by the interaction of communal motivational states at $t1$ and communal experiences between $t1$ and $t2$. Figure created with ggplot2 (Wickham, 2009) and available at <https://osf.io/b8pu6/>, under a CC-BY4.0 license.

¹⁰This result was similar when using the discarded item of state relationship satisfaction instead (see Methods), $p_{\text{one-tailed}} = .041$, or both items as a scale, $p_{\text{one-tailed}} = .018$. Note that in both cases the main effect of communal motivational states was not significant, $ps > .123$.

However, no significant interaction was found when predicting individuals' mood,¹¹ which is the analysis more directly representing the affect-amplifying function described in motive theory (although the coefficient was in the expected positive direction). We explored whether this result can be attributed to gender differences, but the original model was not significantly worse than a model with gender moderating the interaction, $\Delta \chi^2(3) = 5.51$, $p = .138$, $\Delta \text{AIC} = 1$.

In exploratory analyses we substituted communal relationship experiences by a dummy variable indicating whether participants actively spent time with their partners at the moment of the survey (=1) or not (=0). This is an alternative operationalization of a potentially satisfying communal experience, albeit more imprecise because it does not specify the quality of the interaction. In contrast to the original analysis, it does not refer to a past time interval, but was assessed simultaneously with the two outcomes of interest. Therefore, the concurrent effect of this communal experience with the partner on mood and satisfaction can be evaluated. Again, we found significant positive main effects and an interaction between motivational states at $t1$ and time spent with the partner at $t2$ predicting state relationship satisfaction – but not mood – at $t2$ (see Models 2A and 2B in Table 4). Regarding mood as criterion, the model comparison pointed to a moderation by gender, as the model without this moderation fit the data significantly worse, $\Delta \chi^2(3) = 9.57$, $p = .023$, $\Delta \text{AIC} = 4$. An examination of the gender-specific model revealed that for women and for men the main effect of the partner time dummy was positive and significant. While the interaction effect for women was stronger than for men, it was also not significant (see Model 3 in Table 4). Tables S5-S9 show the complete results.

¹¹This result did not change when additionally controlling for the arousal dimension of mood, that was assessed simultaneously with the valence dimension of mood, $p_{\text{one-tailed}} = .178$.

Table 4

Multilevel Analyses (Fixed Effects) for Predicting Mood and State Relationship Satisfaction by the Interaction of Communion Motivation and Communal Experiences

Models 1A and 1B for H4A and H4B	Mood (valence) at $t2$ (z)				State relationship satisfaction at $t2$ (z)					
	Estimate	SE	df	t value	p	Estimate	SE	df	t value	p
Female Intercept	-0.190	0.059	90.986	-3.221	.002	-0.095	0.079	76.125	-1.203	.233
Male Intercept	0.032	0.066	80.538	0.487	.628	-0.130	0.072	68.201	-1.818	.073
Weekend Dummy ^a	0.118	0.027	4864.060	4.295	< .001	0.055	0.023	4862.332	2.403	.016
Survey Index ^a	0.002	0.001	4783.977	3.025	.002	0.002	0.001	4825.007	4.015	< .001
Mean Communion Motivation ^b (z)	0.010	0.065	119.991	0.155	.877	0.208	0.077	123.918	2.708	.008
Mean Communal Experience ^b (z)	0.238	0.078	105.326	3.035	.003	0.291	0.096	118.805	3.040	.003
$t1$ Communion Motivation ^a (z)	0.010	0.020	86.669	0.516	.607	0.053	0.020	56.722	2.624	.011
$t1-t2$ Communal Experience ^a (z)	0.254	0.022	64.589	11.655	< .001	0.341	0.024	60.637	14.349	< .001
$t1$ Communion Motivation ^a (z) \times $t1-t2$ Communal Experience ^a (z)	0.024	0.024	59.112	1.019	.156^c	0.040	0.018	106.953	2.234	.014^c
Models 2A and 2B	Estimate	SE	df	t value	p	Estimate	SE	df	t value	p
Female Intercept	-0.296	0.061	89.800	-4.863	< .001	-0.194	0.084	75.975	-2.299	.024
Male Intercept	-0.088	0.069	78.796	-1.274	.206	-0.258	0.078	67.269	-3.312	.001
Weekend Dummy ^a	0.127	0.028	4792.692	4.505	< .001	0.081	0.024	4896.264	3.331	< .001
Survey Index ^a	0.002	0.001	4743.701	3.366	.001	0.002	0.001	4875.379	4.062	< .001
Mean Communion Motivation ^b (z)	0.109	0.063	122.358	1.730	.086	0.271	0.074	122.682	3.686	< .001
$t1$ Communion Motivation ^a (z)	0.004	0.027	59.880	0.165	.870	0.054	0.026	73.622	2.042	.045
$t2$ Partner Time Dummy ^a	0.368	0.042	58.858	8.847	< .001	0.394	0.046	55.510	8.497	< .001
$t1$ Communion Motivation ^a (z) \times $t2$ Partner Time Dummy ^a	0.069	0.043	107.527	1.609	.111	0.104	0.034	624.172	3.057	.002
Model 3	Estimate	SE	df	t value	p					
Female Intercept	-0.318	0.061	92.018	-5.183	< .001					
Male Intercept	-0.058	0.070	81.728	-0.823	.413					
Weekend Dummy ^a	0.125	0.028	4796.788	4.421	< .001					
Survey Index ^a	0.002	0.001	4743.415	3.310	< .001					
Mean Communion Motivation ^b (z)	0.110	0.063	122.595	1.745	.083					
$t1$ Female Communion Motivation ^a (z)	0.002	0.032	118.554	0.074	.941					
$t1$ Male Communion Motivation ^a (z)	0.009	0.036	153.418	0.258	.797					
$t2$ Female Partner Time Dummy ^a	0.448	0.050	118.618	8.992	< .001					
$t2$ Male Partner Time Dummy ^a	0.279	0.052	130.885	5.374	< .001					
$t1$ Female Communion Motivation ^a (z) \times $t2$ Female Partner Time Dummy ^a	0.100	0.056	286.382	1.778	.077					
$t1$ Male Communion Motivation ^a (z) \times $t2$ Male Partner Time Dummy ^a	0.043	0.059	304.412	0.734	.463					

Note. $N = 5035$ (Model 1) and 5083 (Models 2 and 3) observations in 68 couples. Motivation refers to motivational appetence. $z = z$ -standardized (focal level 1 variables are additionally person-mean centered). The effects focal to our hypotheses are printed in boldface. A full report including random effect variances can be found in Tables S5-S9. ^a Level 1 variable, ^b Level 2 variable, ^c One-tailed p s.

H5: From Relationship Satisfaction States to Global Assessments

Multilevel regression model analyses confirmed our assumption that global relationship satisfaction can be predicted by average relationship satisfaction states, for the Couple Satisfaction Index, $b = 0.61$, $SE_b = 0.13$, $p < .001$, for Positive Relationship Quality, $b = 0.45$, $SE_b = 0.14$, $p = .002$, and Negative Relationship Quality, $b = -0.56$, $SE_b = 0.14$, $p < .001$ (Table S10).¹² Note that global relationship satisfaction was assessed before measuring average relationship states. However, we assume that global relationship satisfaction is relatively stable, thus warranting the presented analysis.

Discussion

Drawing on motive disposition theory (McClelland, 1987), the Rubicon model of action phases (Back & Vazire, 2015; Heckhausen & Heckhausen, 2008), and the Zurich Model of Social Motivation (Bischof, 1975), this study suggests the “DynaMoS” model (Dynamics of Motive Satisfaction) illustrated in Figure 1, that can be used to examine intra-individual motivational processes in couple relationships. Applying an experience sampling approach we tested several focal paths of the model in the domain of communion motives. We not only replicated previous findings on positive inter-individual associations between the implicit partner-related need for communion and relationship satisfaction; the data also supported most preregistered hypotheses regarding the intra-individual process model: Individuals who have a stronger disposition to strive for communal experiences feel the urge to seek out emotional closeness more frequently in everyday life. When motivated in that regard, they also behave in a communal way more often. The communal experiences individuals make as a consequence of their own and their partner’s behavior generally improve their mood and their momentary relationship satisfaction. Moreover, these experiences are even more beneficial for state relationship satisfaction if individuals had previously experienced strong motivation for closeness, compared to weak previous motivation. In turn, individuals who had frequent and intense experiences of state satisfaction reported higher global relationship quality than others.

The Emergence of Motivational States

Individuals with a strong dispositional communion motive (pnCommunion) experienced more communal motivational states across two weeks. However, this contribution of pnCommunion to the prediction of mean motivation

¹²The results were similar when using the discarded item of state relationship satisfaction or both items as a scale instead (see Methods), $ps < .001$.

was not incremental to the contribution of the explicit relationship-specific desire for closeness. This is a very strict test, as the desire for closeness is measured as the typical frequency of motivational states and is therefore conceptually similar with the average of the self-reported motivational states. In addition, our sample size might have not provided enough statistical power to detect an incremental contribution, therefore this question must be examined in a larger sample.

Further, it is theoretically assumed that implicit motives translate to self-reported motivational states if barriers hinder their fulfillment (Bischof, 2008). Accordingly, Hagemeyer et al. (2015) found in the complementary domain of agency that a strong implicit motive in men predicted higher agentic states only in potentially frustrating living arrangements. The current study did not look at such interactions between disposition and situation, a dynamic that has yet to be considered. We also did not examine motivational incentives that were present in the situation, being also relevant to the emergence of motivational states (McClelland, 1987). Nonetheless, we found that on average implicit communion motives are represented in self-reported motivational states.

Motivated Behavior

The experience of motivational states predicted instrumental communal behavior that was shown afterwards. This means that a current desire for emotional closeness was followed by the implementation of behavior that served this goal within a short time span of a few hours. This is not self-evident, as many circumstances can undermine the implementation of individuals' motivation, such as situational barriers or competing desires.

We found in exploratory analyses that communal behavior also predicted future motivational states – that is, when individuals acted communally they were motivated to continue receiving fulfilling experiences. It has to be mentioned that looking at this direction of the process means to look at variables that were assessed during the same survey: We asked participants how they behaved since the last survey during the same survey they indicated their momentary motivational state. This simultaneous assessment could have biased the answers. Still, this result is consistent with a plausible bidirectional influence between motivational states and behavior, at least on a short time scale: Not only does communion motivation lead to according behavior; communally satisfying behavior also reinforces the motivation to get more of the same.

Applying systems theory, though, at a larger time scale we would expect that motivation decreases after (enough) consummatory experiences, that is when the perfect level of a specific state is achieved. Thinking further, we

would even expect aversive reactions (“too much closeness”) that translate into the wish to avoid further communal experiences. It is therefore noteworthy that our analyses were limited to associations between adjacent time points. Thus, the extent of communal experiences reported at one time point was likely not sufficient to satisfy the previously reported motivation, which therefore persisted. At this point, little is known about the time scale of motivational arousal and satisfaction (see also Hagemeyer et al., 2015). Future studies using the ESM should aim at a better understanding of this basic feature of motivational dynamics.

The Affect-Amplifying Nature of Motivation

Situations and behaviors that satisfy communion motivation generally increased emotional valence and state relationship satisfaction. Moreover, and crucial for the test of our “motivation as affect-amplifier” hypothesis, we found that communal experiences were more rewarding for individuals if they were highly motivated before. This result, however, was only found for state relationship satisfaction (H4B), not for general emotional valence (H4A). The latter outcome measure, however, is conceptually closer to the actual affective experience, which is central to the original formulation of motives as affect-amplifiers (McClelland, 1987; Schultheiss & Wirth, in press; Schultheiss, 2008). Hence, this hypothesis could only be partly confirmed.

As a post-hoc explanation, one could argue that the measurement of state relationship satisfaction is conceptually closer to the partner-related communion motive than the more domain-general mood assessment. Hence, one could assume a motive-specific version of the affect-amplification hypothesis, namely that a motive or a motivational state amplifies predominantly *motive-specific* emotions, such as joy and love in the domain of communion (see also Job et al., 2012; Zurbriggen & Sturman, 2002).

Nonetheless, general mood should still be affected according to motive theory (McClelland, 1987; Schultheiss & Wirth, in press; Schultheiss, 2008). It could be that too much time passed between the positively valenced experience and the moment we asked for individuals’ mood: Participants reported on the behavior that happened in the last couple of hours, but indicated their momentary mood. Compared with state relationship satisfaction, mood is more volatile (see within-person variances in Table 2) and might be more susceptible to influences outside of the relationship (e.g. the satisfaction of a certain motivation by other persons), while relationship satisfaction is mainly driven by experiences inside the relationship. On the other hand, the main effect of communal experiences was of similar strength for both outcomes.

We therefore additionally examined the interaction with a communally satisfying measure that was assessed simultaneously with mood: whether individuals actively spent time with their partner at the moment of the survey. Whereas the interaction with this alternative measure was in the expected direction, it was not significant and, if at all, only present for women. Relationship satisfaction, in contrast, showed the expected result irrespective of gender. The momentary involvement with the partner, though, is not an optimal measure of communal satisfaction – individuals might have had an argument and would still indicate that they had spent time with their partner. Other situational influences on participants' mood and the time passed between communal behavior and mood assessment might therefore play a role in the unexpected results.

It has to be mentioned, that the fixed interaction effect on relationship satisfaction was rather small on an absolute scale, and might be as small for mood. One reason for not finding an effect for mood might therefore be that the statistical power provided by the current sample size was not sufficient to reliably detect similar interaction effects for mood. Furthermore, there were substantial inter-individual random variations of the main effect of motivation (95% range of random slopes: -0.17 to 0.28) and the focal interaction effect (95% range of random interaction coefficients: -0.07 to 0.15). Hence, there is a lot of unexplained inter-individual variation in the strength of these associations.

Taken together, we could not find that fulfilling experiences boosted general mood more strongly for motivated individuals than for those who do not care (H4A). But we found this amplifying effect of motivation with regard to the more motive-specific measure of momentary relationship satisfaction (H4B).

The Relationship Between State and Global Relationship Satisfaction

In a final step we showed that individuals' average relationship satisfaction states retrodicted a global measure of relationship satisfaction (the CSI) and a more specific measure representing the positive and negative assessments of the relationship (the PNRQ). In total, the results point to the fact that inter-individual differences in relationship satisfaction are associated with the average experience of relationship satisfaction states.

Study Limitations and Recommendations for Future Research

Our results have to be interpreted under the consideration of some limitations. For one thing, our sample consisted of primarily young, highly educated, heterosexual, happy couples who mostly did not have children. These individuals were mostly at the beginning of their relationships and at different phases in their lives than for instance couples having been married for years or those raising kids. Yet, we expect the DynaMoS model to reflect basic motivational processes, that should also be observable in couples who are in different circumstances. Especially in difficult stages of a relationship, for example when having to step back more from one's own needs to meet the demands of a newborn, the motivational dynamics should become visible even more strongly. Further, we excluded homosexual couples solely for methodological reasons, as distinguishable dyads allowed us to compute double-intercept-models (Bolger & Laurenceau, 2013). A study analyzing differences between hetero- and homosexual couples showed that the factors predicting relationship quality can be generalized (Kurdek, 2004). We therefore expect the motivational relationship dynamics to be similar for homosexual couples.

Potential biases that might come along with our study should be considered: All variables, except for the implicit motives, were measured by self-report. Some of our results could therefore be inflated through shared method variance. To overcome this bias, future studies could use smartphones not only for ESM, but also as a tool to complement self-report with (more objective) behavioral data, such as logging the actual contacts individuals have with their partners via telephone, messenger or assessing the proximity to the partner via bluetooth or geopositioning (cf. Miller, 2012; Harari, Gosling, Wang, & Campbell, 2015).

Furthermore, the items we used to measure communion motivation did not have a good reliability in terms of internal consistency at the intra-individual level. This lack of reliability could in turn lead to an underestimation of the true effects. With regard to the validity of the items, we asked participants in the feedback questionnaire to give examples of situations in which they indicated to be motivated and what exactly they wished for. This qualitative data supports the assumption that participants did at least understand the items correctly. Generally, the ESM approach justifies the generalization to real life situations more strongly than artificial laboratory studies or questionnaire vignettes. Individuals report on their feelings and behavior in their usual living context without the researcher being present. The prompt response to their momentary situations is less susceptible to recall bias com-

pared to questionnaires usually asking participants to mentally average across a large variety of situations and report on these averages (Bolger, Davis, & Rafaeli, 2003).

Finally, the unequal time intervals between surveys ($M = 3$ hours, $SD = 47$ minutes, Range = 1 to 6 hours) were ignored in the lagged analyses. As the timing is a crucial component for lagged analyses in general (Bolger & Laurenceau, 2013; Collins & Graham, 2002; Gollob & Reichardt, 1987) and the dynamics we look at in particular, these unequally spaced intervals may have created some distortion for which the current analyses did not account for. It is therefore very important when designing future studies to consider carefully different time intervals.

The available results provide evidence for the proposed DynaMoS model in the domain of communion. Although we preregistered the model, hypotheses and analysis strategy, a direct replication of our results is necessary. We expect that the model generally applies to other motive domains, such as agency, which has also been shown to be an important motivational factor for relationship functioning (Hagemeyer et al., 2015). Additionally, explicit motives play an important role in relationships (Hagemeyer, Neyer, et al., 2013) and deserve separate attention. In this regard it is relevant that our study worked with self-reported motivational states, relating them to implicit motives. Future research should be devoted to ways of measuring implicit motivational states, for example by using contingencies between motivationally relevant stimuli and affective or behavioral reactions. This was already done for the dispositional affiliation motive in a study of Dufner et al. (2015), but has yet to be investigated for motivational states. The DynaMoS model can further be adapted to motivational dynamics in other contexts than couple relationships, for example in friendships (see e.g. McAdams, Healy, & Krause, 1984). It could also be used to take a look at the circumstances under which physiological expressions of affect (e.g. smiling) instead of relationship satisfaction is influenced by the interaction of motivational states and situations (see e.g. Fodor & Wick, 2009).

The current study was based on data from dyads, but we primarily focused on intra-individual processes and investigated no within-dyad hypotheses. The partner is not only needed for the satisfaction of social motives, his or her own motives and experiences constitute a strong dynamical, situational factor in itself. For instance, Kanat-Maymon et al. (2017) recently showed that contingent positive or negative reactions from the partner to the satisfaction or frustration of needs (“conditional positive or negativ regard”) plays a crucial role in explaining relationship satisfaction at the intra-individual, as well as

the inter-individual level. We already found at the inter-individual level that individuals' pnCommunion is positively related to their partners' global relationship satisfaction, thereby replicating prior results (Hagemeyer & Neyer, 2012). Looking at the interaction between individuals' and their partners' motivational states, behavior, and reactions at the state level could produce further insights about the dyadic processes that occur in couple relationships and contribute to such partner effects.

Finally, the DynaMoS model (Figure 1) is not intended to and does by far not capture all processes that could be derived from motive theories and existing empirical research. Motivational processes can be examined from different perspectives applying higher or lower resolutions. In the DynaMoS model, each variable and each path could be decomposed into more specific components and associations, respectively, to achieve a higher resolution. For instance, although theoretically relevant, we omitted analyses of the distinction between approach and avoidance motivation and also did not analyze all of the potential moderating situational influences. Neither did we account for all theoretically plausible feedback loops and reciprocal paths. It should be kept in mind that theoretical models can never capture the complex reality of psychological phenomena as a whole. Rather, they depict theoretically sound assumptions about such phenomena that are never comprehensive and always simplified. The usefulness of a specific model with a specific level of resolution is (among other factors) determined by the research question at hand. Therefore, we focused our hypotheses and analyses to what is presented in this article. Future research could turn to motivational dynamics on other levels of resolution that are still consistent with the model presented here.

The Explanatory Power and Limitations of the Experience Sampling Approach

While intensive longitudinal data on individuals' states collected with the Experience Sampling Method (Csikszentmihalyi & Larson, 1987) is observational and does not allow direct causal statements as in "when manipulating x, changes in y result", it still has several strengths with regard to the interpretations that are possible: The repeated sampling from the same individuals allows to make statements about the within-subject effects, and to separate them from the between-subject effects. Conclusions can then be drawn about which associations hold at which level, and if they differ, how they differ. Statements at the between-subject level refer to average states or stable traits (e.g. "a person who feels *on average* more motivated / has a *stronger motive disposition* than another person, is on average more satisfied"). Statements

on the within-subject level, however, refer to relationships between temporally changing states, either collected at a single time point or at different time points (e.g. “a person who feels *at one moment* more motivated than usual, is typically more satisfied *at that moment / the next moment*”). These interpretations have different implications for theory and interventions, therefore it is important to disentangle these variances.

It has been argued that the temporal sequence of observed processes can serve as a “proxy for causality” (Nezlek et al., 2017, p.3). When doing lagged analyses, i.e. investigating participants’ answers at different time points, one can examine whether the relationship holds up when reversing the order of variables. For example, we hypothesized that motivational states at one time point predicted future behavior reported at the next time point. We also looked at the alternative direction of the effect, i.e. behavior predicting future motivational states. The results suggest that the influence is bidirectional; or that the motivational activation persists for a longer time, implying that not enough satisfying experiences took place during the time span we looked at. Had only one of these relationships persisted in the analyses, the temporal precedence would indicate which variable might cause the other variable (Hamaker, Kuiper, & Grasman, 2015; Nezlek, 2017; West & Hepworth, 1991). In addition, if researchers are interested in changes on a specific variable, analyses can include the lagged dependent variable as predictor (“autoregression”, see Gollob & Reichardt, 1987; Hamaker & Grasman, 2015; Hamaker et al., 2015). This is meant to control for the stability of the criterion over time, which is a potential confound. Additionally though, it assumes equally spaced intervals (Bolger & Laurenceau, 2013; Nezlek, 2017) – which was not the case in the present study – and alters the interpretation of results, because rather than predicting absolute values, relative values contingent on previous values (i.e. changes) are considered.

It is important to reiterate that data from ESM studies cannot provide the causal interpretations that are often warranted by randomized experiments (Bolger & Laurenceau, 2013; but see Deaton & Cartwright, 2017 who emphasize that even with randomization one still has to consider confounding, post-randomization differences or the representativeness of the sample to be able to draw meaningful conclusions). In purely within-subject MLM analyses of observational ESM data, individuals are held constant (which leads to “participants as their own controls”, Bolger et al., 2003, p.587), and changes on variables over time are addressed – contingent on individuals’ own averages on specific variables. The problem of unobserved confounds (which is usually tackled by randomization in experiments) is present in such analyses, but lies

on the within-subject-level. Here, confounding variables are not randomly distributed across the levels of a state variable the researcher is interested in. If, for instance, an unobserved confounding variable co-occurs with motivation and is related to subsequent behavior, one would wrongfully assign the cause to motivation and not to the disregarded confound. Bolger and Laurenceau (2013, p.71) suggest including a variable representing the elapsed time of the study (as was done in the present study), to control for influences that are due to the duration of data collection (e.g. “measurement-as-treatment”, Collins & Graham, 2002, p.95). In the next step, to achieve an improved causal interpretation with longitudinal data, the assumed predictor would have to be manipulated accordingly and subsequent changes in the criterion (measured at the within-subject-level) would have to be monitored.

Conclusion

Intensively measuring variation of individuals’ behavior, motivational states, feelings, and experiences is an important endeavor when trying to understand differential outcomes of dispositions. With a process model (“DynaMoS”) as theoretical framework we showed that these kind of measures can help to uncover how different states are dynamically intertwined in couple relationships. Our results support the conception that implicit motives are reflected in motivational states, that these motivational states direct individuals’ behavior towards satisfying end-states, and that motivational states boost the positive effect communal experiences have on individuals’ evaluation of their relationship. We suggest that the DynaMoS model can be applied to further research on motives, while at the same time serving as an example of how to approach the thorough study of intra-individual processes, and how to connect them to inter-individual differences and associations.

References

- Arslan, R., & Tata, C. (2016). formr.org survey software (version v0.16.12). Retrieved from <https://formr.org/>
- Augustine, A. A., & Larsen, R. J. (2012). Is a trait really the mean of states? Similarities and differences between traditional and aggregate assessments of personality. *Journal of Individual Differences, 33*, 131–137. doi:10.1027/1614-0001/a000083
- Back, M. D., Baumert, A., Denissen, J. J. A., Hartung, F. M., Penke, L., Schmukle, S. C., . . . et al. (2011). Persoc: A unified framework for understanding the dynamic interplay of personality and social relationships. *European Journal of Personality, 25*, 90–107. doi:10.1002/per.811
- Back, M. D., & Vazire, S. (2015). The social consequences of personality: Six suggestions for future research. *European Journal of Personality, 29*, 296–307. doi:10.1002/per.1998
- Bakan, D. (1966). *The duality of human existence*. Boston, IL: Beacon Press.
- Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language, 68*, 1–43. doi:10.1016/j.jml.2012.11.001
- Barton, K. (2016). *Mumin: Multi-model inference*. R package version 1.15.6. Retrieved from <https://CRAN.R-project.org/package=MumIn>
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*, 1–48. doi:10.18637/jss.v067.i01
- Bischof, N. (1975). A systems approach toward the functional connections of attachment and fear. *Child Development, 46*, 801–817.
- Bischof, N. (1995). *Struktur und Bedeutung: Eine Einführung in die Systemtheorie*. Bern: Huber.
- Bischof, N. (2008). *Psychologie: Ein Grundkurs für Anspruchsvolle*. Stuttgart: Kohlhammer.
- Bolger, N., Davis, A., & Rafaeli, E. (2003). Diary methods: Capturing life as it is lived. *Annual Review of Psychology, 54*, 579–616. doi:10.1146/annurev.psych.54.101601.145030
- Bolger, N., & Laurenceau, J.-P. (2013). *Intensive longitudinal methods*. New York, NY: Guilford.
- Brunstein, J. C., Schultheiss, O. C., & Grassmann, R. (1998). Personal goals and emotional well-being: The moderating role of motive dispositions. *Journal of Personality and Social Psychology, 75*, 494–508. doi:10.1037/0022-3514.75.2.494

- Collins, L. M., & Graham, J. W. (2002). The effect of the timing and spacing of observations in longitudinal studies of tobacco and other drug use: Temporal design considerations. *Drug and Alcohol Dependence*, *68*, 85–96. doi:10.1016/S0376-8716(02)00217-X
- Csikszentmihalyi, M., & Larson, R. (1987). Validity and reliability of the experience-sampling method. *Journal of Nervous and Mental Disease*, *175*, 526–536.
- Curran, P. J., & Bauer, D. J. (2011). The disaggregation of within-person and between-person effects in longitudinal models of change. *Annual Review of Psychology*, *62*, 583–619. doi:10.1146/annurev.psych.093008.100356.
- Deaton, A., & Cartwright, N. (2017). Understanding and misunderstanding randomized controlled trials. *Social Science & Medicine*. doi:https://doi.org/10.1016/j.socscimed.2017.12.005
- Debrot, A., Schoebi, D., Perrez, M., & Horn, A. B. (2013). Touch as an interpersonal emotion regulation process in couples' daily lives: The mediating role of psychological intimacy. *Personality and Social Psychology Bulletin*, *39*, 1373–85. doi:10.1177/0146167213497592
- Dufner, M., Arslan, R. C., Hagemeyer, B., Schönbrodt, F. D., & Denissen, J. J. A. (2015). Affective contingencies in the affiliative domain: Physiological assessment, associations with the affiliation motive, and prediction of behavior. *Journal of Personality and Social Psychology*, *109*, 662–676. doi:10.1037/pspp0000025
- Finnigan, K. M., & Vazire, S. (2017). The incremental validity of average state self-reports over global self-reports of personality. *Journal of Personality and Social Psychology*. doi:10.1037/pspp0000136
- Fleeson, W. (2001). Toward a structure-and process-integrated view of personality: Traits as density distributions of states. *Journal of Personality and Social Psychology*, *80*, 1011–1027. doi:10.1037/0022-3514.80.6.1011
- Fleeson, W., & Gallagher, P. (2009). The implications of Big Five standing for the distribution of trait manifestation in behavior: Fifteen experience-sampling studies and a meta-analysis. *Journal of Personality and Social Psychology*, *97*, 1097–1114. doi:10.1037/a0016786
- Fodor, E. M., & Wick, D. P. (2009). Need for power and affective response to negative audience reaction to an extemporaneous speech. *Journal of Research in Personality*, *43*, 721–726. doi:10.1016/j.jrp.2009.06.007
- Funk, J. L., & Rogge, R. D. (2007). Testing the ruler with item response theory: Increasing precision of measurement for relationship satisfaction

- with the Couples Satisfaction Index. *Journal of Family Psychology*, *21*, 572–583. doi:10.1037/0893-3200.21.4.572
- Gable, S. L., & Poore, J. (2008). Which thoughts count? Algorithms for evaluating satisfaction in relationships. *Psychological Science*, *19*, 1030–1036. doi:10.1111/j.1467-9280.2008.02195.x
- Gollob, H. F., & Reichardt, C. S. (1987). Taking account of time lags in causal models. *Child Development*, *58*, 80–92.
- Hagemeyer, B., Dufner, M., & Denissen, J. J. A. (2016). Double dissociation between implicit and explicit affiliative motives: A closer look at socializing behavior in dyadic interactions. *Journal of Research in Personality*, *65*, 89–93. doi:10.1016/j.jrp.2016.08.003
- Hagemeyer, B., Neberich, W., Asendorpf, J. B., & Neyer, F. J. (2013). (In)Congruence of implicit and explicit communal motives predicts the quality and stability of couple relationships. *Journal of Personality*, *81*, 390–402. doi:10.1111/jopy.12016
- Hagemeyer, B., & Neyer, F. J. (2012). Assessing implicit motivational orientations in couple relationships: The Partner-Related Agency and Communion Test (PACT). *Psychological Assessment*, *24*, 114–28. doi:10.1037/a0024822
- Hagemeyer, B., Neyer, F. J., Neberich, W., & Asendorpf, J. B. (2013). The ABC of Social Desires: Affiliation, being alone, and closeness to partner. *European Journal of Personality*, *27*, 442–457. doi:10.1002/per.1857
- Hagemeyer, B., Schönbrodt, F. D., Neyer, F. J., Neberich, W., & Asendorpf, J. B. (2015). When "together" means "too close": Agency motives and relationship functioning in coresident and living-apart-together couples. *Journal of Personality and Social Psychology*, *109*, 813–835. doi:10.1037/pspi0000031
- Hamaker, E. L. (2012). Why researchers should think “within-person”: A paradigmatic rationale. In M. R. Mehl & T. S. Conner (Eds.), *Handbook of research methods for studying daily life* (pp. 43–61). Guilford. Retrieved from <http://www.geo.uzh.ch/~screde/share/ReadingGroup/Hamaker-12.pdf>
- Hamaker, E. L., & Grasman, R. P. P. P. (2015). To center or not to center? Investigating inertia with a multilevel autoregressive model. *Frontiers in Psychology*, *5*, 1–15. doi:10.3389/fpsyg.2014.01492
- Hamaker, E. L., Kuiper, R. M., & Grasman, R. P. P. P. (2015). A critique of the cross-lagged panel model. *Psychological Methods*, *20*, 102–116. doi:10.1037/a0038889

- Harari, G. M., Gosling, S. D., Wang, R. U. I., & Campbell, A. T. (2015). Comment Capturing Situational Information with Smartphones and Mobile Sensing Methods. *European Journal of Personality, 29*, 509–511. doi:10.1002/per.2032
- Heckhausen, J., & Heckhausen, H. (2008). *Motivation and action* (2nd ed.). Cambridge University Press.
- Heller, D., Watson, D., & Ilies, R. (2004). The role of person versus situation in life satisfaction: A critical examination. *Psychological Bulletin, 130*, 574–600. doi:10.1037/0033-2909.130.4.574
- Heyman, R. E., Sayers, S. L., & Bellack, A. S. (1994). Global marital satisfaction versus marital adjustment: An empirical comparison of three measures. *Journal of Family Psychology, 8*, 432–446.
- Hofer, J., & Busch, H. (2011a). Satisfying one's needs for competence and relatedness: Consequent domain-specific well-being depends on strength of implicit motives. *Personality and Social Psychology Bulletin, 37*, 1147–1158. doi:10.1177/0146167211408329
- Hofer, J., & Busch, H. (2011b). When the needs for affiliation and intimacy are frustrated: Envy and indirect aggression among German and Cameroonian adults. *Journal of Research in Personality, 45*, 219–228. doi:10.1016/j.jrp.2011.02.003
- Hofer, J., & Hagemeyer, B. (in press). Social bonding: Affiliation motivation and intimacy motivation. In J. Heckhausen & H. Heckhausen (Eds.), *Motivation and action* (3rd ed.). Berlin, Germany: Springer.
- Hofmann, W., Baumeister, R. F., Förster, G., & Vohs, K. D. (2012). Everyday temptations: An experience sampling study of desire, conflict, and self-control. *Journal of Personality and Social Psychology, 102*, 1318–35. doi:10.1037/a0026545
- Hofmann, W., Finkel, E. J., & Fitzsimons, G. M. (2015). Close relationships and self-regulation: How relationship satisfaction and self-regulation facilitates momentary goal pursuit. *Journal of Personality and Social Psychology, 109*, 434–452.
- Horowitz, L. M., Wilson, K. R., Turan, B., Zolotsev, P., Constantino, M. J., & Henderson, L. (2006). How interpersonal motives clarify the meaning of interpersonal behavior: A revised circumplex model. *Personality and Social Psychology Review, 10*, 67–86. doi:10.1207/s15327957pspr1001_4
- Howell, R. T., Ksendzova, M., Nestingen, E., Yerahian, C., & Iyer, R. (2017). Your personality on a good day: How trait and state personality predict daily well-being. *Journal of Research in Personality, 69*, 250–263. doi:10.1016/j.jrp.2016.08.001

- Job, V., Bernecker, K., & Dweck, C. S. (2012). Are implicit motives the need to feel certain affect? Motive–affect congruence predicts relationship satisfaction. *Personality and Social Psychology Bulletin*, *38*, 1552–1565. doi:10.1177/0146167212454920
- Johnson, P. C. (2014). Extension of nakagawa and schielzeth’s r2glmm to random slopes models. *Methods in Ecology and Evolution*, *5*, 944–946. doi:10.1111/2041-210X.12225
- Kanat-Maymon, Y., Argaman, Y., & Roth, G. (2017). The association between conditional regard and relationship quality: A daily diary study. *Personal Relationships*, *24*, 27–35. doi:10.1111/per.12164
- Kelley, K. (2016). *Mbess: The mbess R package*. R package version 4.1.0. Retrieved from <https://CRAN.R-project.org/package=MBESS>
- Kenny, D. A., Kashy, D. A., & Cook, W. L. (2006). *Dyadic Data Analysis*. New York, NY: Guilford.
- Kurdek, L. A. (2004). Are gay and lesbian cohabiting couples really different from heterosexual married couples? *Journal of Marriage and Family*, *66*, 880–900.
- Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2016). *Lmertest: Tests in linear mixed effects models*. R package version 2.0-33. Retrieved from <https://CRAN.R-project.org/package=lmerTest>
- Laurenceau, J.-P., Feldman Barrett, L., & Pietromonaco, P. R. (1998). Intimacy as an interpersonal process: The importance of self-disclosure, partner disclosure, and perceived partner responsiveness in interpersonal exchanges. *Journal of Personality and Social Psychology*, *74*, 1238–1251. doi:10.1037/0022-3514.74.5.1238
- Laurenceau, J.-P., Feldman Barrett, L., & Rovine, M. (2005). The interpersonal process model of intimacy in marriage: A daily-diary and multilevel modeling approach. *Journal of Family Psychology*, *19*, 314–323. doi:10.1037/0893-3200.19.2.314
- Laurenceau, J.-P., Troy, A. B., & Carver, C. S. (2005). Two distinct emotional experiences in romantic relationships: Effects of perceptions regarding approach of intimacy and avoidance of conflict. *Personality and Social Psychology Bulletin*, *31*, 1123–1133. doi:10.1177/0146167205274447
- Le, B., & Agnew, C. R. (2001). Need fulfillment and emotional experience in interdependent romantic relationships. *Journal of Social and Personal Relationships*, *18*, 423–440. doi:10.1177/0265407501183007
- Le, B., & Agnew, C. R. (2003). Commitment and its theorized determinants: A meta-analysis of the Investment Model. *Personal Relationships*, *10*, 37–57. doi:10.1111/1475-6811.00035

- Malouff, J. M., Thorsteinsson, E. B., Schutte, N. S., Bhullar, N., & Rooke, S. E. (2010). The Five-Factor Model of personality and relationship satisfaction of intimate partners: A meta-analysis. *Journal of Research in Personality, 44*, 124–127. doi:10.1016/j.jrp.2009.09.004
- Markey, P. M., Funder, D. C., & Ozer, D. J. (2003). Complementarity of Interpersonal Behaviors in Dyadic Interactions. *Personality and Social Psychology Bulletin, 29*, 1082–1090. doi:10.1177/0146167203253474
- Mason, A., & Blankenship, V. (1987). Power and affiliation motivation, stress, and abuse in intimate relationships. *Journal of Personality and Social Psychology, 52*, 203–210. doi:10.1037/0022-3514.52.1.203
- McAdams, D. P., & Constantian, C. A. (1983). Intimacy and affiliation motives in daily living: An experience sampling analysis. *Journal of Personality and Social Psychology, 45*, 851–861. doi:10.1037/0022-3514.45.4.851
- McAdams, D. P., Healy, S., & Krause, S. (1984). Social motives and patterns of friendship. *Journal of Personality and Social Psychology, 47*, 828–838. doi:10.1037/0022-3514.47.4.828
- McAdams, D. P., & Vaillant, G. E. (1982). Intimacy motivation and psychosocial adjustment: A longitudinal study. *Journal of Personality Assessment, 46*, 586–593. doi:10.1207/s15327752jpa4606_5
- McAdams, D. P., Jackson, R. J., & Kirshnit, C. (1984). Looking, laughing, and smiling in dyads as a function of intimacy motivation and reciprocity. *Journal of Personality, 52*, 261–273.
- McClelland, D. C. (1987). *Human Motivation*. New York: Cambridge University Press.
- McClelland, D. C., Koestner, R., & Weinberger, J. (1989). How do self-attributed and implicit motives differ? *Psychological Review, 96*, 690–702. doi:10.1037/0033-295X.96.4.690
- Miller, G. (2012). The Smartphone Psychology Manifesto. *Perspectives on Psychological Science, 7*, 221–237. doi:10.1177/1745691612441215
- Molenaar, P. C. M. (2004). A manifesto on psychology as idiographic science: Bringing the person back into scientific psychology, this time forever. *Measurement Interdisciplinary Research and Perspectives, 2*, 201–218. doi:10.1207/s15366359mea0204
- Molenaar, P. C. M., & Campbell, C. G. (2009). The new person-specific paradigm in psychology. *Current Directions in Psychological Science, 18*, 112–117. doi:10.1111/j.1467-8721.2009.01619.x
- Nakagawa, S., & Schielzeth, H. (2013). A general and simple method for obtaining r^2 from generalized linear mixed-effects models. *Methods in Ecology and Evolution, 4*, 133–142. doi:10.1111/j.2041-210x.2012.00261.x

- Neubauer, A. B., & Voss, A. (2016). The structure of need fulfillment. *European Journal of Psychological Assessment*. doi:10.1027/1015-5759/a000326
- Nezlek, J. B. (2017). A practical guide to understanding reliability in studies of within-person variability. *Journal of Research in Personality*, *69*, 149–155. doi:10.1016/j.jrp.2016.06.020
- Nezlek, J. B., Newman, D. B., & Thrash, T. M. (2017). A daily diary study of relationships between feelings of gratitude and well-being. *The Journal of Positive Psychology*, *12*, 323–332. doi:10.1080/17439760.2016.1198923
- Noftle, E. E., & Shaver, P. R. (2006). Attachment dimensions and the big five personality traits: Associations and comparative ability to predict relationship quality. *Journal of Research in Personality*, *40*, 179–208. doi:10.1177/0146167292185003
- R Core Team. (2016). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. Vienna, Austria. Retrieved from <https://www.R-project.org/>
- Reis, H. T., & Shaver, P. (1988). Intimacy as an interpersonal process. In *Handbook of personal relationships* (pp. 367–389).
- Revelle, W. (2016). *Psych: Procedures for psychological, psychometric, and personality research*. R package version 1.6.12. Northwestern University. Evanston, Illinois. Retrieved from <https://CRAN.R-project.org/package=psych>
- Riediger, M., & Freund, A. M. (2004). Interference and facilitation among personal goals: Differential associations with subjective well-being. *Personality and Social Psychology Bulletin*, *30*, 1511–1523. doi:10.1177/0146167204271184
- Rogge, R. D., Fincham, F. D., Crasta, D., & Maniaci, M. R. (2017). Positive and negative evaluation of relationships: Development and validation of the Positive – Negative Relationship Quality (PN-RQ) Scale. *Psychological Assessment*, *29*, 1028–1043. doi:10.1037/pas0000392
- Rosseel, Y. (2012). lavaan: An R package for structural equation. *Journal of Statistical Software*, *48*, 1–36.
- Russell, J. A., Weiss, A., & Mendelsohn, G. A. (1989). Affect Grid: A single-item scale of pleasure and arousal. *Journal of Personality and Social Psychology*, *57*, 493–502. doi:10.1037/0022-3514.57.3.493
- Sadikaj, G., Moskowitz, D. S., & Zuroff, D. C. (2016). Negative affective reaction to partners dominant behavior influences satisfaction with romantic relationship. *Journal of Social and Personal Relationships*, 1–23. doi:10.1177/0265407516677060

- Sadikaj, G., Moskowitz, D. S., & Zuroff, D. C. (2015). Felt security in daily interactions as a mediator of the effect of attachment on relationship satisfaction. *European Journal of Personality, 29*, 187–200. doi:10.1002/per.1988
- Schneider, M. (2001). Systems theory of motivational development. In N. J. Smelser & P. B. Baltes (Eds.), *International encyclopedia of the social & behavioral sciences* (pp. 10120–10125). Elsevier.
- Schönbrodt, F. D., & Gerstenberg, F. X. R. (2012). An IRT analysis of motive questionnaires: The Unified Motive Scales. *Journal of Research in Personality, 46*, 725–742. doi:10.1016/j.jrp.2012.08.010
- Schultheiss, O. C., & Wirth, M. M. (in press). Biopsychological aspects of motivation. In J. Heckhausen (Ed.), *Motivation and action* (3rd ed.). New York: Cambridge University Press.
- Schultheiss, O. C. (2001). An information processing account of implicit motive arousal. In M. L. Maehr & P. Pintrich (Eds.), *Advances in motivation and achievement* (12th ed., pp. 1–41). Greenwich, CT: JAI Press.
- Schultheiss, O. C. (2008). Implicit motives. In O. P. John, R. W. Robins, & L. A. Pervin (Eds.), *Handbook of personality: Theory and research* (3rd ed., pp. 603–633). New York: Guilford.
- Schultheiss, O. C., & Brunstein, J. C. (2010). *Implicit motives*. New York: Oxford University Press.
- Schultheiss, O. C., Jones, N. M., Davis, A. Q., & Kley, C. (2008). The role of implicit motivation in hot and cold goal pursuit: Effects on goal progress, goal rumination, and emotional well-being. *Journal of Research in Personality, 42*, 971–987. doi:10.1016/j.jrp.2007.12.009
- Shaver, P. R., & Brennan, K. A. (1992). Attachment styles and the "Big Five" personality traits: Their connections with each other and with romantic relationship outcomes. *Personality and Social Psychology Bulletin, 18*, 536–545.
- Sheldon, K. M., Ryan, R., & Reis, H. T. (1996). What makes for a good day? Competence and autonomy in the day and in the person. *Personality and Social Psychology Bulletin, 22*, 1270–1279. doi:10.1177/01461672962212007
- Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2012). A 21 word solution. *SPSP Dialogue, 1–4*. doi:10.2139/ssrn.2160588
- Stewart, A. J., & Rubin, Z. (1976). The power motive in the dating couple. *Journal of Personality and Social Psychology, 34*, 305–309. doi:10.1037/0022-3514.34.2.305

- Strack, F., & Deutsch, R. (2004). Reflective and impulsive determinants of social behavior. *Personality and Social Psychology Review*, *8*, 220–247.
- Vater, A., & Schröder-Abé, M. (2015). Explaining the link between personality and relationship satisfaction: Emotion regulation and interpersonal behaviour in conflict discussions. *European Journal of Personality*, *29*, 201–215. doi:10.1002/per.1993
- Weinberger, J., Cotler, T., & Fishman, D. (2010). The Duality of Affiliative Motivation. In O. C. Schultheiss & J. C. Brunstein (Eds.), *Implicit motives* (pp. 71–88).
- West, S. G., & Hepworth, J. T. (1991). Statistical issues in the study of temporal data: Daily experiences. *Journal of Personality*, *59*, 609–662. doi:10.1111/j.1467-6494.1991.tb00261.x
- Wickham, H. (2009). *Ggplot2: Elegant graphics for data analysis*. Springer-Verlag New York. Retrieved from <http://ggplot2.org>
- Wilson, R. E., Harris, K., & Vazire, S. (2015). Personality and friendship satisfaction in daily life: Do everyday social interactions account for individual differences in friendship satisfaction? *European Journal of Personality*, *29*, 173–186. doi:10.1002/per.1996
- Winter, D. G., John, O. P., Stewart, A. J., Klohnen, E. C., & Duncan, L. E. (1998). Traits and motives: Toward an integration of two traditions in personality research. *Psychological Review*, *105*, 230–250. doi:10.1037/0033-295X.105.2.230
- Wrzus, C., & Mehl, M. R. (2015). Lab and/or field? Measuring personality processes and their social consequences. *European Journal of Personality*, *29*, 250–271. doi:10.1002/per.1986
- Zhang, Z., Zyphur, M. J., & Preacher, K. J. (2009). Testing multilevel mediation using hierarchical linear models: Problems and solutions. *Organizational Research Methods*, *12*, 695–719. doi:10.1177/1094428108327450
- Zurbriggen, E. L. (2000). Social motives and cognitive power-sex associations: Predictors of aggressive sexual behavior. *Journal of Personality and Social Psychology*, *78*, 559–581. doi:10.1037//0022-3514.78.3.559
- Zurbriggen, E. L., & Sturman, T. S. (2002). Linking motives and emotions: A Test of McClelland's hypotheses. *Personality and Social Psychology Bulletin*, *28*, 521–535.
- Zygar, C., Hagemeyer, B., Pusch, S., & Schönbrodt, F. D. (2018). From motive dispositions to states to outcomes: Research data of an intensive experience sampling study on communal motivational dynamics in couples [translated title] (version 2.0.0) [data and documentation]. doi:10.5160/psychdata.zrce16dy99_v20000

Chapter 5

Paper 2: Motivated Behavior in Intimate Relationships: Comparing the Predictive Value of Motivational Variables



Motivated Behavior in Intimate Relationships: Comparing the Predictive Value of Motivational Variables

Caroline Zygar-Hoffmann
Ludwig-Maximilians-Universität München

Sebastian Pusch
Friedrich-Schiller-Universität Jena

Birk Hagemeyer
Friedrich-Schiller-Universität Jena

Felix D. Schönbrodt
Ludwig-Maximilians-Universität München

**Earlier version (before revision) of the published article in the
Social Psychological Bulletin. Readers are kindly advised to
switch to the final, published version.**

Zygar-Hoffmann, C., Pusch, S., Hagemeyer, B., & Schönbrodt, F. D. (2020).
Motivated behavior in intimate relationships: Comparing the predictive value
of motivational variables. *Social Psychological Bulletin*, 15(2), Article e2873.
<https://doi.org/10.32872/spb.2873>

Abstract

Motivational variables are considered as fundamental factors influencing the occurrence of behavior. The current study compared different types of motivational variables (implicit and explicit motive dispositions, motivation as states and as aggregated person-level variables) in their ability to predict communal and agentic behavior reports in intimate relationships. 510 individuals completed measures of dispositional communion and agency motives and participated in a dyadic experience sampling study with five assessments per day across four weeks. They reported on their momentary communal and agentic motivation and on behaviors of their partner and of themselves. All examined types of motivational variables predicted certain behavior reports on the between-person or within-person level, and had incremental effects beyond the other motivational variables in at least one motive domain. Directly replicating and conceptually extending prior research, the effects of motivational states and their aggregates were consistently found across behavioral outcomes, across self- and partner-reports and across the motive domains of communion and agency. Using the example of motivational states, the general value of assessing within-person variables for psychological phenomena in ESM-designs is discussed.

Keywords: implicit motives, explicit motives, motivation, behavior reports, intimate relationships, partner-report, experience sampling method

Motivational variables have been described as one major determinant of behavior. A fundamental notion about motive dispositions, for example, is their function, to orient, select and energize behavior towards certain incentives or away from certain disincentives (McClelland, 1987; McClelland, Koestner, & Weinberger, 1989), providing insights into the “why” of behavior (McClelland, 1987; Schultheiss, 2008). Researchers have investigated the ability of motivational variables to predict behavior in several motive domains, for implicit and explicit motives, and sometimes on the between-person as well as on the within-person level (e.g., Craig, Koestner, & Zuroff, 1994; McAdams & Constantian, 1983; Spangler, 1992).

Department of Psychology, Ludwig-Maximilians-Universität München. We embrace the values of openness and transparency in science (<http://www.researchtransparency.org/>). This research was funded by grants from the German Research Foundation to Felix Schönbrodt (SCHO 1334/5-1) and Birk Hagemeyer (HA 6884/2-1). The authors thank Paula Fehrmann, Nicole Horn, and Lara Lietge for the PACT scoring, Lukas Müller and Ludwig Zellner for the app development, and Franziska Eckerskorn for English proofreading.

Correspondence concerning this article should be addressed to Caroline Zyggar-Hoffmann, Leopoldstr. 13, 80802 München, Germany. E-mail: caroline.zygar@psy.lmu.de

Motives from different domains (e.g., communion or agency) orient people's attention to different incentives (e.g., closeness or distance), which results in selecting different behavior that is instrumental to goal attainment (e.g., spending time with others or alone), and energizing this behavior despite potential barriers. The distinction between implicit and explicit motives showed to be further relevant for the specific nature of the instrumental behavior (e.g., nonverbal or verbal, Hagemeyer, Dufner, & Denissen, 2016; McAdams et al., 1984a; McClelland et al., 1989; Schultheiss, 2001). However, a direct comparison of motive dispositions and motivational states for the prediction of behavior is missing so far, although motive theories emphasize the distinction (e.g., Atkinson, 1957; Bischof, 1975), and despite recent calls to focus on within-person processes in the prediction of behavior (Back & Vazire, 2015; Baumert et al., 2017).

For the domains of communion and agency, the current study compares three different classes of motivational variables in their ability to predict self- and partner-reported behavior in intimate relationships: 1) Implicit and explicit motive dispositions, 2) Motivational states as within-person variables, and 3) Aggregated motivational states as between-person variables. Data from an intensive longitudinal study employing the experience sampling method (ESM; Csikszentmihalyi & Larson, 1987) in a dyadic design is used to differentiate within-person from between-person effects of motivational states.

Motivational Variables

Implicit and Explicit Motive Dispositions

Motive dispositions are defined as relatively stable, inter-individual differences in preferences for certain classes of incentives (McClelland, 1987; Schultheiss, 2001). In a dual-systems model of motivation, McClelland et al. (1989) provide arguments and early evidence that implicit motives (assessed with projective measures) and explicit motives (assessed with self-report questionnaires) are two qualitatively different motive dispositions, which are at most weakly correlated (see Köllner & Schultheiss, 2014 for a more recent meta-analysis). Implicit motives are considered to be non-consciously represented, while explicit motives are considered part of one's accessible self-concept. A defining difference between these two constructs is the nature of the incentives they pertain to and the behavior they are assumed to evoke (McClelland et al., 1989): For implicit motives, the rewarding nature of incentives is supposed to be intrinsic to the behavior that attains the goal, being

represented by rather uncontrolled behavior or in rather unstructured situations. According to Schultheiss (2001, see also Schultheiss, 2008), they are assumed to mainly respond to non-verbal cues (e.g., facial expressions) and to impact bottom-up behavioral regulation (indicated by “non-declarative measures”, e.g., physiological responses, expressive and automatic behavior). For explicit motives, a more extrinsic drive is assumed that becomes effective in more structured situations, for example when making a choice while having one’s self-attributed goals or an external social demand in mind. They are assumed to mainly respond to verbal-symbolic cues (e.g., conversations) and to impact top-down reflective behavior (indicated by “declarative measures”, e.g., deliberate judgments). A detailed review of studies linking implicit and explicit motive dispositions to different kinds of behavior is provided by Stanton, Hall, and Schultheiss (2010).

Motivational States

Whereas motive dispositions are conceptualized as relatively stable inter-individual differences, motivational states are transient, intra-individually varying, and influenced by the situation. The term motivation describes an aroused motive (for an overview, see McClelland, 1987, p. 84), which is an “affectively charged state that energizes and directs action aimed at the attainment of a reward (or avoidance of a punishment)” (Schultheiss, Strasser, Rösch, Kordik, & Graham, 2012, p. 650). Motivation is expected to emerge from an interaction of motive dispositions and situational influences: First, moderately strong situational cues (e.g., the opportunity to socialize) may arouse a motivational state only for those with a strong correspondent motive (McClelland, 1987; Schultheiss et al., 2012). Second, situations may differ in their average level of need satisfaction, resulting in more experiences of unsatiated motivation for individuals with a strong motive disposition compared to those with a weak motive (see Bischof, 1975; Zygar, Hagemeyer, Push, & Schönbrodt, 2018 for a more detailed take on this). Within-person variation in explicit motivational states can be measured via self-report by repeatedly asking participants about their current motivation for certain goals.

Aggregated Motivational States

Beyond that, applying the ESM provides the possibility of aggregating motivational states for each individual across a certain period, with the mean of states representing the inter-individually varying average experience of motivational states as a new between-person variable.

Construct status. Stapleton, Yang, and Hancock (2016) differentiate two boundary cases for the construct status of aggregated constructs in the situation of persons nested in groups. We apply the logic here to repeated measurements nested in persons. The first type are (purely) *shared cluster constructs*. In this case, within-person measurements are intended to measure a between-person construct, and a reflective causal relationship is assumed (Lüdtke, Marsh, Robitzsch, & Trautwein, 2011) where variation in the between-person constructs causes variation in the within-person measurements. In absence of measurement error, ideally all within-person measurements reflect the same latent value of the person construct, with minimal variability at the within-person level. The second type are (purely) *configurational cluster constructs*, which represent summary indices of the responses on the within-person level. Typically, the mean is taken as an index of central tendency, but one could also use measures of dispersion or other indices. In contrast to shared cluster constructs, here a formative measurement model is assumed (Lüdtke et al., 2011). Hence, it is *not* expected that the responses at the within-person level are interchangeable or at least highly similar. Although the aggregate measure is derived from the within-person measurements, it is theoretically distinct from them (Bliese, 2000).

Several authors emphasized that between-person variables based on an aggregation of within-person measurements can also represent a mixture of shared/reflective and configurational/formative aspects (Bliese, 2000; Lüdtke et al., 2011; Stapleton et al., 2016). In such “fuzzy composition processes” (Bliese, 2000), the main difference between the within-person variable and the between-person aggregate is that other between-person predictors can explain the between-person variance part of the aggregate variable – representing contextual influences which are not captured by the within-person measure.

Such a mixture would fit the nature of aggregated motivational states: As we assume that motivational states are influenced by time-varying factors, such as situations, their aggregate cannot be a purely reflective construct. At the same time, the aggregated variable suggests a certain reflective status as it is theoretically expected and has empirically been shown that dispositional motive strength (i.e., of implicit and explicit motives) predicts the propensity to experience a motivational state (Zygar et al., 2018). This implies that aggregated states contain stable and systematic (between-person) variance, that is reflective of an individual’s motive strength. In fact, a variance decomposition of motivational states has shown that the variance between persons and couples is roughly of the same size as the variance within persons (Schönbrodt,

Zygar-Hoffmann, Nestler, Pusch, & Hagemeyer, 2019).

Hence, we conceptualize aggregated motivational states as simultaneously shared and configurational cluster constructs (Stapleton et al., 2016). These are, on the one hand, summary indices of the experienced states and therefore formative summaries of the motivational climate during the assessment period. On the other hand, they contain systematic and meaningful variance at the between-person level.

Delineation from motive dispositions. What differentiates aggregates of motivational states from explicit and implicit motives? A theoretically assumed feature of motive dispositions is a certain stability over time (see Denzinger & Brandstätter, 2018 for an overview on implicit motives). Although changes are not impossible, shifts would be assumed to occur moderately: Implicit motive dispositions are proposed to be a result of learning experiences (see Schultheiss, 2008). Thus, a constant long-term experience of a situation that does not match an individual’s implicit motive may lead to a slow situational adaptation of the disposition over time (a process that has been described as “acclimatization”, see Bischof, 2008; Gubler & Bischof, 1991).

The stability of average motivational states, in contrast, can vary: As a motivational state is assumed to emerge from an interaction of (relatively stable) motive dispositions and (potentially unstable) situational influences, the stability of aggregated motivational states depends on the permanence of the situation over the measurement period. For example, if a relationship provides rather constant and steady (dis-)incentives and (dis-)satisfaction, one would assume that individuals in this relationship have a rather stable motivational climate during the considered time. Similarly, average experiences of motivation should change profoundly within a person when qualitatively different measurement periods are compared (e.g., a longer absence of the partner compared to a shared holiday trip).

In our view, motivational states, aggregated motivation, and motive dispositions could therefore be best placed on a continuum representing the permanence of situational features needed to evoke change in the measure: Motivational states are based on direct, rather immediate situational contingencies; aggregated motivational states represent typically experienced contingencies during a specific time-frame; and repeated time-frames of experienced contingencies are necessary for a slow, and thus time-delayed adaptation of implicit and probably also explicit motive dispositions.

The Influence of Motive Dispositions and Motivational States on Behavior

Many classical theoretical accounts on motivation are formulated on an intra-individual level, emphasizing the waxing and waning of motivation and behavior within persons over time and situations (see, e.g., Atkinson & Birch, 1970). In contrast, many empirical studies linking motivational variables to behavior do this on a between-person level. Although between-person variables like motive dispositions are indicative of which individuals are generally more prone to show certain types of behavior, this result cannot be seamlessly transferred to the within-person level (Molenaar, 2004; Molenaar & Campbell, 2009). Within-person variables capturing varying motivational states of each person are necessary to predict in which instances a single individual is more prone to show a certain behavior, and thereby to align a study's design to the theoretical models on motivation.

Furthermore, many motivational theories (implicitly or explicitly) imply that the influence of motive dispositions on behavior is fully mediated via state motivation (McClelland, 1987; Rheinberg, 2002, see also Zygar et al., 2018). This strict interpretation implies that a strong latent motive disposition does not per se trigger more behavior than a weak motive disposition – only with higher levels of corresponding current motivation more instrumental behavior should emerge.

In contrast to this perspective, a between-person analysis could still find incremental effects of motive dispositions beyond the state effects of motivation. Repeated enactment of instrumental behavior that gets reinforced by a satiation of needs, for example, can lead to habits. Consequently, if strong motive dispositions are correlated with corresponding behavioral schemata and habits, a between-person analysis could reveal that persons with strong motives habitually show more instrumental behavior, even when no current motivation is present. This would become evident in effects of globally assessed implicit and/or explicit motive dispositions predicting behavior beyond the between-person effect of aggregated motivational states capturing the motivational climate during the assessment period.

Analyzing the predictive effects of between- and within-person motivational variables requires, in a first step, determining to what extent behavior actually varies between and within persons. Dispositions as time-invariant predictors can only explain the between-person variance of the outcome (Hoffman & Stawski, 2009). In contrast, motivational states (calculated as momen-

tary deviations from a person's typical motivation) are time-varying predictors that can only explain the pure within-person variance. To summarize, when between- and within-person processes are disentangled, it is theoretically relevant to investigate (a) to what extent behavior varies between and within persons, (b) whether motivational states (as within-person variables) can predict the within-person portion of behavioral variance, and (c) whether motive dispositions and/or aggregated motivation (as between-person variables) can predict the between-portion of behavioral variance.

What Kind of Behavior? The Domains of Communion and Agency Motives

Motives can be classified based on the (dis-)incentives they correspond to. One possible, rather broad distinction is the orientation towards communal versus agentic interpersonal incentives. This classification was introduced by Bakan (1966), who postulated communion and agency as two fundamental concepts for the study of human behavior. While the term communion reflects the participation of the individual in a larger social context, by forming connections to and cooperating with others, agency entails a focus on the individual, on separations and mastery (Bakan, 1966). Different theoretical frameworks emphasize different aspects of communion and agency (Abele, Cuddy, Judd, & Yzerbyt, 2008). Based on the definitions of partner-related needs by Hagemeyer and Neyer (2012), we will discuss which concrete classes of goal states communion and agency motives aim for in couple relationships.

Closeness as Core Incentive of the Communion Motive in Couple Relationships

According to Hagemeyer and Neyer (2012), a core incentive for the partner-related communion motive is the experience of closeness and a sense of unity with the partner. This may be achieved by warm and kind interactions, for example by actively engaging with each other or disclosing thoughts and feelings to each other. On an explicit level, this is captured by the partner-specific desire for closeness (Hagemeyer, Neyer, Neberich, & Asendorpf, 2013), and by the more global intimacy motive, which is characterized by the desire for experiencing interpersonal closeness in general (Schönbrodt & Gerstenberg, 2012).

Regarding the dispositional level, there are already many studies linking communal motives to diverse types of behavior in relationships (e.g., Craig et

al., 1994; Dufner, Arslan, Hagemeyer, Schönbrodt, & Denissen, 2015; Hagemeyer et al., 2016; Lansing & Heyns, 1959; Mason & Blankenship, 1987; McAdams & Constantian, 1983; McAdams et al., 1984a, 1984b; Schüler, Job, Fröhlich, & Brandstätter, 2008; Zurbriggen, 2000). For example, Craig et al. (1994) showed that the explicit intimacy motive predicts the number of social interactions in daily life, while the implicit intimacy motive predicts the number of dyadic interactions in particular. On a state level, we already showed that communal motivational states are predictive of subsequent specific communal behavioral acts at the within-person level, with aggregated states having an additional effect on the between-person level (Zygar et al., 2018). This study, however, neither contrasted this to dispositional measures (such as implicit or explicit motives) nor examined agency as another relevant motive domain in relationships.

Two Implementation Styles of the Agency Motive

Hagemeyer and Neyer (2012) define the implicit agency motive in the context of romantic relationships as “a recurrent concern for experiences confirming the self as an independent and capable individual.” (p.116). Two facets are considered simultaneously in this definition: Independence and power. For the research questions at hand it is important to distinguish these aspects. On the one hand, both independence and power share the goals of placing the individual first and forming separations from others, constituting the reason as to why they are subsumed under the agency motive. On the other hand, independence can be achieved by individual activities or by creating distance, whereas the realization of power requires some form of contact with the partner. Thus, whereas both facets of the agency motive aim for psychological distance from the partner, realizing power often requires physical proximity while independence is more often experienced when being by oneself.

Individuals with a high power motive experience reward from exercising dominant behavior, for example by influencing, impacting and controlling other people’s behavior, emotions or thoughts (McClelland, 1987; Winter, 1973). They also find it rewarding to enhance their prestige, for example by being in the focus of others’ attention. While the power motive is often regarded as negative and manipulative (“personalized power”), it can also become evident in pro-social actions, such as having an impact by supporting others (e.g., advising) or by cheering them up (i.e., influencing feelings in a positive way, “socialized power”; McClelland, 1970, 1975). Still this kind of pro-social agentic behavior is motivated by the need for the mastery of the

situation and the experiences of one's power, and not by the need to form a connection (as would be the case for the communion motive). The power motive is thus characterized by the need to distinguish the individual from other persons like the partner, realized by the ability to influence and to impress (e.g., enforcing decisions or seeking admiration in the relationship). Although other persons are essential to exert dominance and to receive prestige, the focus lies on the individual creating a psychological separation (Hagemeyer & Neyer, 2012; McClelland, 1987; Winter, 1973).

Concerning the need for independence, the individual is placed first by means of creating autonomy, freedom and individuality. This translates into the goals of pursuing one's own interests independently, being a capable individual without the help of others, and spending time by oneself. Therefore, the fulfillment of this motive will regularly be accomplished in a couple relationship by creating a physical distance from the partner (Hagemeyer & Neyer, 2012).

So far, research has mainly focused on the relations between *dispositional* agency motives and relationship behavior, for example the implicit power motive predicting agentic leadership and persuasive behaviors in friendships (McAdams et al., 1984a; see also Ackerman & Corretti, 2015; Mason & Blankenship, 1987; Zurbriggen, 2000, 2011). Further, dispositional implicit and explicit independence motives were associated with couples' living arrangements, that is, whether men or women at different ages were coresident with their partner or living apart (Hagemeyer, Schönbrodt, Neyer, Neberich, & Asendorpf, 2015).

The Current Study

In this study, we investigate the ability of motivational variables on three conceptual levels (motive dispositions, motivational states, and aggregated motivational states) to predict self-reported and partner-reported instrumental behavior in couple relationships for different motive domains (communion and agency). Behavioral acts in uncontrolled everyday situations are a multi-determined phenomenon, which sets an upper limit on the variance that can be explained by motivational factors. Whether motives and motivation lead to actual instrumental behavior depends on additional factors, both in the situation (e.g., opportunities and barriers), and in the person (e.g., habits, necessary skills, restraining self-control, or conflicting motivations). It is theoretically expected that these additional factors contribute to the prediction of behavior, both as main effects and in interaction with motivation (McClelland,

1987; Schultheiss, Kordik, Kullmann, Rawolle, & Rösch, 2009). Nonetheless, various previous studies operated under the assumption that, averaged across situations, higher motives and/or higher motivation lead as a main effect to more instrumental behavior (e.g., McAdams et al., 1984a; Schultheiss, Dargel, & Rohde, 2003; Zurbriggen, 2011; Zygar et al., 2018). In the current analyses, we also take this stance and focus on the marginal main effects of motivational variables on behavior.

This study aims to replicate existing findings, and to extend the evidence in multiple ways. It specifies four goals concerning the prediction of behavior reports: (1) quantify the unique contributions of motive dispositions as classical trait variables, of within-person motivational states, and of aggregated states as emergent between-person variables, (2) perform a high-powered direct and conceptual replication of previously published results showing that communal motivational states as well as aggregated states predict specific self-reported behavioral acts in the communion domain (Zygar et al., 2018), (3) test the generalizability of these findings by extending them to the domain of agentic motivation, as well as to motive dispositions, and (4) extend these findings beyond self-reports of behavior by including partner-reported behavior.

Our preregistration is available at <https://osf.io/af4yb/>.¹ We had only hypotheses for goals (2) and (3): We preregistered that all motivational variables are positively related to self-reported behavior that promotes incentive attainment, and negatively related to self-reported behavior that hinders it.

For a broad conceptual replication of the finding by Zygar et al. (2018), we used several operationalizations both for predictors and outcomes. These differed in their abstractness and mode of assessment, a summary is presented in Table 1. The measurement of behaviors included three different operationalizations: (i) reports about the occurrence of specific behavioral acts, (ii) an overall evaluation of behavior quality, and (iii) an estimate on how much time individuals spend alone or for their own interests. On the one hand, these types of behavior can be considered as declarative measures, because individuals are explicitly prompted to reflect on their behaviors and report them, thereby providing the possibility to align the answers with their self-concept. On these grounds, explicit motives should be more suitable to predict these answers. On the other hand, the ESM aims to reduce such biases of identity-related beliefs

¹Other research questions were also preregistered, which are not covered in the current manuscript. The preregistration mentions several possible exploratory analyses; in this manuscript we focus on the confirmatory preregistered hypotheses, which are referred to as research goals 4A, 4B, 5A and 5B in the preregistration.

in general, trying to rely on experiential memory by asking about concrete incidents and time-frames (Bolger, Davis, & Rafaeli, 2003; see also Robinson & Clore, 2002). The dyadic nature of the study could also prompt to report as objectively as possible, as own reports can be contrasted with the reports of the partner (e.g., the reported occurrence of a conflict should normally not diverge between partner's answers). Further, the actual behavior that took place in the relationship when individuals report that they behaved in a certain way is most probably a combination of both declarative and non-declarative processes (e.g., affection is shown verbally and non-verbally). Therefore, implicit motives could as well predict these behaviors.

Table 1
Operationalizations of Predictors and Outcomes

Variables	Characteristic			Motive Domain		
	Abstractness	Assessment	Communion	Agency	Independence	Power
Predictors						
Explicit desires	relationship-specific	self-report ¹	×		×	
Explicit motives	global	self-report	×			×
Implicit motives	relationship-specific	indirect	×	×	×	×
Motivational states	relationship-specific	self-report	×	×	×	×
Mean motivation	relationship-specific	self-report	×	×	×	×
Outcomes						
Specific acts: Behavior Index	relationship-specific	self- and partner-report	×	×		×
Overall quality: IPC	relationship-specific	self- and partner-report	×	×		×
Time: Spent alone	global	self-report	×			
Time: Spent for own interests	global	self-report			×	

Note. IPC = interpersonal circumplex. ¹ Explicit desires were also assessed via partner-reports, but only the self-reports were preregistered, and to avoid confounds with perception biases, only those are presented in this paper. However, the corresponding effects for partner-reported explicit desires are reported in the Supplemental Materials.

Method

We report how we determined our sample size, all data exclusions, all manipulations, and all measures in the study (Simmons, Nelson, & Simonsohn, 2012). As this study assessed dyadic data which cannot be fully anonymized without losing the assignment of individuals to dyads, we published our data as a scientific use file accessibly only to academic users (Zygar-Hoffmann, Hagemeyer, Pusch, & Schönbrodt, 2020). All analysis scripts² and Supplemental Materials can be found in the associated OSF repository (<https://osf.io/urwq7/>).

Procedure

The study was advertised for heterosexual couples through social networks, flyers and by couple counselors. To participate, both partners were required to own an Android or iOS smartphone, which they have at hand for regular use. Upon registration, couples chose the time span in which the daily surveys should be sent (ranging from a start at 07:00 to 10:00 am, and an end from 9:00 pm to 11:00 pm), and blocked up to two hours per day as survey-free. Subsequently, each participant completed a preliminary questionnaire by themselves on their personal computer, and received the ESM application “Tellmi” developed at LMU Munich. Upon login with a personalized code, the survey questions were introduced in a video. Participants were instructed not to talk about their answers with their partner during the ESM study, even though the survey timings were identical for both partners. The study started on the next Monday after login, with five daily surveys over four weeks. The first four surveys of each day were identical, and were scheduled semi-randomly across the chosen time-span. Participants were notified by their smartphone about an active survey, and had 45 minutes to complete it before it timed out. The last survey of each day had a different subset of questions and was designed as a daily diary: Individuals were instructed to complete it before going to bed, therefore it had a timeout of five hours. Answering the surveys took a median time of 2.70 minutes (interquartile range = 2.17 minutes). Directly after the four weeks and again a year later individuals were invited to complete

²For our analyses we used *R* (version 3.5.3, R Core Team, 2019) with mainly the packages *dplyr* (Wickham, François, Henry, & Müller, 2018) for data handling, *lme4* (Bates, Mächler, Bolker, & Walker, 2015) and *lmerTest* (Kuznetsova, Brockhoff, & Christensen, 2017) for multilevel modeling, *stats* (R Core Team, 2019) for the Benjamini-Hochberg correction, *MBESS* (Kelley, 2018) for computing McDonald’s omega, and *MuMIn* (Barton, 2018) to determine the explained variance of the models. We wrote this reproducible manuscript with the package *papaja* (Aust & Barth, 2018).

follow-up questionnaires. Participants received compensation or course credit based on their compliance in the ESM part of the study and could opt in for a feedback report on their answers.

Sample

The number of couples to be recruited was constrained by the money available for compensation (up to 190€ per couple). Out of 576 participants (from 293 couples) who completed the preliminary questionnaire, ten individuals were not eligible for the ESM part, because their partners did not answer the preliminary questionnaire; six couples decided not to start with the ESM part, and 18 couples as well as eight individuals provided not enough ESM data (we preregistered to include only data from individuals who answered at least one third of all surveys). As a result, a total of 510 participants from 259 couples (256 women and 254 men) provided data for the ESM part of the study that we could use (age in years: $M = 31.40$, $SD = 9.54$, range = 18 - 68; relationship duration in years: $M = 6.43$, $SD = 6.43$, range = 0.17 - 33.17). Most participants were childless (68 %), not married (67 %), not students (71 %), but had a German Abitur (64 %). This sample had a mean compliance of 88% with a standard deviation of 12%. We further excluded some data on the survey level leading to a total of 60942 (at least partly) answered surveys.³

Measures

The complete codebook including all variables of the study can be found at <https://osf.io/psqx8/>.

Preliminary Questionnaire: Motive Dispositions. The dispositional measures were completed in a preliminary questionnaire which was set up with the survey software *formr* (Arslan, Walther, & Tata, 2020; Arslan & Tata, 2017).

Explicit social desires (relationship-specific).

The partner-specific explicit social desire for closeness and desire for being alone were each assessed as the average of eight items of the ABC scale of social desires (Hagemeyer et al., 2013). Statements like “In the presence of my partner, I feel relaxed.” (closeness) or “I like to be completely alone.” (being

³First, a time zone transition of a couple caused their surveys to be wrongly activated during nighttime, thus all answers on these surveys were excluded ($n = 26$ surveys = 0.04%). Second, we preregistered to exclude answers that were discussed with the partner and where we expected bias due to the discussion (in $n = 171$ surveys = 0.24%), and surveys that were answered in less than 60 seconds ($n = 1855$ surveys = 2.58%).

alone) were rated on subjective frequency scales ranging from 1 = *never* to 7 = *always*.

Explicit motives (global).

Global (i.e., not partner-specific) explicit motives of intimacy and power were assessed as the average of the six-item versions of the Unified-Motives-Scales (UMS-6; Schönbrodt & Gerstenberg, 2012). Participants rated statements and goals such as “Not being separated from the people I really care about” (intimacy) or “Opportunities to influence others” (power) on Likert scales (e.g., 0 = *not important to me* to 5 = *extremely important to me*).

Partner-related implicit motives.

The Partner-Related Agency and Communion Test (Hagemeyer & Neyer, 2012) was used to assess implicit partner-related needs for Communion (pnCommunion) and Agency (pnAgency, including the subscales for pnPower and pnIndependence).⁴ In this projective measure, participants are prompted to write a short relationship story about eight ambiguous pictures based on three question prompts. For each story, a random pair of two out of five trained coders scored the stories for motive-related themes, and codings were averaged (see Table 2 for inter-coder-reliability). Ambiguous cases were discussed and resolved together with all coders. One participant answered less than six pictures, therefore this person’s PACT scores were coded as missing. After averaging the codings across all pictures to obtain raw motive scores, word count of the stories ($M = 334$, $SD = 149$) was partialled out with a robust regression approach, to account for a confounding with verbal fluency (correlations of word count and motive score were between $-.04$ and $.41$, see Hagemeyer & Neyer, 2012 for recommendations of this procedure).

ESM: Motivational States. For communion, independence and power motivation, we preregistered to compute scales of items that were assessed at each of the first four surveys of each day, when the respective event-level reliability exceeded $.40$. This was the case for all three motivational domains (see Schönbrodt et al., 2019) for an extensive reliability analysis of these items on different analysis levels). An agency motivation scale was computed as average across the independence and power items.⁵

⁴The test allows to differentiate pnCommunion and pnAgency by an approach and an avoidance component, representing the focus on seeking out incentives compared to avoiding disincentives. The Supplemental Materials presents the results when examining the effects of these components separately.

⁵We consider agency motivation a more formative construct, therefore we preregistered to compute the scale of power and independence items irrespective of the internal consistency, which was below $.40$, see Schönbrodt et al. (2019).

Depending on whether the partner was present or not, the instruction for the items changed: If the partner was present, individuals were asked “right now, do you wish:” followed by a goal consistent with the motive domain. If the partner was not present, individuals were asked to imagine they had free time to spend with their partner and to indicate what their motivation would be in this situation. This instruction aimed to reduce situational effects undermining the report of a motivation, such as when certain desires seemed impractical or difficult to realize. Participants were asked to report on their momentary desires “to share experiences, thoughts or feelings with [their] partner” and “to receive emotional affection from [their] partner” (both communion) as well as “to act and decide independent of [their] partner” and to solitarily pursue [their] own interests” (both independence). Further assessed desires were “to influence the feelings or behavior of [their] partner in any way”, “that there is an exchange with [their] partner, which is about [them], where [they] are the center of attention”, and “that [their] partner fits in with [their] wishes” (all power). Examples of correspondent behavior and an optional tooltip providing more descriptions were available for each desire (see codebook). These motivational states were answered on the same Likert scale, ranging from -1 = *no, that would bother me right now* over 0 = *no, I don't need this right now*, 1 = *yes, but only weakly*, 2 = *yes, moderately*, 3 = *yes, strongly*, to 4 = *yes, very strongly*.

We further had two additional items pertaining to communal motivation more generally, which we preregistered not to include in the communion scale, as they have a slightly different focus: First, the question “How emotionally close would you want to be to your partner at the moment?” (termed “closeness motivation”) was answered on a discrete seven-point slider, with each position depicting one picture of the Inclusion of Other in the Self Scale (Aron, Aron, & Smollan, 1992). The first position was labeled *distance* and the last position *maximal closeness*. Second, individuals were prompted to imagine how they would like to spend the next two hours, if they could use the time as they wished (termed “time-spending motivation”), with answers on a continuous slider from 0 = *Entirely without your partner (as me-time)* to 10 = *Entirely with your partner (as shared-time)*.⁶ We preregistered our hypotheses for the communion scale and for these two single items separately, and to tackle the problem of multiple testing by considering all analyses in a control of the false-discovery-rate.⁷

⁶This item was the only item which was also assessed in the evening survey, with the variation that individuals were asked how they would like to spend their time *the next day* instead of in the next two hours.

⁷As reported in Schönbrodt et al. (2019), computing a scale with all four communal

ESM: Behavior Reports. Behavior reports were assessed at each of the five daily surveys. For the first survey each day, individuals were instructed to refer to what happened since the morning, for all other surveys they should refer to the time since they answered the last survey.

Occurrence of specific behavioral acts: Communion and Agency Behavior Index.

A list of twenty specific communal and agentic behaviors was presented to the participants in the form of a checklist. Participants were instructed to check a box if the associated behavior had occurred. When the described behavior was reciprocal (e.g., “Intimate or cordial activity, conversation or discussion”) rather than unidirectional (e.g., “You elicited strong positive or negative feelings in your partner”), participants were further asked who initiated the behavior. The list of behaviors was rated by the authors a priori as instrumental or obstructive for the attainment of communal and agentic relationship goals and according weights were preregistered (-1 = very obstructive, -0.5 = a little obstructive, 0 = irrelevant, 0.5 = a little instrumental, 1 = very instrumental; see codebook for the full list of behaviors with their corresponding weights). To calculate communion and agency behavioral indices, the checked behaviors at each survey were multiplied by their assigned weights and the resulting scores were summed up, representing the amount of obstructive or instrumental behavior that was shown for each motive domain in each time span (ranging from -2.5 to 6.5 for the communion index and from -2.0 to 8.0 for the agency index).

Overall evaluation of behavior quality: IPC kindness and IPC dominance.

A grid representing the broad interpersonal dimensions of communion (on the x-axis) and agency (on the y-axis) was presented twice per assessment: Participants should indicate on these interpersonal circumplex grids (IPC; see Horowitz et al., 2006; Wiggins, 1979) how they behaved towards their partner and how their partner behaved towards them. They were instructed to point to the position on the grid that represented both the amount of kindness (from 0 = *rejecting* to 1 = *friendly*) and the amount of dominance (from 0 = *unobtrusive* to 1 = *dominant*).⁸ Examples of combinations of both dimensions

motivational items is a viable alternative. We refer readers who are interested in the results for such a scale to the Supplemental Materials.

⁸The answers on the dimensions were associated ($\beta_{kindness\sim dominance} = 0.18$ and $\beta_{dominance\sim kindness} = 0.16$ for own behavior; $\beta_{kindness\sim dominance} = 0.14$ and $\beta_{dominance\sim kindness} = 0.13$ for partner behavior). Instead of a correlation, we report regression coefficients from a multilevel model, illustrating both directions of regressing the

were presented in the edges of the grid (e.g., *protecting* for a highly kind and dominant behavior).

Quality of time: Time spent alone and for own interests.

Participants answered two questions on how they spent their time: “How much time did you spend alone (without mentionable interactions with your partner or others)?”, and “How much time did you spend to pursue your own interests, e.g., hobbies?”. Both items were answered on a continuous slider ranging from 0 = *no time at all*, over 5.5 = *half of the time* to 10 = *all the time*.

Covariates: Time Spent With the Partner and Amount of Duties. We assessed potential covariates at the within-person level that impose rather strong restrictions or upper limits on the ability to show any partner-related behavior, which we control for in our models. At each ESM survey, participants were asked about the time they had spent with their partner since the last survey or since the morning: “How much time did you actually actively spend together with your partner (technically mediated as well)?”. Responses were given on a continuous slider ranging from 0 = *no time at all*, over 5.5 = *half of the time* to 10 = *all the time*. Additionally, at the first four ESM surveys, participants reported the amount of workload they had upcoming: The question “How many tasks are on your to-do-list for the next two hours (occupationally as well as privately)?” was answered on a continuous slider from 0 = *no tasks* to 10 = *many tasks*.

Data analysis

Our data is structured as surveys nested in individuals nested in couples. To tackle this three-level structure with only two data points on the upper level, we estimate double-intercept models (Bolger & Laurenceau, 2013) with two levels for all analyses: These models specify the survey answers on level 1 and couples on level 2, but include separate intercepts for men and women as fixed and random effects to account for individual differences. We *z*-standardized all continuous variables on the grand-mean and grand-standard deviations across the whole sample.

Our behavior-related outcome variables are on level 1 in all analyses, and we preregistered to include the following two covariates for such models: A variable representing the survey index (0 = first survey) to control for

indizes on each other, because the multilevel structure would not have been considered in a simple correlation across all surveys.

potential effects on the outcomes over the course of the study (see Bolger & Laurenceau, 2013), and a dummy variable indicating whether the survey was on a weekday (=0) or during the weekend (=1). When predicting the outcomes by motivational states, we preregistered to include the random effect of the motivational state if the model converges (Barr, Levy, Scheepers, & Tily, 2013), and to add the following covariates on level 1: a) If the outcome is one of the behavior indices, or a dimension of the interpersonal circumplex, then we include the time spent with the partner as covariate to account for the amount of time individuals could behave towards their partner; b) If the outcome is the time individuals spent alone or for own interests, then we add the reported momentary amount of duties as covariate, to account for situational constraints influencing how individuals spend their time. Finally, for all analyses with one dimension of the interpersonal circumplex as outcome, regardless of the predictor being on level 1 or level 2, we added the other dimension as covariate to account for a potential confounding of the answers on the grid.

When mixed preregistered and exploratory analyses are presented together (see Tables 5 and 6), the results of preregistered analyses are printed in black and one-sided p -values are reported. Results of exploratory analyses are printed in gray, with two-sided and uncorrected p -values.⁹ As some of our preregistered hypotheses refer to the same underlying idea, but with varying operationalizations, we preregistered exactly how we aim to control the false-discovery rate (FDR) at $\alpha = 0.05$ with the Benjamini-Hochberg correction (Benjamini & Hochberg, 1995).¹⁰ In tables shown in this paper, we printed preregistered analyses in bold when they were significant after controlling the FDR.

⁹We preregistered to consider exploratory effects noteworthy when the p -value is smaller than $\alpha = .01$. However, the exploratory analyses reported in this paper are direct extensions of our preregistered analyses to partner-reports, to similar operationalizations or to models with more covariates. Compared to pure exploration of the data for noteworthy patterns, we deem these analyses as theoretically more founded. We therefore discuss also results with a two-sided p -value smaller than .05. However, as we report exact p -values, readers can regard exploratory results with a p -value between .01 and .05 as less trustworthy by themselves.

¹⁰When multiple models were calculated with one type of motivational variable (e.g., implicit motives, explicit motives, state motivation, or aggregated motivation) pertaining to one motive domain (e.g., communion or agency) for the same or a similar outcome (the behavior index and the IPC measure were regarded as similar), then these were treated as one subset of models for which a control of the FDR was indicated. For example, for implicit motives in the domain of agency, four models were calculated in total: Two models with pnAgency (one model predicting the agency behavior index and one model predicting the IPC measure), and similarly two models with pnPower. The FDR was controlled for the subset of these four models. An even stricter error control could be applied, but as we considered the different motivational variables and motive domains as having separate theoretical foundations pertaining to different processes and effects, we did not control the FDR across all analyses.

Results

Descriptives

In Table 2, means and standard deviations of motivational variables on the between-person level are presented separately for women and men. Regressing these variables on gender in a multilevel model showed that women had significantly higher means than men on the explicit intimacy motive, $t(255) = 5.69$, $p < .001$, on raw pnCommunion motive scores, $t(258) = 3.46$, $p = .001$, on mean communion motivation, $t(253) = 5.14$, $p < .001$, and on mean power motivation, $t(254) = 2.48$, $p = .014$. However, on average, women had a significantly lower explicit power motive than men, $t(255) = -4.95$, $p < .001$. All other comparisons were not significant ($ps > .094$).

The correlations in Table 3 show convergence of the motivational variables on the between-person level. Corresponding implicit and explicit motives correlate with a maximum of $r = .28$, observed for the explicit desire for closeness and pnCommunion in women. The maximum correlation between implicit motives and aggregated states was of similar size, with $r = .29$ for mean closeness motivation and pnCommunion, again found in women. Overall, explicit motives and aggregated motivational states correlated more strongly, with a maximum observed $r = .50$ for the explicit desire for closeness and aggregated closeness motivation in women. Intra-dyad correlations were significant for both explicit communion motives, but not for the other motive dispositions. Regarding aggregated motivational states, intradyad-correlations were substantial for all motive domains (with rs ranging from .25-.53).

In Table 4, the Intra-Class-Correlations (ICCs) of state measures show that there is considerable variance of the predictor and outcome variables on both analysis levels. Regarding the outcomes, the variance on the between-person level was generally lower than the variance on the within-person level: Time spent alone had the lowest between-person variance (15.60% for women, 17.01% for men), and kindness behavior had the highest between-person variance (36.42% for women, 38.60% for men). It has to be noted that the variance on the within-person level is a mix of systematic and error variance. As our outcome measures are formative indices or single item measures, their reliability is unknown. However, even if we assume low reliabilities there is still a substantial amount of true within-person variance. It is therefore meaningful to proceed with comparing within- and between-level predictors to explain the total variance observed in the behavioral outcomes.

Table 2
Descriptive Statistics of Between-Person Motivational Variables

Variables	Rel	Range	Women		Men	
			<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Explicit motives						
Explicit desire: Closeness	.85	1.4 to 7	6.03	0.85	6.03	0.66
Explicit desire: Being alone	.89	1 to 6.9	4.17	1.02	4.23	0.94
Explicit intimacy motive	.70	2.2 to 6	4.92	0.60	4.63	0.70
Explicit power motive	.87	1 to 6	3.16	0.95	3.57	1.03
Implicit motives						
pnCommunion	.96	0 to 10	5.55	2.02	4.95	2.04
pnAgency	.93	0 to 10.5	5.07	2.00	4.79	1.94
pnIndependence	.93	0 to 6.5	1.69	1.23	1.56	1.20
pnPower	.90	0 to 5.5	1.12	1.07	1.24	1.19
Aggregated motivational states						
Mean communion motivation	.97	0 to 4	2.02	0.8	1.75	0.76
Mean closeness motivation	.97	1.3 to 7	4.85	1.28	4.83	1.29
Mean time-spending motivation	.96	1.1 to 10	7.08	1.9	6.92	1.88
Mean agency motivation	.98	-0.6 to 2.6	1.00	0.57	0.94	0.53
Mean independence motivation	.97	-0.7 to 3.2	1.09	0.68	1.13	0.66
Mean power motivation	.98	-0.6 to 2.8	0.94	0.69	0.82	0.59

Note. N = 509-510 individuals in 259 couples (255-256 women, 254 men). Rel = McDonald's omega total for explicit measures, intra-class-correlation ICC(1,2) of content codings for implicit measures, between-person reliability estimate based on Schönbrodt et al. (2019) for aggregated state measures. pn = partner-related need. *M* and *SD* of implicit motives refer to raw motive scores (number of motive categories).

Table 3
Correlations of Between-Person Motivational Variables

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1.Explicit desire: Closeness	.28	-.19	.48	-.05	.20	-.18	-.12	-.11	.32	.47	.38	-.08	-.16	.00
2.Explicit desire: Being alone	-.32	.12	-.36	-.12	.05	.00	.21	-.23	-.06	-.13	-.27	.06	.20	-.06
3.Explicit intimacy motive	.38	-.40	.19	.12	.24	-.13	-.14	-.04	.23	.39	.39	-.04	-.14	.04
4.Expl power motive	-.08	.02	-.04	.11	-.02	.01	-.04	.11	-.03	-.00	.05	.15	.13	.12
5.pnCommunion	.28	-.06	.17	.00	.06	-.40	-.24	-.23	.13	.19	.08	-.01	-.10	.05
6.pnAgency	-.24	-.02	-.06	.02	-.49	.07	.53	.42	-.00	-.18	-.06	.12	.16	.05
7.pnIndependence	-.07	.14	-.08	.04	-.27	.57	-.05	-.14	-.08	-.20	-.14	.03	.14	-.05
8.pnPower	-.11	-.08	-.03	.06	-.26	.39	-.07	-.01	-.03	-.04	.06	.01	-.02	.03
9.Mean communion motivation	.33	-.18	.31	-.02	.19	-.00	-.09	.03	.44	.58	.54	.36	.10	.48
10.Mean closeness motivation	.50	-.18	.36	-.06	.29	-.21	-.14	-.15	.63	.50	.67	.01	-.17	.14
11.Mean time-spending motivation	.47	-.40	.40	-.04	.25	-.18	-.20	-.01	.56	.68	.53	-.08	-.34	.12
12.Mean agency motivation	-.14	.10	-.02	.16	-.03	.16	.05	.09	.39	.03	-.11	.29	.81	.90
13.Mean independence motivation	-.26	.31	-.22	.14	-.11	.15	.10	.08	-.00	-.26	-.46	.74	.30	.47
14.Mean power motivation	-.01	-.07	.12	.13	.03	.12	-.00	.07	.54	.21	.15	.90	.36	.25

Note. N = 509-510 individuals in 259 couples (255-256 women, 254 men). pn = partner-related need (implicit motive). Correlations below the diagonal refer to associations between women, correlations above the diagonal refer to associations between men, correlations on the diagonal refer to dyadic associations. Correlations of implicit motives were calculated with motive scores corrected for word count. Correlations with $p < .05$ are printed in bold italic, those with $p < .01$ are printed in bold.

Table 4

Descriptive Statistics and Intra-Class-Correlations of Experience Sampling Measures

	Range	Grand-Mean (SD)	Grand-SD (SD)	Women		Men	
				ICC (%)	1-ICC (%)	ICC (%)	1-ICC (%)
Predictors							
State communion motivation	-1 to 4	1.88 (0.79)	0.85 (0.28)	44.19	55.81	41.47	58.53
State closeness motivation	1 to 7	4.84 (1.28)	1.04 (0.43)	55.97	44.03	56.48	43.52
State time-spending motivation	0 to 10	7.00 (1.89)	2.18 (0.96)	38.59	61.41	38.20	61.80
State agency motivation	-1 to 4	0.97 (0.55)	0.50 (0.19)	53.09	46.91	49.38	50.62
State independence motivation	-1 to 4	1.11 (0.67)	0.78 (0.28)	39.53	60.47	38.13	61.87
State power motivation	-1 to 4	0.88 (0.64)	0.62 (0.26)	50.60	49.40	42.72	57.28
Outcomes							
Communion Behavior Index	-2.5 to 6.5	1.79 (0.89)	1.22 (0.31)	31.36	68.64	33.96	66.04
Agency Behavior Index	-2 to 7.5	0.68 (0.47)	0.82 (0.27)	18.71	81.29	26.10	73.90
IPC kindness	0 to 1	0.74 (0.11)	0.13 (0.04)	36.42	63.58	38.60	61.40
IPC dominance	0 to 1	0.54 (0.08)	0.12 (0.05)	29.16	70.84	25.71	74.29
Time spent alone	0 to 10	3.00 (1.38)	2.95 (0.70)	15.60	84.40	17.01	82.99
Time spent for interests	0 to 10	2.72 (1.64)	2.57 (0.80)	25.99	74.01	26.67	73.33
Covariates							
Time spent with partner	0 to 10	4.39 (1.45)	3.51 (0.67)	12.51	87.49	14.12	85.88
Amount of duties	0 to 10	4.40 (1.62)	3.17 (0.78)	20.49	79.51	17.81	82.19

Note. N = 510 individuals in 259 couples (256 women, 254 men). Grand-Mean = Mean of the intra-individual (person) means, with the standard deviation of these (person) means from the Grand-Mean in parentheses, Grand-SD = Mean of the intra-individual (person) standard deviations, with the standard deviation of these (person) SDs from the Grand-SD in parentheses, ICC = between-person variance, 1-ICC = within-person variance and error variance, IPC = interpersonal circumplex.

Prediction of Behavior and Time Reports in Reduced Models

Following our preregistration, we predicted self-reported behavior by different motivational variables first in *separate* models, and adopt analogous models for the prediction of partner-reports. That is, each of these models (termed “reduced models”) contained only one of the assessed motivational variables as predictor. An exception to this procedure is applied to the models with motivational states as predictors (assessed at surveys prior to the outcomes¹¹): We preregistered to person-mean center these state variables and additionally include the aggregated states in the model (“centered within context with reintroduction of the subtracted means at level 2” = CWCM method; Zhang, Zyphur, & Preacher, 2009, p. 709). This allows disentangling within-subject effects by motivational states from between-subject effects by aggregated states.¹²

The results are shown in Tables 5 and 6 (see “Reduced Models”): All examined types of motivational variables were predictive of one or more of the examined self-reported outcomes in at least one motive domain.

Communion. The preregistered communal outcomes were a) an index of poi communal behaviors, b) an overall evaluation of kind behavior towards the partner, and c) the amount of time spent alone. The observed effects went in the assumed directions.

All state and aggregated communal motivation measures predicted the communal behavior index and the kindness evaluation, irrespective of the report stemming from the individuals themselves or their partners. The amount of time individuals spent alone was only predicted by the two more outcome-specific closeness and time spending motivation items, but not by the communion motivation scale.

The two explicit dispositional measures, the intimacy motive and the desire for closeness, had similar effects. For all self-reports and the partner-reported kindness evaluation, both measures were significantly predictive. Ad-

¹¹That is, motivation at time-point t_i predicted the report given at time-point t_{i+1} about the behavior that happened since t_i . Answers were therefore only included in the analyses, when both surveys t_i and t_{i+1} were completed. Outcomes assessed in morning surveys were preregistered to be excluded as the prior t_i time-point would refer to the evening survey of the last day and overnight effects might bias the results.

¹²In the preregistration we only specified this CWCM method for analyzing the effect of motivational states. For the effect of aggregated states, we did not mention this procedure. However, a model containing only the aggregated states as predictor disregards confounding with level-1 variation. Therefore, models with aggregated states, but without the state motivation are presented only in the Supplemental Materials. For CWCM results shown in tables in this paper, it is highlighted that the analyses for the aggregated motivation as predictors do not follow the preregistered analysis plan.

ditionally, the more global explicit intimacy motive significantly predicted the partner-reported communal behavior index. The last examined dispositional measure, the implicit pnCommunion, significantly predicted the self-reported kindness evaluation and the time spent alone, but neither the self-reported communal behavior index nor any partner-report.

Independence. For the domain of independence, we preregistered that motivational variables would predict the amount of time individuals spent for own interests.¹³ While both state independence motivation and aggregated state independence motivation predicted the time spent for interests in the assumed direction, neither explicit nor implicit motives had significant effects.

Agency/Power. For the agentic power domain, we preregistered two behavioral reports analogous to the communal behavior reports: a) An index of specific agentic behaviors and b) an overall evaluation of dominant behavior towards the partner. In our preregistration, we were inconsistent across the different analysis levels, regarding whether agentic motivation or power motivation is assumed to predict the behavior reports: On level 1 as a state, we preregistered agentic motivation as predictor, but on level 2 as aggregated measure we preregistered power motivation. Therefore, we report the results of both scales as predictors on both analysis levels. In general, the effects were of similar sizes, with agency motivation on level 2 performing slightly better in predicting the agency behavior index.

State power and agency motivation, as well as aggregated power and agency motivational states were significantly predictive for both the self- and the partner-reported agency behavior index, but only for the self-reported dominance evaluation. The explicit power motive on the other hand was predictive for both the self- and partner-reported dominance evaluation, but only for the self-reported agency behavior index. The implicit pnPower only predicted the conceptually closer dominance self-report significantly, but not the partner-report or any report of the behavior index. The implicit pnAgency yielded no significant predictive effects on any outcome.

¹³Compared to time spent alone, which might be more strongly influenced by external demands (e.g., when one is usually alone at work), we had a priori considered time spent for own interests as the more meaningful behavioral outcome of independence motivation. However, as spending time completely for oneself also facilitates the experience of independence, the Supplemental Materials show the results of independence motivational variables predicting the time spent alone.

Comparison of Effect Sizes

We descriptively compared the standardized regression coefficients from the between-person effects in the reduced models, that is, of the aggregated states and the motive dispositions (we did not include the state motivation in this comparison, as it explains within-person variance rather than between-person variance, thus having a different interpretation of the effect size). As a general pattern across motive domains, outcomes, and for both self- and partner-reports, the aggregated motivational states had the strongest effects in the majority of cases. They were followed by the effect sizes of the explicit motives and concluded with the lowest effects by the implicit motives.

Tables 5 and 6 also show the increase in explained variance (see “ ΔR^2 ”) after adding the fixed effects of a single motivational variable to baseline models (marginal R^2 , Johnson, 2014; Nakagawa & Schielzeth, 2013). We defined baseline models as models with the same dyadic structure and covariates, but without the respective motivational variables: Hence, for all but the CWCM models, these baseline models included only the survey index and a weekend dummy variable as predictors. CWCM baseline models included further covariates (see analysis plan) and, additionally, the motivational variable on the level that is not under consideration (e.g., when examining the incremental effect of motivational states, then the aggregated motivation is already included in the baseline model). This comparison of the ΔR^2 allows taking into account the variance that was additionally explained by the level 1 motivational states. In general, the discussed pattern of the standardized beta coefficient sizes for between-person effects is mirrored in the explained variance. Overall, the magnitude of the added explained variances was higher in the communion domain (maximum observed $\Delta R^2 = 9.3\%$) compared to the agency domain (maximum observed $\Delta R^2 = 1.1\%$). On the within-person level, the motivational states explained mixed amounts of total outcome variances up to 1.8% ¹⁴, rarely higher than the variance explained by aggregated states (this was only the case when predicting outcomes related to time), and often between the range of explicit and implicit motives.

The analyses mentioned so far all relate to the amount of *total* variance of the outcomes that could be explained by the motivational variables. The total variance is however a mix of between- and within-person variance (see ICCs in Table 4), and each variance part can only be explained by predictors on the respective level. Hence, even when the amount of total explained variance by

¹⁴In comparison, the baseline models for motivational states that included situational constraints on level 1 explained up to 10.4% of variance.

a certain predictor is low, this could be misleading if there is a low amount of variance on the analysis level of the predictor (Rights & Sterba, 2019). As complementary analysis, we calculated the explained variance of the motivational states relative to the *available* within-person variance in the outcomes (see values in parentheses in the “ ΔR^2 ” column). However, even when taking into account the amount of within-person variance that could have been explained by the motivational states, the effect sizes remain similarly low. This shows that a large amount of the available within-person variance could not be explained by motivational states.

Incremental Contributions in Full Models

Complementary to the effects of single motivational variables in the reduced models, we further calculated models that included all motivational variables simultaneously (termed “full models”, see again Tables 5 and 6), to examine which predictors have incremental effects. In our case, incremental effects are only meaningful for predictors on the between-person level, as motivational states are the only predictors on the within-person level.¹⁵

Irrespective of the motive domain and the source of the report, an incremental contribution of aggregated state motivation beyond the other measures could be observed for almost all outcomes for which these variables were already predictive in the reduced models. This was also the case for explicit and implicit power motives: If they were predictive of the agency behavior index or dominant behavior reports when they were the only predictors, they were still significant when all other motivational variables were included in the models as well.

For communion, some variables predicted outcomes significantly in the reduced models, but did not provide significant incremental contributions in the full models: This was the case for pnCommunion, which did not provide any incremental contributions beyond the other motivational variables. Further, the explicit intimacy motive only had an incremental effect for the prediction of self-reported kindness, whereas in the reduced models it was a significant predictor for every outcome, irrespective of self- or partner-report. For time spent alone as outcome, no between-person motivational variable had a significant incremental effect, whereas almost all variables were significant predictors in the reduced models.

¹⁵For those variables, incremental effects could only be observed beyond other within-person variables. The results of motivational states in the full models therefore generally mirror the effects found in the reduced models. Any differences should be due to listwise deletion generated by missing values on the other variables.

Table 5

Prediction of Communion Behavior and Time Spent Alone by Different Communal Motivational Variables

	Self-Reported Behavior: Reduced / Full Model			Partner-Reported Behavior: Reduced / Full Model		
	$\beta \pm$ Range of 95% CI	p	ΔR^2 (within)	$\beta \pm$ Range of 95% CI	p	ΔR^2 (within)
DV: Comm Behavior Index						
State Comm Motivation CWCM	0.12 ± 0.02 / 0.12 ± 0.02	< .001 ^a / < .001	0.6% (0.9%)	0.08 ± 0.02 / 0.08 ± 0.02	< .001 / < .001	0.3% (0.5%)
Mean Comm Motivation CWCM	0.31 ± 0.07 / 0.27 ± 0.07	< .001 / < .001	4.4%	0.13 ± 0.07 / 0.13 ± 0.07	< .001 / .001	0.9%
State CI Motivation CWCM	0.13 ± 0.02 / 0.13 ± 0.02	< .001 ^a / < .001	0.5% (0.8%)	0.12 ± 0.02 / 0.12 ± 0.02	< .001 / < .001	0.4% (0.6%)
Mean CI Motivation CWCM	0.20 ± 0.06 / 0.14 ± 0.06	< .001 / < .001	3.0%	0.16 ± 0.06 / 0.16 ± 0.06	< .001 / < .001	2.0%
State TS Motivation CWCM	0.11 ± 0.02 / 0.11 ± 0.02	< .001 ^a / < .001	0.4% (0.6%)	0.08 ± 0.02 / 0.08 ± 0.02	< .001 / < .001	0.4% (0.5%)
Mean TS Motivation CWCM	0.17 ± 0.08 / 0.08 ± 0.08	< .001 / .056	1.6%	0.14 ± 0.08 / 0.13 ± 0.08	< .001 / .002	1.2%
Explicit Desire: Closeness	0.14 ± 0.04 / 0.08 ± 0.04	< .001 ^a / .001	1.8%	0.03 ± 0.05 / -0.01 ± 0.05	.205 / .692	0.1%
Explicit Intimacy Motive	0.07 ± 0.05 / 0.01 ± 0.05	.001^a / .811	0.5%	0.05 ± 0.04 / 0.04 ± 0.04	.045 / .166	0.2%
pnCommunion	0.03 ± 0.04 / 0.00 ± 0.04	.100 ^a / .907	0.1%	0.02 ± 0.04 / 0.03 ± 0.04	.252 / .170	0.0%
DV: IPC Kindness						
State Comm Motivation CWCM	0.12 ± 0.02 / 0.12 ± 0.02	< .001 ^a / < .001	0.7% (1.1%)	0.07 ± 0.02 / 0.07 ± 0.02	< .001 / < .001	0.4% (0.6%)
Mean Comm Motivation CWCM	0.30 ± 0.07 / 0.21 ± 0.07	< .001 / < .001	4.5%	0.14 ± 0.08 / 0.12 ± 0.08	< .001 / .005	1.0%
State CI Motivation CWCM	0.22 ± 0.02 / 0.21 ± 0.02	< .001 ^a / < .001	1.8% (2.8%)	0.14 ± 0.02 / 0.14 ± 0.02	< .001 / < .001	0.7% (1.1%)
Mean CI Motivation CWCM	0.37 ± 0.06 / 0.29 ± 0.06	< .001 / < .001	9.3%	0.21 ± 0.07 / 0.18 ± 0.07	< .001 / < .001	3.0%
State TS Motivation CWCM	0.16 ± 0.02 / 0.16 ± 0.02	< .001 ^a / < .001	1.6% (2.6%)	0.10 ± 0.02 / 0.10 ± 0.02	< .001 / < .001	0.7% (1.1%)
Mean TS Motivation CWCM	0.37 ± 0.08 / 0.26 ± 0.08	< .001 / < .001	6.6%	0.23 ± 0.08 / 0.19 ± 0.08	< .001 / < .001	2.5%
Explicit Desire: Closeness	0.22 ± 0.05 / 0.14 ± 0.05	< .001 ^a / < .001	4.7%	0.10 ± 0.05 / 0.08 ± 0.05	< .001 / .011	1.0%
Explicit Intimacy Motive	0.14 ± 0.05 / 0.05 ± 0.05	< .001 ^a / .038	2.2%	0.05 ± 0.05 / 0.03 ± 0.05	.033 / .256	0.4%
pnCommunion	0.05 ± 0.05 / 0.00 ± 0.05	.020^a / .864	0.3%	0.01 ± 0.04 / 0.00 ± 0.04	.597 / .977	0.0%
DV: Time Spent Alone						
State Comm Motivation CWCM	-0.01 ± 0.02 / -0.01 ± 0.02	.208 ^a / .386	0.0% (0.0%)			
Mean Comm Motivation CWCM	-0.01 ± 0.06 / 0.02 ± 0.06	.685 / .553	0.0%			
State CI Motivation CWCM	-0.05 ± 0.02 / -0.05 ± 0.02	< .001 ^a / < .001	0.1% (0.1%)			
Mean CI Motivation CWCM	-0.07 ± 0.05 / -0.03 ± 0.05	.009 / .227	0.2%			
State TS Motivation CWCM	-0.10 ± 0.02 / -0.10 ± 0.02	< .001 ^a / < .001	0.6% (0.7%)			
Mean TS Motivation CWCM	-0.08 ± 0.06 / -0.04 ± 0.06	.010 / .251	0.2%			
Explicit Desire: Closeness	-0.05 ± 0.04 / -0.04 ± 0.04	.005^a / .057	0.2%			
Explicit Intimacy Motive	-0.04 ± 0.04 / -0.02 ± 0.04	.023^a / .334	0.1%			
pnCommunion	-0.03 ± 0.03 / -0.02 ± 0.03	.038^a / .274	0.1%			

Note. ^a one-sided p -value. $N = 29405$ -60809 surveys from 501-510 individuals. Reduced model = model with preregistered covariates and only one motivational variable as predictor (exception: CWCM = model including both level 1 state variable (person-mean centered) and level 2 aggregated state variable, to disentangle within- and between-person variance). Full model = model with preregistered covariates and all motivational variables of a domain as predictors (exception: to avoid collinearity problems only communion motivation (state/aggregated), but not closeness or time-spending motivation (states/aggregated) is included in the full models of motive dispositions; vice versa closeness and time-spending motivation (states/aggregated) are *only* included in their respective full models). CI = Confidence Interval, ΔR^2 = additional total outcome variance explained by the motivational variable as fixed effect compared to a model containing only the covariates of the reduced model; for state motivation additionally the amount of explained variance relative to the amount of within-person variance (rather than the total variance) is shown in parentheses. DV = dependent variable, Comm = Communion, CI = Closeness, TS = Time-Spending, pnCommunion = partner-related need for communion (implicit motive), IPC = interpersonal circumplex. Covariates included in the models, but not shown: weekend dummy-variable, survey index; additional covariate for the IPC as DV; the other IPC dimension 'dominance'; additional covariate for CWCM models and behavior as DV; time spent with partner; additional covariate for CWCM models with time as DV; amount of duties. Preregistered analyses are printed in black, and additionally bold if the results are significant after controlling the false-discovery-rate at $\alpha = 0.05$. Other analyses are printed in gray, and additionally bold if the results are significant at $\alpha = 0.05$.

Table 6
Prediction of Time Spent for Own Interests and Agentic Behavior by Different Agentic Motivational Variables

	Self-Reported Behavior: Reduced / Full Model			Partner-Reported Behavior: Reduced / Full Model		
	$\beta \pm$ Range of 95% CI	p	ΔR^2t (within)	$\beta \pm$ Range of 95% CI	p	ΔR^2t (within)
DV: Time Spent For Own Interests						
State independence motivation CWCM	0.09 ±0.02 / 0.09 ±0.02	< . 001 ^a / < .001	0.7% (0.9%)			
Mean independence motivation CWCM	0.10 ±0.07 / 0.09 ±0.07	.005 / .014	0.3%			
Explicit desire: Being alone	0.03 ±0.04 / 0.03 ±0.04	.066 ^a / .180	0.1%			
pnAgency	-0.02 ±0.04 / -0.03 ±0.04	.771 ^a / .130	0.0%			
pnIndependence	-0.01 ±0.04 / -0.01 ±0.04	.643 ^a / .668	0.0%			
DV: Agency Behavior Index						
State agency motivation CWCM	0.03 ±0.02 / 0.03 ±0.02	< . 001 ^a / .003	0.1% (0.1%)	0.02 ±0.02 / 0.02 ±0.02	.043 / .042	0.0% (0.0%)
Mean agency motivation CWCM	0.20 ±0.07 / 0.19 ±0.07	< .001 / < .001	1.0%	0.12 ±0.08 / 0.11 ±0.08	.005 / .009	0.3%
State power motivation CWCM	0.04 ±0.02 / 0.04 ±0.02	< .001 / < .001	0.0% (0.0%)	0.02 ±0.02 / 0.02 ±0.02	.025 / .024	0.0% (0.0%)
Mean power motivation CWCM	0.15 ±0.06 / 0.14 ±0.06	< .001 / < .001	0.9%	0.09 ±0.06 / 0.09 ±0.06	.004 / .006	0.4%
Explicit power motive	0.05 ±0.04 / 0.04 ±0.04	.011 ^a / .037	0.2%	0.03 ±0.04 / 0.03 ±0.04	.162 / .146	0.1%
pnAgency	-0.02 ±0.04 / -0.02 ±0.04	.888 ^a / .195	0.1%	-0.01 ±0.04 / -0.02 ±0.04	.562 / .269	0.0%
pnPower	0.02 ±0.04 / 0.02 ±0.04	.104 ^a / .380	0.1%	-0.01 ±0.04 / 0.00 ±0.04	.764 / .853	0.0%
DV: IPC Dominance						
State agency motivation CWCM	0.02 ±0.02 / 0.02 ±0.02	.030 ^a / .056	0.0% (0.0%)	0.00 ±0.02 / 0.00 ±0.02	.909 / .878	0.0% (0.0%)
Mean agency motivation CWCM	0.12 ±0.08 / 0.10 ±0.08	.004 / .016	0.4%	0.03 ±0.08 / 0.00 ±0.08	.486 / .921	0.0%
State power motivation CWCM	0.02 ±0.02 / 0.02 ±0.02	.031 / .031	-0.1% (-0.1%)	0.00 ±0.02 / 0.00 ±0.02	.640 / .651	0.0% (0.0%)
Mean power motivation CWCM	0.12 ±0.06 / 0.11 ±0.06	< .001 / .001	0.7%	0.07 ±0.07 / 0.05 ±0.07	.054 / .158	0.2%
Explicit power motive	0.07 ±0.04 / 0.07 ±0.04	.001 ^a / .004	0.5%	0.07 ±0.05 / 0.07 ±0.05	.001 / .002	0.6%
pnAgency	-0.01 ±0.04 / -0.01 ±0.04	.602 ^a / .523	0.0%	0.02 ±0.04 / 0.02 ±0.04	.300 / .429	0.0%
pnPower	0.05 ±0.04 / 0.04 ±0.04	.009 ^a / .046	0.2%	0.04 ±0.04 / 0.04 ±0.04	.060 / .085	0.2%

Note. ^a one-sided p -value. $N = 32664$ -60809 observations in 501-510 individuals. Reduced model = model with preregistered covariates and only one motivational variable as predictor (exception: CWCM = model including both level 1 state variable (person-mean centered) and level 2 aggregated state variable, to disentangle within- and between-person variance). Full model = model with preregistered covariates and all motivational variables of a domain as predictors (exception: to avoid collinearity problems pnIndependence and pnPower are *not* included in the full models of pnAgency, and power motivation (state/aggregated) is *not* included in the full models of agency motivation; vice versa pnAgency is *only* included in the full models of pnAgency, and agency motivation (state/aggregated) is *only* included in the full models of agency motivation). CI = Confidence Interval, ΔR^2t = additional total outcome variance explained by the motivational variable as fixed effect compared to a model containing only the covariates of the reduced model; for state motivation additionally the amount of explained variance relative to the amount of within-person variance (rather than the total variance) is shown in parentheses. DV = dependent variable, IPC = interpersonal circumplex, pn = partner-related need (implicit motive). Covariates included in the models, but not shown: weekend dummy-variable, survey index; additional covariate for the IPC as DV: the other IPC dimension 'kindness'; additional covariate for CWCM models with time as DV: amount of duties; additional covariate for CWCM models and behavior as DV: time spent with partner. Preregistered analyses are printed in black, and additionally bold if the results are significant after controlling the false-discovery-rate at $\alpha = 0.05$. Other analyses are printed in gray, and additionally bold if the results are significant at $\alpha = 0.05$.

Discussion

A central goal of the current study was the comparison of motivational variables as predictors of behavior reports in intimate relationships. We approached this comparison from various perspectives: (a) We took a within-person perspective in addition to a between-person perspective on motivational variables, enabling to predict different variance sources of the behavior reports; (b) We looked at different motive domains, specifically communal, independence and power motivational variables predicting corresponding behavioral outcomes; (c) We included more general and more specific behavioral reports; (d) We examined partner-reports next to self-reports of the behavioral outcomes; (e) We compared bivariate to incremental contributions.

To begin with, our results replicate and extend previous research, which demonstrated that between-person differences in motivational variables predict average behavioral levels: For different kinds of behaviors, persons with higher average levels of motivation and those with stronger motive dispositions showed on average more instrumental behavior. This corroborates a central assumption of motive disposition theory for the domain of intimate relationships: that motives as inter-individual differences select and energize corresponding instrumental behavior. From a functional perspective, however, a crucial test of the theory lies on the within-person level of motivational states. In this regard, our results directly replicate a prior study showing that communal motivational states predict subsequent specific communal behavioral acts as self-reports (Zygar et al., 2018). The present study extended this finding conceptually with other operationalizations of communal motivation, and with other outcomes, which were additionally partner-reported. Finally, our results support the idea of a process in which motivational states precede instrumental behavior likewise for the independence and power motive domains in intimate relationships.

The Special Case of Motivational States

The effects of state motivation on a within-person level were generalizable and robust, in terms of being consistent across outcomes, motive domains and source of the behavior reports (self vs. partner): The occurrence of a motivational state at a certain time point predicted more instrumental self- and partner-reports of behavior in the next hours, which was observable for all kinds of behaviors. For example, a stronger state motivation for communion came along with more subsequent self- and partner-reported communal behav-

ior and a stronger state motivation for independence came along with subsequently more time spent for own interests. Only the time individuals spent alone was not consistently predicted by the different operationalization of communal state motivation (two out of three operationalizations were significant, namely those that especially emphasized the closeness aspect of communion).

The effects on self-reported behavior largely corresponded with those on partner-reported behavior, with only one exception: State agency and power motivation predicted self-, but not partner-reported dominance behavior. The fact that almost all results also applied to partner-reports of behavior (albeit, with smaller effect sizes) makes a strong case for the importance of (self-reported) motivational states, as person-specific common-method biases are less of a problem for these partner-reports. That is, if the effects were only this consistent because both predictor and outcome were assessed with ESM and individuals have a bias in their response to experience sampling items in general, then the effects would not generalize to partner-reports (at least not unless the partner shared the bias).

From a theoretical perspective, it is highly plausible to find such consistent effects as the link between motivational *states* and behavior is at the core of the assumed function of motive *dispositions* (McClelland, 1987, see also DynaMoS model in Zygar et al., 2018): Primarily the within-person process of aroused motivation that selects and energizes behavior facilitates a satisfaction of the underlying motive of a person. The approximation of this process by looking at between-person differences in behavior and aggregated motivation or globally assessed motives is therefore a proxy that can be potentially misleading (see Molenaar, 2004; Molenaar & Campbell, 2009). In the specific context of intimate relationships, our results replicate past research showing that between-person as well as within-person effects of motivation on behavior reports point in the same direction (see Zygar et al., 2018). Still, this correspondence of between- and within-person effects cannot be taken for granted and has to be shown for each new domain. Therefore, effects that theoretically are located on a within-person level should always be investigated with appropriate within-person research designs.

As a caveat, the effect sizes of motivational states were generally small in terms of additional explained variance compared to a model with only the aggregated states and some covariates included. They were, however, of comparable size to the between-person effects of explicit or implicit motive dispositions. In general, a notable advantage of state variables compared to dispositional variables is their ability to explain within-person variance, which would

otherwise be completely treated as error variance in between-person analyses. Still, even when relativizing the amount of explained variance by the motivational states to the amount of within-person variance in the outcome (instead of to the total amount of outcome variance), the effect sizes remained to be low. That is, there were substantial amounts of within-person variance in the outcomes, which could not be explained by the motivational states.

Situational influences in everyday life are certainly one reason for the large amount of unexplained variance, as they are important factors for the occurrence of behavior, which were considered in our analyses only remotely as covariates that pose absolute barriers to the implementation of behavior. Opportunities that are present in situations, or less restrictive situational constraints (e.g., other persons being present or not) can also influence the occurrence of certain behaviors (e.g., what would be deemed appropriate behavior). In the current analysis, we similarly disregarded interpersonal influences on motivation and behavior, as the behavior of one partner constitutes again an important situational influence for the behavior of the other partner. The energizing function of motives (McClelland, 1987) would further suggest that differences in motive dispositions could explain why some individuals overcome certain situational barriers, while others do not, which could be represented by an interaction between motive disposition strength and motivational state predicting behavior in certain situations. All of these factors were disregarded in our analyses, which were only concerned with main effects of motivational variables controlled for certain situational barriers. Hence, a promising avenue for future research is to specify the main effects we found on the within-person level by considering the situations individuals are in, the dyadic influences in the relationship, as well as cross-level interactions with motive dispositions.

The Effects of Between-Person Motivational Variables

The assessment of state measures does not only have the advantage of enabling within-person analyses, it also allows the computation of a person average of these states, which can be compared with other person-level variables, in our case, implicit and explicit motive dispositions. As a general result, between-person motivational measures do predict the average occurrence of certain behaviors. Hence, these variables can explain some of the between-person variance observed in behavior reports.

Further, in some cases, the classical motive dispositions had an incremental effect beyond the effects of state and average motivational states. This

could be explained by learned habits which trigger behavior without a corresponding motivation. It could serve as a shortcut in relationship situations which routinely follow a pattern (e.g., a sign of affection on a reunion) or which do not require realizing behavior against any barriers. A central question would be how such habituated behavior differs from motivated behavior, for example in the affective satisfaction gained from these behavior types (see the affect-amplifying nature of motivation, Zygar et al., 2018). This result seems to provide evidence against the mediation hypothesis assuming that motives only influence behavior via concrete motivational states. However, this conclusion would require that all variance of states is adequately measured. For various reasons this is unrealistic. For example, an ESM study cannot continuously measure the current motivational state, and our time windows between ESM surveys provide plenty of opportunities for states to change after an assessment. Hence, the relationship between motivation and behavior could differ depending on the time scale that is considered.¹⁶

Finally, the effects of aggregated motivation on behavior might represent - on the one hand - the effect of a motivational climate due to more or less stable situations in the relationship: For example, when situations repeatedly provide motivationally relevant incentives for individuals, they repeatedly experience motivation and seem to repeatedly get the opportunities to realize this in according behavior. Individuals in situations with less incentives might experience less motivation and therefore have no need or no opportunities to act in an instrumental way. On the other hand, aggregated states might directly reflect inter-individual differences in motive dispositions through the propensity to experience motivational states in general.

Comparison: Effect Sizes. Comparing the different between-person motivational measures showed that across motive domains aggregated states generally had the strongest effects that were also, in most cases, incremental to the other effects. Again, this was also the case for most partner-reports of behavior, underlining the value of assessing motivational states also for explaining between-person differences.

It has to be considered, however, that the aggregated states were measured with higher reliabilities than the other dispositional measures, which influences the effect sizes. Moreover, as the aggregated states refer to the

¹⁶To that end, the Supplemental Materials present preregistered analyses showing that evening motivation to spend time together with the partner on the next day predicts communal behavior towards the partner on that day, while aggregated states still provide incremental effects. This could be extended in further research to even larger and also even shorter time frames.

same time period as the reported outcomes, they have a crucial advantage compared to the motive dispositions which were assessed before the experience sampling. A future study would need to predict behavior that is assessed separately from the occurrence of average motivational states, for example in a delayed experience sampling period. This could reveal if aggregated states still maintain larger predictive effects than motive dispositions if they do not relate to the same time frame as the outcomes. Additionally, future research could provide insights about the stability of aggregated states: This would allow to interpret differences across time periods in terms of conceptual meaning of aggregated states compared to motive dispositions.

Comparison: Types of Predicted Outcomes. Regarding the types of outcomes that the between-person motivational variables predicted, no systematic differences could be observed between aggregated states and explicit measures: For most outcomes, both kinds of measures were predictive. However, some differences emerged for implicit motives: First, the only time-related variable that was predicted by implicit motives was the time spent alone by pnCommunion, although we had also hypothesized that pnIndependence would predict the time spent for own interests. As spending time alone crucially hinders the fulfillment of intimacy, it represents a barrier to a basic boundary condition to be able to satisfy the communion motive, and might therefore be especially relevant (see Bischof, 2008). Spending time for one's own hobbies might not be enough to feel as an independent individual (although it should facilitate it), an experience needed for the satisfaction of independence motivation. Future research should therefore investigate behavior that is even more closely related to independent behavior, such as making autonomous decisions or spending time for one's own interests *while simultaneously being alone* (compared to e.g., group sports).

Second, pnCommunion and pnPower were predictive of the two evaluations of overall behavior quality captured by the IPC measure, but not for the occurrence of specific behavior acts captured by the behavioral indices. Post-hoc, it might be argued that compared to the specific behavior reports of the index, the IPC behavior reports are more indirect, thereby providing more opportunities to factor non-declarative cues into the assessment. That is, the rather vague assessment of kindness and dominance with the IPC grid might encourage individuals to include nonverbal behavior towards their partner they perceived as kind or dominant, such as an encouraging smile. Vice versa, the behaviors comprised in the index might be more declarative (e.g., the threshold to indicate that affection was shown might be perceived as requiring some

form of verbal or physical interaction).

It has to be noted though, that the study was not designed to differentiate declarative from non-declarative measures or to evaluate the effect of hindering behavior (see e.g., Hagemeyer et al., 2016). We expected that all motivational variables would predict all behavior self-reports (although difference in strength of effects were to be expected), hence our interpretations of the differences in results for different outcomes can only remain speculative.

Finally, whereas pnPower did predict dominant behavior, the superordinate pnAgency did not. This might be due to the specificity of the outcome measure, which directly matches a power goal, rather than a broader differentiation from other persons. Nonetheless, the result suggests that for the prediction of behavior, future studies should investigate the subcomponents of pnAgency separately (for a similar argument from a psychometric perspective, see Schönbrodt et al., 2019). Ideally, this might be done with partner-specific projective measures which were specifically designed to assess these two facets of agency.

Limitations of the Assessed Behavior Reports

Although the assessment of behavior reports with the experience sampling method has higher ecological validity than traditional questionnaire methods by assessing the behavior during individuals' daily lives and reducing memory biases (Bolger et al., 2003), our study still did not assess objective behavior. Arguably, partner-reports of behavior can be expected to be less influenced by biases that stem from "common rater effects" (Podsakoff, MacKenzie, Lee, & Podsakoff, 2003), and might be more objective to that respect. This, however, does not preclude that other biases can distort the perceptual process of the partner. Future research should combine the experience sampling with mobile sensing (Harari, Gosling, Wang, & Campbell, 2015) to objectively assess whether the partner was contacted or what kind of interactions occurred. Additionally, studies using participant observation might provide more valid assessments of the actual agentic and communal qualities of behavior.

Conclusion

For a basic research question in motivational psychology our study demonstrated how the assessment of psychological states with the experience sampling method can extend our understanding of daily processes. Specifically, our results show that the prediction of behavior reports profits from a

within-person perspective on motivational variables, that is, from the assessment of motivational states in addition to motive dispositions. These motivational states predict how individuals behave in their relationships in daily life, as reported by themselves and their partners. As a within-person effect, this complements the between-person effects of aggregated states and motive dispositions that in turn illustrate which individuals on average behave more in certain ways than other individuals. Future research should therefore consider both levels of analysis to capture the full picture of examined effects, and to better understand the psychological processes going on in individuals' lives.

References

- Abele, A. E., Cuddy, A. J. C., Judd, C. M., & Yzerbyt, V. Y. (2008). Fundamental dimensions of social judgement. *European Journal of Social Psychology, 38*, 1063–1065. doi:10.1002/ejsp.574
- Ackerman, R. A., & Corretti, C. A. (2015). Pathological personality traits and intimacy processes within roommate relationships. *European Journal of Personality, 29*, 152–172. doi:10.1002/per.1991
- Aron, A., Aron, E. N., & Smollan, D. (1992). Inclusion of Other in the Self Scale and the structure of interpersonal closeness. *Journal of Personality and Social Psychology, 63*(4), 596–612. doi:10.1037/0022-3514.63.4.596
- Arslan, R. C., Walther, M. P., & Tata, C. S. (2020). Formr: A study framework allowing for automated feedback generation and complex longitudinal experience-sampling studies using R. *Behavior Research Methods, 52*, 376–387. doi:10.3758/s13428-019-01236-y
- Arslan, R., & Tata, C. (2017). *formr.org survey software (version v0.16.13-version v0.17.16)*. Retrieved from <https://formr.org/>
- Atkinson, J. W. (1957). Motivational determinants of risk-taking behavior. *Psychological Review, 64*(6), 359–372.
- Atkinson, J. W., & Birch, D. &. (1970). *The dynamics of action*. Oxford, United Kingdom: John Wiley.
- Aust, F., & Barth, M. (2018). *Papaja: Create APA manuscripts with R Markdown*. Retrieved from <https://github.com/crsh/papaja>
- Back, M. D., & Vazire, S. (2015). The social consequences of personality: Six suggestions for future research. *European Journal of Personality, 29*(2), 296–307. doi:10.1002/per.1998
- Bakan, D. (1966). *The duality of human existence*. Boston, IL, USA: Beacon Press.
- Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language, 68*(3), 1–43. doi:10.1016/j.jml.2012.11.001
- Barton, K. (2018). *MuMIn: Multi-model inference*. Retrieved from <https://CRAN.R-project.org/package=MuMIn>
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1–48.

doi:10.18637/jss.v067.i01

- Baumert, A., Schmitt, M., Perugini, M., Johnson, W., Blum, G., Borkenau, P., . . . Wrzus, C. (2017). Integrating personality structure, personality process, and personality development. *European Journal of Personality, 31*(5), 503–528. doi:10.1002/per.2115
- Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B. Methodological, 57*(1), 289–300. doi:10.1111/j.2517-6161.1995.tb02031.x
- Bischof, N. (1975). A systems approach toward the functional connections of attachment and fear. *Child Development, 46*, 801–817. doi:10.2307/1128384
- Bischof, N. (2008). *Psychologie: Ein Grundkurs für Anspruchsvolle*. Stuttgart, Germany: Kohlhammer.
- Bliese, P. D. (2000). Within-group agreement, non-independence, and reliability: Implications for data aggregation and analysis. In *Multilevel theory, research, and methods in organizations: Foundations, extensions, and new directions*. (pp. 349–381). San Francisco, CA, USA: Jossey-Bass.
- Bolger, N., Davis, A., & Rafaeli, E. (2003). Diary methods: Capturing life as it is lived. *Annual Review of Psychology, 54*, 579–616. doi:10.1146/annurev.psych.54.101601.145030
- Bolger, N., & Laurenceau, J.-P. (2013). *Intensive longitudinal methods*. New York, NY, USA: Guilford.
- Craig, J.-A., Koestner, R., & Zuroff, D. C. (1994). Implicit and self-attributed intimacy motivation. *Journal of Social and Personal Relationships, 11*, 491–507.
- Csikszentmihalyi, M., & Larson, R. (1987). Validity and reliability of the Experience-Sampling Method. *The Journal of Nervous and Mental Disease, 175*(9), 526–536. doi:10.1097/00005053-198709000-00004
- Denzinger, F., & Brandstätter, V. (2018). Stability of and changes in implicit motives. A narrative review of empirical studies. *Frontiers in Psychology, 9*, Article 777. doi:10.3389/fpsyg.2018.00777
- Dufner, M., Arslan, R. C., Hagemeyer, B., Schönbrodt, F. D., & Denissen, J. J. A. (2015). Affective contingencies in the affiliative domain: Physiological assessment, associations with the affiliation motive, and predic-

tion of behavior. *Journal of Personality and Social Psychology*, *109*(4), 662–676. doi:10.1037/pspp0000025

Gubler, H., & Bischof, N. (1991). A systems theory perspective. In M. E. Lamb & H. Keller (Eds.), *Infant development : Perspectives from german-speaking countries* (pp. 35–66). Hillsdale, NJ, USA: Lawrence Erlbaum Associates.

Hagemeyer, B., Dufner, M., & Denissen, J. J. A. (2016). Double dissociation between implicit and explicit affiliative motives: A closer look at socializing behavior in dyadic interactions. *Journal of Research in Personality*, *65*, 89–93. doi:10.1016/j.jrp.2016.08.003

Hagemeyer, B., & Neyer, F. J. (2012). Assessing implicit motivational orientations in couple relationships: The Partner-Related Agency and Communion Test (PACT). *Psychological Assessment*, *24*(1), 114–28. doi:10.1037/a0024822

Hagemeyer, B., Neyer, F. J., Neberich, W., & Asendorpf, J. B. (2013). The ABC of Social Desires: Affiliation, being alone, and closeness to partner. *European Journal of Personality*, *27*(5), 442–457. doi:10.1002/per.1857

Hagemeyer, B., Schönbrodt, F. D., Neyer, F. J., Neberich, W., & Asendorpf, J. B. (2015). When "together" means "too close": Agency motives and relationship functioning in coresident and living-apart-together couples. *Journal of Personality and Social Psychology*, *109*(5), 813–835. doi:10.1037/pspi0000031

Harari, G. M., Gosling, S. D., Wang, R. U. I., & Campbell, A. T. (2015). Capturing situational information with smartphones and mobile sensing methods. *European Journal of Personality*, *29*, 509–511. doi:10.1002/per.2032

Hoffman, L., & Stawski, R. S. (2009). Persons as contexts: Evaluating between-person and within-person effects in longitudinal analysis. *Research in Human Development*, *6*(2-3), 97–120. doi:10.1080/15427600902911189

Horowitz, L. M., Wilson, K. R., Turan, B., Zolotsev, P., Constantino, M. J., & Henderson, L. (2006). How interpersonal motives clarify the meaning of interpersonal behavior: A revised circumplex model. *Personality and Social Psychology Review*, *10*(1), 67–86. doi:10.1207/s15327957pspr1001_4

Johnson, P. C. D. (2014). Extension of nakagawa and schielzeth's r2glmm

- to random slopes models. *Methods in Ecology and Evolution*, 5(9), 944–946. doi:10.1111/2041-210X.12225
- Kelley, K. (2018). *MBESS: The MBESS R package*. Retrieved from <https://CRAN.R-project.org/package=MBESS>
- Köllner, M. G., & Schultheiss, O. C. (2014). Meta-analytic evidence of low convergence between implicit and explicit measures of the needs for achievement, affiliation, and power. *Frontiers in Psychology*, 5, Article 826. doi:10.3389/fpsyg.2014.00826
- Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, 82(13), 1–26. doi:10.18637/jss.v082.i13
- Lansing, J. B., & Heyns, R. W. (1959). Need affiliation and frequency of four types of communication. *The Journal of Abnormal and Social Psychology*, 58(3), 365–372. doi:10.1037/h0045906
- Lüdtke, O., Marsh, H. W., Robitzsch, A., & Trautwein, U. (2011). A 2×2 taxonomy of multilevel latent contextual models: Accuracy-bias trade-offs in full and partial error correction models. *Psychological Methods*, 16(4), 444–467. doi:10.1037/a0024376
- Mason, A., & Blankenship, V. (1987). Power and affiliation motivation, stress, and abuse in intimate relationships. *Journal of Personality and Social Psychology*, 52(1), 203–210. doi:10.1037/0022-3514.52.1.203
- McAdams, D. P., & Constantian, C. A. (1983). Intimacy and affiliation motives in daily living: An experience sampling analysis. *Journal of Personality and Social Psychology*, 45(4), 851–861. doi:10.1037/0022-3514.45.4.851
- McAdams, D. P., Healy, S., & Krause, S. (1984a). Social motives and patterns of friendship. *Journal of Personality and Social Psychology*, 47(4), 828–838. doi:10.1037/0022-3514.47.4.828
- McAdams, D. P., Jackson, R. J., & Kirshnit, C. (1984b). Looking, laughing, and smiling in dyads as a function of intimacy motivation and reciprocity. *Journal of Personality*, 52(3), 261–273.
- McClelland, D. C. (1970). The two faces of power. *Journal of International Affairs*, 24(1), 29–47.
- McClelland, D. C. (1975). *Power: The inner experience*. Oxford, United Kingdom: Irvington.

- McClelland, D. C. (1987). *Human Motivation*. New York, NY, USA: Cambridge University Press.
- McClelland, D. C., Koestner, R., & Weinberger, J. (1989). How do self-attributed and implicit motives differ? *Psychological Review*, *96*(4), 690–702. doi:10.1037/0033-295X.96.4.690
- Molenaar, P. C. M. (2004). A manifesto on psychology as idiographic science: Bringing the person back into scientific psychology, this time forever. *Measurement Interdisciplinary Research and Perspectives*, *2*(4), 201–218. doi:10.1207/s15366359mea0204
- Molenaar, P. C. M., & Campbell, C. G. (2009). The new person-specific paradigm in psychology. *Current Directions in Psychological Science*, *18*(2), 112–117. doi:10.1111/j.1467-8721.2009.01619.x
- Nakagawa, S., & Schielzeth, H. (2013). A general and simple method for obtaining R^2 from generalized linear mixed-effects models. *Methods in Ecology and Evolution*, *4*(2), 133–142. doi:10.1111/j.2041-210x.2012.00261.x
- Podsakoff, P. M., MacKenzie, S. B., Lee, J.-Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology*, *88*(5), 879–903. doi:10.1037/0021-9010.88.5.879
- R Core Team. (2019). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from <https://www.R-project.org/>
- Rheinberg, F. (2002). *Motivation*. Kohlhammer.
- Rights, J. D., & Sterba, S. K. (2019). Quantifying explained variance in multilevel models: An integrative framework for defining R-squared measures. *Psychological Methods*, *24*(3), 309–338. doi:10.1037/met0000184
- Robinson, M. D., & Clore, G. L. (2002). Belief and feeling: Evidence for an accessibility model of emotional self-report. *Psychological Bulletin*, *128*(6), 934–960. doi:10.1037/0033-2909.128.6.934
- Schönbrodt, F. D., & Gerstenberg, F. X. R. (2012). An IRT analysis of motive questionnaires: The Unified Motive Scales. *Journal of Research in Personality*, *46*, 725–742. doi:10.1016/j.jrp.2012.08.010
- Schönbrodt, F. D., Zygar-Hoffmann, C., Nestler, S., Pusch, S., & Hagemeyer, B. (2019). *Measuring motivational relationship processes in experience*

sampling: A reliability model for moments, days, and persons nested in couples. doi:10.31234/osf.io/6mq7t

- Schultheiss, O. C. (2001). An information processing account of implicit motive arousal. In M. L. Maehr & P. Pintrich (Eds.), *Advances in motivation and achievement* (12th ed., pp. 1–41). Greenwich, CT, USA: JAI Press.
- Schultheiss, O. C. (2008). Implicit motives. In O. P. John, R. W. Robins, & L. A. Pervin (Eds.), *Handbook of personality: Theory and research* (3rd ed., pp. 603–633). New York, NY, USA: Guilford.
- Schultheiss, O. C., Dargel, A., & Rohde, W. (2003). Implicit motives and sexual motivation and behavior. *Journal of Research in Personality*, *37*(3), 224–230. doi:10.1016/S0092-6566(02)00568-8
- Schultheiss, O. C., Kordik, A., Kullmann, J. S., Rawolle, M., & Rösch, A. G. (2009). Motivation as a natural linchpin between person and situation. *Journal of Research in Personality*, *43*(2), 268–269. doi:10.1016/j.jrp.2008.12.018
- Schultheiss, O., Strasser, A., Rösch, A., Kordik, A., & Graham, S. C. (2012). Motivation. In V. S. Ramachandran (Ed.), *Encyclopedia of human behavior* (pp. 650–656). Oxford, United Kingdom: Elsevier.
- Schüler, J., Job, V., Fröhlich, S. M., & Brandstätter, V. (2008). A high implicit affiliation motive does not always make you happy: A corresponding explicit motive and corresponding behavior are further needed. *Motivation and Emotion*, *32*(3), 231–242. doi:10.1007/s11031-008-9096-y
- Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2012). A 21 word solution. *SPSP Dialogue*, *26*(2), 1–4. doi:10.2139/ssrn.2160588
- Spangler, W. D. (1992). Validity of questionnaire and TAT measures of need for achievement: Two meta-analyses. *Psychological Bulletin*, *112*(1), 140–154. doi:10.1037//0033-2909.112.1.140
- Stanton, S. J., Hall, Julie L., & Schultheiss, O. C. (2010). Properties of motive-specific incentives. In O. C. Schultheiss & J. C. Brunstein (Eds.), *Implicit motives* (pp. 245–278). Oxford, United Kingdom: Oxford University Press.
- Stapleton, L. M., Yang, J. S., & Hancock, G. R. (2016). Construct meaning in multilevel settings. *Journal of Educational and Behavioral Statistics*, *41*(5), 481–520. doi:10.3102/1076998616646200

- Wickham, H., François, R., Henry, L., & Müller, K. (2018). *Dplyr: A grammar of data manipulation*. Retrieved from <https://CRAN.R-project.org/package=dplyr>
- Wiggins, J. S. (1979). A psychological taxonomy of trait-descriptive terms: The interpersonal domain. *Journal of Personality and Social Psychology, 37*(3), 395–412. doi:10.1037/0022-3514.37.3.395
- Winter, D. G. (1973). *The power motive*. New York, NY, USA: Free Press.
- Zhang, Z., Zyphur, M. J., & Preacher, K. J. (2009). Testing multilevel mediation using hierarchical linear models: Problems and solutions. *Organizational Research Methods, 12*(4), 695–719. doi:10.1177/1094428108327450
- Zurbriggen, E. L. (2000). Social motives and cognitive power-sex associations: Predictors of aggressive sexual behavior. *Journal of Personality and Social Psychology, 78*(3), 559–581. doi:10.1037//0022-3514.78.3.559
- Zurbriggen, E. L. (2011). Implicit motives and sexual conservatism as predictors of sexual behaviors. *The Journal of Social Psychology, 151*(5), 535–555. doi:10.1080/00224540903365372
- Zygar, C., Hagemeyer, B., Pusch, S., & Schönbrodt, F. D. (2018). From motive dispositions to states to outcomes: An intensive experience sampling study on communal motivational dynamics in couples. *European Journal of Personality, 32*(3), 306–324. doi:10.1002/per.2145
- Zygar-Hoffmann, C., Hagemeyer, B., Pusch, S., & Schönbrodt, F. D. (2020). *A large longitudinal study on motivation, behavior and satisfaction in couples: Research data from a four-week experience sampling study with a pre-, post-, and one-year follow-up-assessment [Translated title] (Version 1.0.0)* [data and documentation]. Trier, Germany: Center for Research Data in Psychology PsychData of the Leibniz Institute for Psychology Information ZPID. doi:10.5160/psychdata.zrce18mo99

Chapter 6

Paper 3: Recalling Experiences: Looking at Momentary, Retro- spective and Global Assessments of Relationship Satisfaction



Recalling Experiences: Looking at Momentary, Retrospective and Global Assessments of Relationship Satisfaction

Caroline Zygar-Hoffmann¹ & Felix D. Schönbrodt¹

¹ Ludwig-Maximilians-Universität München

Manuscript published in the Open Access Journal *Collabra:
Psychology*.

Zygar-Hoffmann, C., & Schönbrodt, F. D. (2020). Recalling experiences:
Looking at momentary, retrospective and global assessments of relationship
satisfaction. *Collabra: Psychology*, 6(1), Article 7.
<https://doi.org/10.1525/collabra.278>

Abstract

Relationship satisfaction can be assessed in retrospection, as a global evaluation, or as a momentary state. In two experience sampling studies ($N = 130$, $N = 510$) the specificities of these assessment modalities are examined. We show that 1) compared to other summary statistics like the median, the mean of relationship satisfaction states describes retrospective and global evaluations best (but the difference to some other summary statistics was negligible); 2) retrospection introduces an overestimation of the average annoyance in the relationship reported on a momentary basis, which results in an overall negative mean-level bias for retrospective relationship satisfaction; 3) this bias is most strongly moderated by global relationship satisfaction at the time of retrospection; 4) snapshots of momentary relationship satisfaction get representative of global evaluations after approximately two weeks of sampling. The findings extend the recall bias reported in the literature for retrospection of negative affect to the domain of relationship evaluations and assist researchers in designing efficient experience sampling studies.

Keywords: relationship satisfaction, retrospection, recall, experience sampling method, romantic relationships

Global evaluations of individuals' experiences should correspond to their daily experiences. Fleeson (2001) elaborated on this relationship between global evaluations and momentary behavior in the personality domain and described personality traits as density distributions of personality states. The reasoning that traits reflect to some degree characteristics of the occurrence of corresponding states (such as the amount or intensity) is also common for other psychological constructs, such as affective traits, mood and emotions (Rosenberg, 1998).

States are often assumed to be dynamic and affected by situational influences and must therefore be assessed in the moment, for example with the

Department of Psychology, Ludwig-Maximilians-Universität München. We embrace the values of openness and transparency in science (<http://www.researchtransparency.org/>). We report how we determined our sample size, all data exclusions, all manipulations, and all measures in the study (Simmons, Nelson, & Simonsohn, 2012). The preregistration of Study 1 can be found at <https://osf.io/hafsx/>, the preregistration of Study 2 at <https://osf.io/af4yb/>. Both preregistrations contain additional hypotheses on other research questions than the ones reported here. The Supplemental Materials for this paper and complete codebooks can be found at <https://osf.io/sq7mw/>. The codebooks include all variables of the studies, also those not included in the current paper.

Correspondence concerning this article should be addressed to Caroline Zygar-Hoffmann, Leopoldstr. 13, 80802 München, Germany. E-mail: caroline.zygar@psy.lmu.de

experience sampling method (ESM; Csikszentmihalyi & Larson, 1987). Traits on the other hand are most commonly conceptualized as stable dispositions, typically assessed with self-reports of individuals' global representations of their behaviors and experiences. These trait evaluations have much in common with a third assessment mode: The summative recall of experiences during a certain time period, also called retrospective assessment (e.g., used for the Positive and Negative Affect Schedule, which asks individuals to evaluate their affect during the last day(s), week(s), month(s) or year(s), Watson, Clark, & Tellegen, 1988).

Retrospective assessments can introduce recall biases: For instance, studies find discrepancies between individuals' recall of affective experiences and their momentary report in ESM during that time. More global (trait) evaluations are prone to similar biases as well, as they require to appraise an even wider and more unspecific range of situations and time (Baumert et al., 2017; Reis & Gable, 2000; Robinson & Clore, 2002b).

As a lot of emotional experiences happen within relationships, we explore the correspondence between individuals' state assessments of their relationship satisfaction, measured repeatedly with ESM, and their global as well as retrospective assessment of their relationship satisfaction in two studies. Our aim is to inform researchers about (1) the way ESM data on relationship satisfaction relates to classical measurement tools, by investigating to what extent the average, most intense, or more recent experience corresponds to retrospection and global assessments; (2) the differential validity of retrospective assessments, by investigating what kind of bias in retrospection occurs; (3) the role individual differences have in recalling the past, by investigating the moderation of recall biases by traits, global relationship satisfaction, and other individual or relationship characteristics; (4) the optimal design of ESM studies with high accuracy, by investigating what level of aggregation is sufficient to approach a reliable measurement of the global index.

The Special Case of Relationship Satisfaction

Our study focused on a dyadic setting and the assessments of individuals' relationship satisfaction. While this construct naturally plays a vital role for the study of relationships, it is also of special interest from an assessment perspective. On the one hand, the affective component of relationship satisfaction allows for a comparison with the study of concrete affective experiences, like pain or specific emotions. On the other hand, the construct has trait-like features: It reflects an inter-individual difference, is mainly assessed by asking

individuals to globally evaluate their feelings, behavior and experiences (with regard to their relationships; Fincham & Rogge, 2010) and is related to the average of correspondent everyday states (e.g., Hofmann, Finkel, & Fitzsimons, 2015; Zygar et al., 2018a). Furthermore, global relationship satisfaction typically shows medium to strong stability in couples that do not break up (e.g., $r = .61 - .69$ over two years, which is close to typical personality trait stabilities across the same period of time, Fallis, Rehman, Woody, & Purdon, 2016; McCrae, Bond, Yik, Trapnell, & Paulhus, 1998). Studying the assessment of relationship satisfaction can therefore not only contribute to the understanding of this specific construct, it can also provide insights that might be relevant for the related literature on biases occurring during the assessment of affective experiences and traits more generally.

What Summary Statistic of States Corresponds Best to Retrospection and Global Assessments? (RQ1)

Our first goal was to examine the way ESM assessments relate to classical measurement tools. The distribution of an individual's momentary feelings or behaviors can be summarized across different time periods by various measures, such as the central tendency or extreme values. Which measure best represents what individuals do when they retrospectively assess a time period or globally evaluate their relationship?

For the recall of daily mood, studies found that the peak mood describes retrospection better than or incremental to the average mood (Hedges, Jandorf, & Stone, 1985; Parkinson, Briner, Reynolds, & Totterdell, 1995). This is in line with findings from personality, showing that while the average of personality states is the best indicator for global trait measures, the maximum of the state experience is incrementally relevant (Fleeson & Gallagher, 2009). For the recall of pain and various affective experiences during single, discrete events, a series of studies found that not only the most intense, but also the most recent events are predominant for the evaluation of the experience, termed the peak-and-end rule (see Fredrickson, 2000 for a review). However, this rule seems to have only limited value for multi-episodic events like days, where longer time periods are considered, which are characterized by a mix of events and emotions (Miron-Shatz, 2009).

In sum, previous research found evidence for the informational value of averages, peaks and recent experiences. For relationship satisfaction, we a priori did not have a hypothesis about what summary statistic best describes the retrospective and global assessment. We therefore examined the central

tendency (mean and median), extreme values (90% and 10% quantile), and recency effects (mean during the last week and the last day of the ESM period), contrasted with a primacy effect (mean during the first week).

What Bias Occurs in Retrospection? (RQ2)

Our second goal was to investigate whether individuals are biased in their retrospective assessment of their relationship satisfaction. When it comes to evaluating the convergence of judgments, it is possible to differentiate at least two aspects (see e.g., Fletcher & Kerr, 2010; Neubauer, Scott, Sliwinski, & Smyth, 2019; West & Kenny, 2011): First, mean-level bias (also called directional bias or level convergence), which refers to the sample mean of a judgment being different from the sample mean of another judgment that is used as an external reference category (i.e., as truth criterion). In our case, the external reference is a certain summary of an individual's own repeated assessment of relationship satisfaction with ESM, which is compared to that individual's retrospective assessment. A second aspect that can be considered is tracking accuracy (also called truth force or correspondence convergence), which refers to the actual relationship between the reference category (or truth criterion) and the judgments. In our studies, we investigate tracking accuracy in form of the between-person effect of the aggregated ESM assessments on individuals' retrospective judgments.

In this reasoning, discrepancies between retrospection and mean of ESM states are regarded as systematic recall errors caused during retrospection. However, as already pointed out by others (e.g., Conner & Feldman Barrett, 2012; Feldman Barrett, 1997), it may be that retrospective evaluations are in fact more accurate or have higher validity in some contexts, also because they target all experiences during the examined period, even those moments that were not captured by the ESM surveys. It seems to depend on the type of construct and the type of prediction, whether aggregated ESM states, retrospection or global self-reports are more appropriate to represent meaningful between-person differences (Finnigan & Vazire, 2017; Forbes et al., 2012; Oishi & Sullivan, 2006). For example, in the study by Oishi and Sullivan (2006), daily relationship satisfaction predicted later relationship status better than retrospective evaluations; however, the effect of daily relationship satisfaction was not incremental to global evaluations of relationship satisfaction. Studies applying a more continuous assessment or the Day Reconstruction Method (DRM, Kahneman, Krueger, Schkade, Schwarz, & Stone, 2014) might further help to disentangle which variance in retrospection can and which cannot be

explained by actual experiences (but see Lucas, Wallsworth, Anusic, & Donnellan, 2019 for a critical comparison of ESM and DRM), as well as more studies examining the predictive power of each measure for different outcomes. In a first step in the current paper, however, the goal is to illustrate the degree of convergence between the different assessment modalities of relationship satisfaction. This requires to set one of both measures as reference category; in our case, we decided on the ESM state measures, but the research question could equally be examined using retrospection as reference category.

In the domain of intimate relationships, Fletcher and Kerr (2010) conducted a meta-analysis on the mean-level bias and accuracy of individuals' judgments. They differentiated six judgment categories, of which one dealt with retrospective evaluations of one's own assessments ("memories"). The authors report a positive mean-level bias for this category (i.e., an overestimation of relationship quality during retrospection); however, a closer look at the four studies that were included revealed that these studies dealt with different phenomena pertaining to a different interpretation of the mean-level bias. Specifically, three studies (Karney & Coombs, 2000; Karney & Frye, 2002; Sprecher, 1999) reported a positive mean-level bias of individuals' perception of *change* in relationship quality after time periods of 6 months to 10 years. A biased perception of change may differ from a biased perception of actual past experiences, because – depending on the concurrent assessment – a positively biased perception of change could mean a negatively biased perception of the actual experiences in the past. Indeed, a comparison of the level of relationship quality in retrospection with the *actual* assessment in the past indicates a negative mean-level bias in the studies of Karney and Coombs (2000) and Karney and Frye (2002; see also Holmberg and Holmes, 1994; Sprecher, 1999 did not examine retrospection of actual levels).

The fourth study that was included in the meta-analysis (Oishi & Sullivan, 2006) differed in some aspects from the other studies. First, the authors found a positive mean-level bias in retrospection with regard to actual past aspects of the relationship (i.e., not with regard to changes). Specifically, individuals overestimated the occurrence of partner-related behaviors (positive and negative ones), as well as their satisfaction for specific relationship domains in retrospection. Second, the retrospection occurred directly after a period of 14 days in which individuals rated these aspects of their relationship on a momentary basis. This difference in time between retrospection and experience across the studies included in the meta-analysis might be relevant for the bias that is occurring (see Robinson & Clore, 2002b; Walentynowicz,

Schneider, & Stone, 2018 for effects of short vs. long time periods).

To summarize, the meta-analytic estimate of an overall positive mean-level bias for memories (Fletcher & Kerr, 2010) is a heterogeneous mix of findings which should not be interpreted without further consideration. In Study 1, we explored the mean-level bias of retrospective relationship satisfaction without any hypothesis in mind. Based on preliminary analyses in Study 1, for Study 2 we preregistered that we expect a negative mean-level bias (i.e., an underestimation of relationship satisfaction).

With regard to tracking accuracy, the meta-analysis of Fletcher and Kerr (2010) showed robust, significant and positive effects across all judgment categories. In line with these findings, we preregistered in both studies that we expect a positive association between the average ESM state and retrospection, translating into a positive tracking accuracy.

What Moderates Mean-Level Bias? (RQ3)

A third goal of the current study concerned the exploration of possible moderators of a general mean-level bias. Regarding the retrospection of affective experiences, various moderators were identified in previous research, like personality (Feldman Barrett, 1997; Lay, Gerstorf, Scott, Pauly, & Hoppmann, 2017; Mill, Realo, & Allik, 2016), coping style (Schimmack & Hartmann, 1997), subjective well-being (Diener, Larsen, & Emmons, 1984), gender (Robinson, Johnson, & Shields, 1998), self-esteem (Christensen, Wood, & Feldman Barrett, 2003) or daily tiredness and age (Mill et al., 2016; Neubauer et al., 2020). The accessibility model of Robinson and Clore (2002a) suggests different sources of information individuals use when they report on their emotions. Momentary reports of individuals' emotions are described to be mainly driven by the experiential knowledge in the emotional situation, whereas retrospective reports shift from relying on accessible, episodic memory in short-term retrospection to relying on semantic memory and thereby to stable situation-specific or identity-related beliefs and heuristics in long-term retrospection (see Conner & Feldman Barrett, 2012 for a related account). This would explain why individual characteristics were found to moderate mean-level bias, when these are associated with beliefs about one's experiences and behavior in general (e.g., enhanced levels of remembered negative affect for individuals high in neuroticism, see Feldman Barrett, 1997; Lay et al., 2017; Mill et al., 2016).

Early research examining moderators of bias in the retrospection of relationship feelings indicates that individuals with low trust in their partner

underestimate their own feelings for their partner (Holmberg & Holmes, 1994; see Luchies et al., 2013 for the role of trust in biased memories of the partner). The meta-analysis by Fletcher and Kerr (2010) also looked at moderators of mean-level biases and tracking accuracy. Bearing in mind that this meta-analysis was concerned with other judgment categories than memories as well, their results suggest that relationship quality, relationship length, and gender are important moderators for the mean-level bias observed across these different judgment categories. Specifically, individuals who are globally satisfied with their relationship seem to overall show an especially positive mean-level bias, although this relationship decreases with increasing length of the relationship. Attachment styles are also considered as potential influences (see also Pietromonaco & Feldman Barrett, 1997), which is in line with recent research showing that individuals overestimate their partner's negative emotions when they are high in attachment avoidance (Overall, Fletcher, Simpson, & Fillo, 2015).

Another line of research examined the influence of concurrent experiences on the biases occurring during retrospection. Two studies (Holmberg & Holmes, 1994; McFarland & Ross, 1987) found that relationship feelings during recall have an incremental effect on the retrospective assessment, in the way that the recall was similar to the present evaluation of the relationship (for a similar effect for mood and negative emotions see Chang, Overall, Madden, & Low, 2018; Parkinson et al., 1995). In a longitudinal study covering three decades Karney and Coombs (2000) observed this pattern of consistency of retrospective assessments with current relationship satisfaction in a later stage of the relationship. These findings are in line with a theory by Ross (1989), which states that individuals reconstruct their autobiographical experiences based on their current status and then incorporating implicit theories of the malleability or stability of the experiences at hand. Such expectations may indeed play a role, as a study by Galak and Meyvis (2011) showed that individuals overestimate aversive experiences if they expect them to be repeated in the future.

In our studies, we thus explored individual differences that might invoke situation-specific or identity-related beliefs; global evaluations of the relationship or the partner; objective person and relationship characteristics; attachment styles; and concurrent global evaluations. As the current research focuses on the moderation of mean-level bias, we will shortly report, but not discuss the results concerning a moderation of tracking accuracy.

What Level of Aggregation is Sufficient to Approach a Reliable Measurement of the Global Index? (RQ4)

Our last goal was to explore which number of ESM assessments of relationship satisfaction states account for what amount of variance of a global evaluation of relationship satisfaction. Epstein (1979) investigated a similar question for behavior, studying changes in reliability with an increasing number of daily behavioral assessments. The results showed that it takes around 14 days to achieve a satisfying correlation between behavioral samples of one person. For a time span up to four weeks, we will explore how strongly the association between the ESM assessments and the global index will rise with an increasing number of assessments, depending on the timing of the sampling (e.g., in the morning, evening, or a random survey during the day).

Overview of Studies

For RQ2, the following hypotheses were preregistered : 1) For Study 1 (p.8) and Study 2 (p.41): “Individuals’ relationship satisfaction retrospectively assessed after the experience sampling study is positively related to mean levels of individuals’ state relationship satisfaction during the study (mean of states).” This translates to a positive tracking accuracy. 2) Only for Study 2 (p.41): “Individuals’ relationship satisfaction retrospectively assessed after the experience sampling study is lower than mean levels of individuals’ state relationship satisfaction.” This translates to a negative mean-level bias when regressing the retrospection on the average ESM states. We did not preregister how we were planning to analyze these specific hypotheses, but we preregistered some general exclusion criteria (see Sample), and how to handle multiple operationalizations (see Measures and Table 1). These preregistered decisions and deviations from them are highlighted accordingly in the respective sections. We did not have hypotheses concerning the performance of the different summary statistics in RQ1¹, nor for RQ3 and RQ4, these analyses were exploratory.

Couples were recruited (via social networks, newsletters, flyers, notices

¹We did preregister in both studies that the mean of relationship satisfaction is positively related to retrospection (see tracking accuracy hypothesis described for RQ2) as well as to global relationship satisfaction (see p.9 and p.41). However, our main goal in RQ1 was to descriptively compare the different summary statistics for the prediction of retrospection and global relationship satisfaction, but our preregistrations do not mention other summary statistics than the mean. Hence, even though the preregistered hypotheses also correspond to two analyses reported for RQ1, we refrain to draw special attention to these analyses being preregistered.

at a German university and in Study 2 additionally with a website, and the help of therapists offering couple counseling) separately for two ESM studies with different study periods (14 days in Study 1, 28 days in Study 2). Requirements for participation were the affirmation to be at least 18 years old, to be in a heterosexual relationship with the declared partner, and to individually own an Android or iOS smartphone, which one could use regularly during the day. Participants provided a global evaluation of their relationship satisfaction and a range of other trait measures before they repeatedly rated their state relationship satisfaction five times a day. The studies finished with a retrospective assessment of the study period (and in Study 2 again with a more global evaluation of relationship satisfaction).

All measures were administered in German, if own translations were used, this is indicated accordingly. If not mentioned otherwise, for computation of scales, item responses were averaged. We used R (version 3.5.3, R Core Team, 2018) with the package *dplyr* for data handling (Wickham, François, Henry, & Müller, 2018), and the package *papaja* for manuscript writing (Aust & Barth, 2018). Both studies were part of a project funded by the German Research Foundation, which was approved by the local ethics committee. The data of Study 1 has previously been used by Zygar et al. (2018a), the data of both studies by Pusch, Schönbrodt, Zygar-Hoffmann, and Hagemeyer (2020), as well as Schönbrodt, Zygar-Hoffmann, Nestler, Pusch, and Hagemeyer (2019). The results of these papers overlap with the analyses reported in the current paper only in basic descriptive statistics.²

Study 1: Methods

Detailed Procedure

Couples who signed up for the study could chose a time span of 13.5 hours (starting from 08:00 to 10:30 am, ending from 9:30 pm to midnight³) in which the daily, five ESM surveys were scheduled in a semi-random manner (approximately evenly distributed throughout the day) for a study period of 2 weeks. Next, individuals were invited to answer an online pre-ESM questionnaire on their personal computers (programmed with *formr*, Arslan & Tata, 2016; Arslan, Walther, & Tata, 2020) and received instructions for installing

²Zygar et al. (2018a) also report the result of regressing the global relationship satisfaction evaluation on mean relationship satisfaction states (corresponding to one single coefficient of Table 3), but in that paper the ESM states were z-standardized *before* aggregating them within each person, thereby the result is not equal to the standardized regression coefficient reported in the current manuscript.

³The preregistration contains erroneous time-frames on this matter.

an ESM application on their own smartphones (developed at LMU Munich for Android devices). A personal login-code was assigned to each partner for matching the different data sets and identifying couples.

Right after logging into the ESM application, the questions and survey modalities were explained by written instructions, and the study period with in total 70 ESM surveys started on the day after the login. When a survey became active, individuals were notified by their smartphones and had 45 minutes to answer before the survey timed out. The median time needed to answer the survey was 3.28 minutes (interquartile range = 2.50). The questions were identical in each survey. Both partners were notified at the same time, but were asked to respond to the survey individually without discussing their answers with their partner.

After the ESM period, participants received a link to a post-ESM questionnaire (programmed with *LimeSurvey*, LimesurveyGmbH, 2017) which was to be answered on their personal computers. In this questionnaire individuals could also indicate if they wished to get a report on their answers and receive course credit. When their compliance was at least 80%, participants were also eligible to enter a raffle for a voucher. Due to a technical error, we could not retrieve the exact time difference between the end of the ESM part and the completion of the post-ESM questionnaire, but most participants completed the questionnaire within one to two weeks.

Sample

The sample size in Study 1 was determined by time constraints: As we started data collection in November, we decided to finish it by the Christmas holidays to avoid potential bias during these special days. As one couple started two days later than planned and finished their study during the holidays, we excluded their answers on these days. Two persons participated although they were not in a relationship, so their entire data was excluded. This resulted in data from 152 individuals belonging to 77 couples for the pre-ESM questionnaire (two individuals participated without their partner).

We obtained data from a subset of 130 individuals from 68 couples for the ESM part of the study, as six couples quit after the pre-ESM questionnaire and two couples as well as six individuals answered less than the preregistered threshold of one third of all ESM surveys to be included in the final ESM sample (see p.18 in the preregistration). Compliance for the everyday surveys was on average 84% ($SD = 14\%$). After exclusion of 53 surveys for which par-

ticipants reported that they had talked about their answers with their partner, the total number of (partly) answered measurement points was 7573.⁴

After the ESM study period, 117 individuals completed and one individual started (but did not finish) the post-ESM questionnaire. This sample consists of 66 women (56%), mainly students (83%), not married (97%) and without children (99%). For age and relationship duration, see Table 2, and for more details, see Zygar et al. (2018a).

Measures of Relationship Satisfaction

Global relationship satisfaction (pre-ESM questionnaire). For a global, holistic view on individuals' relationship satisfaction, we used the Couples Satisfaction Index (CSI(16); Funk & Rogge, 2007; Greischel & Johnson, n.d.) and the Positive-Negative Relationship Quality Scale (PNRQ, own translation; Rogge, Fincham, Crasta, & Maniaci, 2016). Whereas the CSI assesses global relationship satisfaction as an unidimensional construct, the PNRQ conceptualizes the evaluation of positive and negative qualities of the relationship as two separate constructs. In both measures, individuals are asked to rate their relationship regarding adjectives, but the CSI uses bipolar Likert scales (e.g., from 0 = *Boring* to 6 = *Interesting*), whereas the PNRQ presents single adjectives (e.g., "pleasant") which are to be evaluated on Likert scales ranging from 1 = *Not at all* to 7 = *Extremely*. The CSI additionally consists of questions such as "In general, how often do you think that things between you and your partner are going well?" with answers on 6- and 7-point Likert scales (see codebook for details). CSI ratings are summed.

State relationship satisfaction (ESM). State relationship satisfaction was assessed with two questions (which we labeled "relationship mood" and "annoyance", see Table 1), with answers given on a continuous slider (without any slider ticks, without any numbers shown, results saved with multiple places after the decimal point, scale from 1 to 7 transformed to a 0-10 scale to match the scale of Study 2; see Schönbrodt et al., 2019 for an analysis of psychometric properties of these items). We considered these items to both reflect state relationship satisfaction, but as a minimum criterion for internal consistency on the between-moments level (also called event-level), we preregistered to only compute a scale if the event-level reliability exceeded .40 (see p.17 in the preregistration). As this was not the case and because the retro-

⁴This number is slightly lower than the one reported in Zygar et al. (2018a), because in that paper we reported the number of measurement points before survey-level exclusions ($n = 53$) and included started surveys without a single answered item ($n = 116$).

spective assessment was only based on the relationship mood item, for Study 1 we only report results for this item.

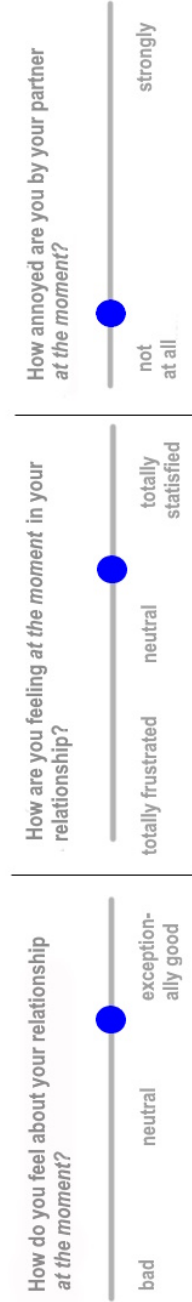
Retrospective relationship satisfaction (post-ESM questionnaire). In the post-ESM questionnaire individuals evaluated the two weeks of the ESM study period on the question “How did you overall feel about your relationship during these two weeks?” with answers on a continuously presented slider ranging from *bad* (=0) to *exceptionally good* (=100; saved as whole numbers, linearly transformed to a 0-10 scale). There were three small differences compared to the state assessment, due to technical limitations (see Figure 1): a) There was no “neutral” label, which was present in the state assessment in the middle of the scale for the relationship mood item, b) The slider started in the middle, whereas no value was preselected in the state assessment, c) Whole numbers were shown as the slider was moved, which was not the case in the state assessment.

Table 1
Slider Items Used for the Assessment of State Relationship Satisfaction

Label	Question	Anchors	ESM	Retro
Item 1: Relationship Mood	How do you feel about your relationship at the moment?	bad (=0) over neutral (=5) to exceptionally good (=10)	Both Studies	Both Studies
Item 2: Annoyance (reverse)	How annoyed are you by your partner at the moment?	not at all (=0) to strongly (=10)	Both Studies	Only Study 2
Item 3: Need Satisfaction	How are you feeling at the moment in your relationship?	frustrated (=0) over neutral (=5) to satisfied (=10)	Only Study 2	Only Study 2
Scale	Average of items			

Note. Experience sampling items used for assessing state relationship satisfaction. The annoyance item was reverse coded for scale calculation. Please note that using only the relationship mood item for the analyses in Study 1 follows our pre-registration. For Study 2, we preregistered to a) use the scale of all ESM relationship satisfaction items, but b) to use only relationship mood when it comes to retrospection; following these decisions would not allow for a commensurable comparison between the ESM measures and retrospection. Therefore, for Study 2, we report the results for all items and the scale separately (see main text for a more detailed description).

Presentation of items during experience sampling in Study 2:



Presentation of items during retrospection in Study 2:

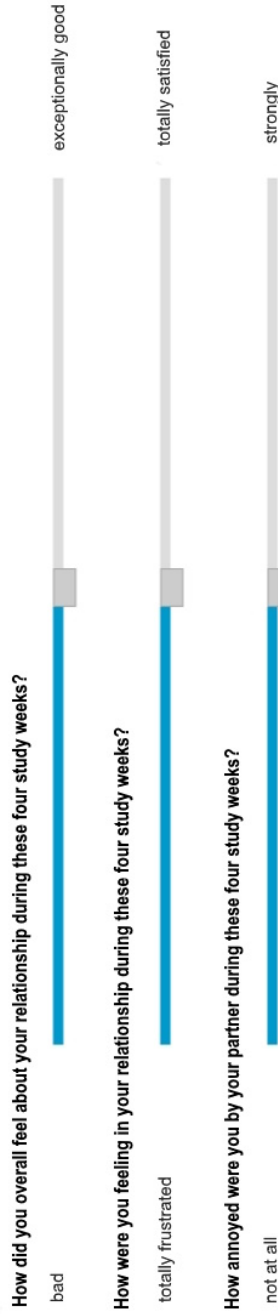


Figure 1. Presentation of items during experience sampling and during retrospection in Study 2 (translated). Presentation differed only slightly in Study 1. Figure available at <https://osf.io/sq7mw/>, under a CC-BY4.0 license.

Potential Moderator Variables

Personality (pre-ESM questionnaire). The Big Five of personality were assessed with the 10-item short version of the Big Five Inventory (Rammstedt & John, 2007). Statements such as “I see myself as someone who gets nervous easily” (Neuroticism) were answered on a Likert scale (1 = *Disagree strongly*, 5 = *Agree strongly*).

Life satisfaction (pre-ESM questionnaire). Individuals’ overall satisfaction with their life was assessed with the Satisfaction with Life Scale (SWLS; Diener, Emmons, Larsen, & Griffin, 1985; Glaesmer, Grande, Braehler, & Roth, 2011). Participants rated five statements like “In most ways my life is close to my ideal.” on a Likert scale (1 = *Strongly disagree*, 7 = *Strongly agree*).

Explicit social desires (pre-ESM questionnaire). Explicit desires for affiliation, being alone and closeness were assessed with the ABC scale of social desires (Hagemeyer, Neyer, Neberich, & Asendorpf, 2013). Participants rated the frequency of 24 experiences related to social desires (e.g., “I enjoy it when my partner wants to be close to me.” for closeness) on Likert scales (1 = *Never*, 7 = *Always*).

Intimacy in the relationship (pre-ESM questionnaire). The amount of intimacy the participants experience in their relationship was measured with two self-constructed items. Individuals rated the frequency of events on questions such as “How often do you tell your partner what you are doing?” on a Likert scale (1 = *Never*, 5 = *Always*).

Further potential moderators (pre- and post-ESM questionnaire). As moderators, we also examined person and relationship characteristics (gender, age, and relationship duration), dominance and autonomy in the relationship, self-reflection and insight (Grant, Franklin, & Langford, 2002), perception of the partner’s explicit social desires (Hagemeyer et al., 2013), explicit motives (UMS-6; Schönbrodt & Gerstenberg, 2012), implicit partner-related needs (PACT; Hagemeyer & Neyer, 2012), and decision-making in the relationship (adaptation of the Allocation of Power in Decision-Making Areas Scale, Bell, 2008; Blood & Wolfe, 1960). As we did not find any effects for these variables as moderators, we refer to the Supplemental Materials and codebook for details.

Study 2: Methods

Detailed Procedure

For Study 2, the general study design was the same as in Study 1, with some exception in details: The ESM period lasted four instead of two weeks (with a total of 140 surveys), and couples were more flexible in their choice of the time span in which the surveys were scheduled. They could choose between a time span of 10 to 16 hours (starting from 07:00 to 10:00 am, ending from 9:00 pm to 11:00 pm) and could block up to two hours per day. A different ESM App was used, namely “Tellmi”, which was developed at LMU Munich not only for Android but also for iOS devices. The questions and survey modalities were explained in a video upon login (instead of text-based in Study 1), and the study period started on the next Monday after the login (instead of on the next day in Study 1). This time, the pre- *and* the post-ESM questionnaire were programmed with *formr* (Arslan et al., 2018; Arslan & Tata, 2017).

The medium time needed to answer the survey was 2.70 minutes (interquartile range = 2.17). The questions were identical for the first four surveys of the day. The evening survey differed with regard to the questions, and had a timeout of five hours instead of 45 minutes, because individuals were instructed to finish it before going to bed.

In addition to the opportunity of receiving a feedback report on their answers as in Study 1, participants were further compensated with course credit or money based on their compliance in the ESM part (up to 170€ per couple). In a follow-up questionnaire a year after the study couples could receive 20€ on top, and participate in a raffle for a voucher.

Sample

Our sample size was constrained by the money available for participant compensation; 576 individuals belonging to 293 couples completed the pre-ESM questionnaire (10 individuals participated without their partner, these could not continue with the ESM part of the study).⁵ We obtained data from a subset of 510 individuals from 259 couples for the ESM part, as six couples quit after the pre-ESM questionnaire and another 18 couples as well as eight

⁵For one couple, we observed an inconsistency in the gender both partners indicated in the pre-ESM compared to the follow-up-questionnaire one year later. We did not exclude this couple from our main analyses, but as the inconsistency might point to careless responding, we report in the Supplemental Materials how minor results for RQ3 change when excluding this couple from the analyses. For all other RQs the pattern of results does not change when excluding their data.

individuals quit during the ESM part or answered less than the preregistered threshold of one third of all ESM surveys⁶ to be included in the final ESM sample (after survey-level exclusions). Compliance for the everyday surveys of the remaining sample was on average 88% ($SD = 12\%$). One couple changed time zone during the study but the survey timing did not adjust to the time transition, so in total 26 surveys (0.04%) were answered during the night and were excluded. As preregistered (see p.59), we further excluded 171 surveys (0.24%) where individuals reported that they had talked about their answers with their partner and additional 1855 entries (2.58%) because of an answering time of less than 60 seconds. In total after all exclusions, 60942 (partly) answered measurement points remained.

After the ESM study period, 508 individuals completed the post-ESM questionnaire. However, we excluded the answers of 22 of these individuals for the retrospective assessment, because of apparently low quality data⁷: These individuals either did not change the default values that were preselected on all sliders ($n = 12$) or probably overlooked the reverse coding of the annoyance item and were thus identified as outliers (Cook's Distance $> 2SD$, $n = 10$).⁸ This resulted in a final sample of 486 individuals, consisting of 249 women (51%), mainly non-students (71%) without children (68%), with roughly one third of them married (32%); for age and relationship duration, see Table 2.

Measures of Relationship Satisfaction

Global relationship satisfaction (pre-ESM questionnaire and post-ESM questionnaire). We used the same measures as in Study 1 (CSI(16); Funk & Rogge, 2007, and PNRQ; Rogge et al., 2016), but also applied them in the post-ESM questionnaire, so we could examine the influence of concurrent relationship evaluations on the retrospective assessment.

State relationship satisfaction (ESM). To achieve a more reliable assessment of state relationship satisfaction, we complemented the two items from Study 1 (but on a scale from 0-10) with an additional question with identical slider properties (which we called “need satisfaction”, see Table 1

⁶In our preregistration we erroneously stated that less than 33% of 140 would be less than 24 rather than the actual 47 surveys (see p.59).

⁷For analyses with data from retrospection, we describe if the pattern of results changes when not excluding this data.

⁸In total, 12 individuals had a Cook's Distance of $> 2SD$. However, two of these individuals were not treated as outliers, as they were just very unhappy with their relationship (and thus different than the majority of the sample), but still consistent in their answers, in contrast to the other 10 individuals who indicated high positive relationship mood and need satisfaction, but also high annoyance.

and Schönbrodt et al., 2019). Again, as a minimum criterion for internal consistency, we preregistered to compute a scale if the event-level reliability exceeded .40, which was the case (see p.42 in the preregistration).

Retrospective relationship satisfaction (post-ESM questionnaire). In the post-ESM questionnaire individuals were asked to evaluate the study period on the questions presented in Figure 1 with answers on a continuously presented slider with the same labels as for the state assessments (scale from 1 to 100, saved as whole numbers, again linearly transformed to a 0-10 scale). In contrast to Study 1, no numbers were shown as the slider was moved in the retrospective assessment, just as it was the case in the state assessment. Yet, two small differences compared to the state assessments remained (see Figure 1): As in Study 1, the “neutral” label was not shown in the retrospective assessment (which was present in the state assessment in the middle of the scale for the relationship mood and need satisfaction items), and the slider started in the middle of the scale instead of no default value being pre-selected.

Although for the retrospective assessment we had questions that were based on all three items, we preregistered to only use the relationship mood item (see p.43 in the preregistration). To deal transparently with these inconsistencies in the preregistration regarding scale calculation of state and retrospective relationship satisfaction, for Study 2 we report the results for all three items and for the scale of all items separately, and correct accordingly for multiple comparisons. Next to providing transparency, this detailed presentation of the results a) allows to illustrate the cumulative evidence across both studies for the relationship mood item, which is the only item that was assessed both in Study 1 and Study 2 both in ESM and retrospection (see Table 1); b) informs which items are more susceptible to bias than others, therefore driving potential biases observed for the scale of all items.

Potential Moderator Variables

We assessed the same moderator variables as in Study 1, but personality was assessed with another measure, attachment styles were additionally included and delay between the ESM period and retrospection was documented.

Personality (pre-ESM questionnaire). The Big Five were measured with the 15 short-item scale developed for the Socio-Economic Panel survey (BFI-S; Gerlitz & Schupp, 2005). Participants rated statements such as “I see myself as someone who does a thorough job” (Conscientiousness) on a Likert scale (1 = *Strongly disagree*, 7 = *Strongly agree*).

Attachment styles (pre-ESM questionnaire). Anxiety and Avoidance in adult relationships were measured with the Experiences in Close Relationships Questionnaire (Ehrental, Dinger, Lamla, Funken, & Schauenburg, 2009). Thirty-six statements such as “I often worry that my partner doesn’t really love me.” (Anxiety) were answered on a Likert scale (1 = *Strongly disagree*, 7 = *Strongly agree*).

Analysis Plan of Both Studies

In both studies state relationship satisfaction was measured repeatedly at the individual level, with individuals belonging to a specific dyad. To account for the resulting nonindependence of the data, we applied multilevel regression models (MLMs; using the packages *lme4* and *lmerTest*, Bates, Mächler, Bolker, & Walker, 2015; Kuznetsova, Brockhoff, & Christensen, 2017). In all models we entered a gender contrast as fixed effect (-1 = women, 1 = men, i.e., regression coefficients of other variables in the models can be interpreted as the average effect across both genders).⁹ We aggregated the ESM data within individuals during preprocessing, hence individuals’ summary of their ESM answers were on level 1 nested in couples on level 2. This pre-aggregation of ESM data was necessary to be able to compare summary statistics (for RQ1), and to be able to compute a slope while accounting for the nonindependence of the dyad data (for RQ2 and RQ3).

The relationship satisfaction variables (global / retrospective / aggregated state) were *z*-standardized for RQ1 to achieve a standardized regression coefficient, using the grand-mean and standard deviation across both genders. For the investigation of bias and accuracy (RQ2 and RQ3), the retrospective assessments and the aggregated ESM answers were grand-mean centered instead, using the grand-mean of the ESM measures (see West & Kenny, 2011): This results in both measures being centered on the variable that is conceptualized as the “truth” (i.e., the ESM answers). As both measures were transformed to the same metric, a mean-level bias would show itself in an intercept different from zero when regressing the retrospective assessment on the ESM answers. The sign of the intercept indicates whether the retrospective assessment is on average an under- or overestimation of the averaged feelings reported during ESM. The coefficient of the aggregated ESM measure shows the tracking ac-

⁹In our preregistrations we stated we would use two-intercept models as default (i.e., separate intercepts for men and women, see p.18 and p.5). However, in the current case, using a gender contrast variable leads to a more meaningful interpretation of the intercept (mean-level bias across both genders, instead of a mean-level bias separately for men and women).

curacy, with a value of one representing perfect accuracy: An increase of one scale point in the aggregated ESM measure would then result in an increase of one scale point in the retrospective assessment. Entering moderators as main effects reveals whether individuals with a high expression of the moderator have an even higher or lower bias (i.e., conditional on the aggregated states as predictor, the main effect of the moderator variable increases or lowers the intercept). An interaction of the moderator variable with the aggregated ESM measure indicates whether tracking accuracy is decreased or increased for certain groups of individuals. The model including a moderator (i.e., for RQ3) is specified as follows (RQ2 uses the same model without all terms involving the moderator variable):

$$\begin{aligned}
 \textit{Retrospection}(GMC_{ESM})_{ij} = & \\
 & (\gamma_{00} \\
 & + \gamma_{10}\textit{GenderContrast}_{ij} + \gamma_{20}\textit{Moderator}(z)_{ij} \\
 + \gamma_{30}\textit{MeanESM}(GMC_{ESM})_{ij} + \gamma_{40}\textit{Moderator}(z)_{ij} \times & \textit{MeanESM}(GMC_{ESM})_{ij} \\
 & + (u_{0j} + r_{ij})
 \end{aligned}$$

with GMC_{ESM} = grand-mean centered on the ESM-mean, i = person-specific index, j = couple-specific index, γ = fixed effect, (z) = z -standardized, u = random intercept, r = error term. This translates into the following between-person interpretation of the estimates:

$$\begin{aligned}
 \textit{Retrospection} = & \\
 (\textit{Mean-LevelBias} & \\
 + \textit{Moderation of Bias by Gender} + \textit{Moderation of Bias by Moderator} & \\
 + \textit{TrackingAccuracy} + \textit{Moderation of Accuracy by Moderator}) & \\
 + (\textit{RandomIntercept for Each Couple} + \textit{Error}) &
 \end{aligned}$$

For all models, we report the marginal R^2 as an effect size, representing the explained variance by the fixed effects ($R^2_{GLMM(m)}$ from the *MuMIn* package, Johnson, 2014; Barton, 2018; Nakagawa & Schielzeth, 2013). When making multiple tests for a single analysis question (i.e., due to multiple items, summary statistics, moderators), we controlled the false discovery rate (FDR) at $\alpha = 5\%$ (two-tailed) with the Benjamini-Hochberg (BH) correction of the p -values (Benjamini & Hochberg, 1995) implemented in the *stats* package (R Core Team, 2018).¹⁰

¹⁰In the preregistration of Study 2 we mention “For controlling the false-discovery-rate

Results of Both Studies

Table 2 shows the descriptive statistics for both studies. Correlations and a complete description of the parameter estimates, confidence intervals, and effect sizes for all results can be found in the Supplemental Materials.

Table 2
Descriptive Statistics

Variables	Study 1				Study 2			
	α/ω	M	SD	Range	α/ω	M	SD	Range
Age in years	-	22.44	4.29	18 to 40	-	31.29	9.49	18 to 68
Relationship duration in years	-	2.30	1.96	0 to 8	-	6.35	6.35	0.2 to 33.2
Global RS: CSI (Pre-ESM)	0.92	66.58	10.55	32 to 81	0.96	64.07	13.51	4 to 81
Concurrent RS: CSI (Post-ESM)	-	-	-	-	0.96	62.33	14.77	5 to 81
Global RS: PRQ (Pre-ESM)	0.92	5.83	0.98	1.5 to 7	0.91	5.74	0.88	2.4 to 7
Concurrent RS: PRQ (Post-ESM)	-	-	-	-	0.94	5.40	1.09	1 to 7
Global RS: NRQ (Pre-ESM)	0.91	1.86	1.05	1 to 5.9	0.94	1.84	1.07	1 to 7
Concurrent RS: NRQ (Post-ESM)	-	-	-	-	0.93	1.73	0.90	1 to 6.8
Mean RS state: Item 1	0.95	7.03	1.14	3.3 to 9.8	0.98	7.25	1.34	2.7 to 10
Retro of RS: Item 1	-	6.83	1.78	1.2 to 10	-	7.23	1.91	0 to 10
Mean RS state: Item 2 (reverse)	0.93	8.95	0.95	4.4 to 9.9	0.96	9.15	0.94	4.8 to 10
Retro of RS: Item 2 (reverse)	-	-	-	-	-	8.28	2.17	0 to 10
Mean RS state: Item 3	-	-	-	-	0.98	7.20	1.38	1.5 to 10
Retro of RS: Item 3	-	-	-	-	-	7.21	2.07	0 to 10
Mean RS state: Scale	-	-	-	-	0.97	7.86	1.11	3.1 to 10
Retro of RS: Scale	-	-	-	-	0.85	7.57	1.79	0.4 to 10
Personality: Conscientiousness	0.49	3.50	0.83	1.5 to 5	0.69	5.25	1.09	1 to 7
Personality: Neuroticism	0.62	2.89	1.10	1 to 5	0.68	4.14	1.34	1 to 7
Satisfaction with life	0.88	5.50	1.07	2 to 7	0.87	5.16	1.16	1.2 to 7
Explicit desire for being alone	0.84	4.07	0.94	1.8 to 6.4	0.85	4.21	0.98	1 to 6.9
Explicit desire for closeness	0.86	6.18	0.67	3.5 to 7	0.90	6.03	0.76	1.4 to 7
Intimacy in the relationship	0.79	4.12	0.78	1.5 to 5	0.82	3.78	0.89	1.5 to 5
AS: Anxiety in the relationship	-	-	-	-	0.90	2.81	1.08	1 to 6.2
AS: Avoidance in the relationship	-	-	-	-	0.89	2.22	0.85	1 to 6.4
Delay of retrospection in days	-	-	-	-	-	2.01	4.05	0 to 63

Note. N (Study 1) = 118-152, N (Study 2) = 486-576, RS = Relationship Satisfaction, CSI = Couples Satisfaction Index, PRQ = Positive Relationship Quality, NRQ = Negative Relationship Quality, Retro = Retrospection, Item 1 = Relationship mood, Item 2 = Annoyance (reverse coded), Item 3 = Need satisfaction, AS = Attachment Style. For state measures the between-person reliability is reported, for scales consisting of only two items Cronbach's α is reported, and for all other measures McDonald's ω_{total} is reported.

(FDR) at 5 % we will apply the Benjamini-Hochberg procedure [...]” (p.6), but we also state that “For exploratory analyses, we consider effects noteworthy when $p < .01$ and $\beta \geq .05$ (for additional moderations of hypotheses) or $\beta \geq .10$ (for additional main effects)” (p.6). Both procedures lead to reporting roughly equivalent exploratory effects in the current paper. We decided on the FDR procedure, as the number of effects to control for could be determined (number of analyses = number of summary statistics or moderators multiplied by the number of items plus the scale; separately for mean-level bias and tracking accuracy) and the other procedure can more easily be applied by the readers themselves.

What Summary Statistic Corresponds Best to Retrospection and Global Assessments? (RQ1)

Table 3 shows the standardized regression coefficients for several ESM summary statistics predicting retrospection after two weeks (Study 1) and four weeks (Study 2) of ESM, separately for the different relationship satisfaction items. For both studies and all items, the best prediction was achieved by the mean of the whole study period, while the mean of the last day and the 90th quantile of the distribution performed the worst. Overall, the highest associations were found for the mean of the scale of all three ESM items predicting the scale of all three retrospective assessments ($\beta = 0.75$), and for the mean of need satisfaction predicting retrospection of this item ($\beta = 0.74$).

The same analysis for the prediction of a global relationship satisfaction measure (the CSI) instead of the retrospective assessment is also shown in Table 3 (for the prediction of PRQ and NRQ see Supplemental Materials). The mean of the last week, of the last day and of the first week were not entered as predictors, as they provide no special meaning to the global evaluation, which was assessed before the ESM part. Again, the mean was the best predictor in all cases. Other summary statistics performed equally well in some cases, but without a systematic pattern. The associations were highest when the mean of the scale, or the mean of need satisfaction (item 3) across four weeks predicted the CSI ($\beta_{Scale} = 0.59$, $\beta_{NeedSatisfaction} = 0.58$).

We additionally checked whether other summary statistics next to the mean provided an incremental contribution to the prediction of retrospection (see Table 4). This was not the case in Study 1 (we controlled the FDR for all incremental effects across studies, all BH-corrected ps of the model comparisons > 0.16). In Study 2, all summary statistics except the 90th quantile and the mean of the first week made incremental contributions for the prediction of retrospection of relationship mood and the scale. For the annoyance item both the 10th and the 90th quantile – but no other summary statistic – had incremental effects. As annoyance was reverse coded, the 10th quantile represents a high level of annoyance, whereas the 90th quantile represents a low level of annoyance. For need satisfaction only the summaries of the end of the study (i.e., mean of the last week and mean of the last day) had additional relevance. Overall the incremental contributions were small (additional explained variance $< 3\%$, compared to baseline explained variance of the mean

as single predictor between 30% and 57%). Whereas the coefficients of the 10th quantile and the means of the last day/week were positive, the median and the 90th quantile had negative coefficients.

Table 3
Prediction of Retrospective and Global Assessment by Different Summary Statistics of ESM Relationship Satisfaction States (All z-Standardized)

Retrospection by summary	Study 1		Study 2		
	Item 1	Item 1	Item 2 (r)	Item 3	Scale
Mean	0.55	0.66	0.61	0.74	0.75
Mean last week	0.54	0.65	0.55	0.72	0.72
10th quantile	0.53	0.63	0.59	0.68	0.71
Median	0.52	0.61	0.49	0.70	0.70
Mean first week	0.48	0.56	0.52	0.64	0.65
Mean last day	0.43	0.57	0.41	0.62	0.59
90th quantile	0.40	0.54	0.28	0.61	0.60
CSI by summary	Item 1	Item 1	Item 2 (r)	Item 3	Scale
Mean	0.38	0.52	0.44	0.58	0.59
10th quantile	0.38	0.50	0.41	0.52	0.55
Median	0.33	0.46	0.35	0.52	0.54
90th quantile	0.30	0.50	0.25	0.55	0.55

Note. N (Study 1) = 115-130, N (Study 2) = 475-510. Item 1 = Relationship mood, Item 2 = Annoyance (reverse coded), Item 3 = Need satisfaction. CSI = Couples Satisfaction Index assessed before the ESM period. Rows ordered by size of average coefficient across all items. The strongest effect is printed in bold.

Table 4
Prediction of Retrospection by Relationship Satisfaction States: Incremental Contributions beyond the Mean

	Study 1			Study 2		
	Relationship mood	Relationship mood	Annoyance (reverse)	Need satisfaction	Scale	
Baseline R ²	30.21%	43.54%	36.95%	56.65%	56.65%	
Mean	b = 0.63, p = .039	b = 1.25, p < .001	b = 0.77, p < .001	b = 0.98, p < .001	b = 1.10, p < .001	
Median	b = -0.08, p = .793	b = -0.60, p < .001	b = -0.18, p = .037	b = -0.24, p = .091	b = -0.35, p = .014	
Δ R ²	0.02%	1.96%	0.70%	0.46%	0.86%	
Mean	b = 0.42, p = .008	b = 0.42, p < .001	b = 0.44, p < .001	b = 0.63, p < .001	b = 0.58, p < .001	
10q	b = 0.15, p = .308	b = 0.27, p < .001	b = 0.20, p = .018	b = 0.12, p = .059	b = 0.19, p = .006	
Δ R ²	1.13%	1.93%	0.93%	0.52%	1.01%	
Mean	b = 0.77, p < .001	b = 0.67, p < .001	b = 0.71, p < .001	b = 0.79, p < .001	b = 0.82, p < .001	
90q	b = -0.26, p = .071	b = -0.02, p = .762	b = -0.16, p = .001	b = -0.05, p = .366	b = -0.09, p = .140	
Δ R ²	2.71%	-0.05%	1.48%	0.05%	0.30%	
Mean	b = 0.54, p < .001	b = 0.54, p < .001	b = 0.58, p < .001	b = 0.63, p < .001	b = 0.66, p < .001	
Mean last day ¹	b = 0.00, p = .970	b = 0.16, p = .003	b = 0.07, p = .115	b = 0.16, p < .001	b = 0.13, p = .003	
Δ R ²	0.17%	1.24%	0.14%	1.22%	0.87%	
Mean	b = 0.32, p = .193	b = 0.39, p < .001	b = 0.53, p < .001	b = 0.49, p < .001	b = 0.51, p < .001	
Mean last week ¹	b = 0.24, p = .327	b = 0.29, p = .002	b = 0.09, p = .241	b = 0.27, p = .001	b = 0.25, p = .002	
Δ R ²	0.61%	1.24%	0.10%	0.81%	0.86%	
Mean	b = 0.85, p < .001	b = 0.73, p < .001	b = 0.56, p < .001	b = 0.86, p < .001	b = 0.74, p < .001	
Mean first week	b = -0.32, p = .170	b = -0.08, p = .321	b = 0.06, p = .394	b = -0.13, p = .083	b = 0.00, p = .971	
Δ R ²	1.23%	0.09%	0.04%	0.24%	-0.06%	

Note. N (Study 1) = 118, N (Study 2) = 486. Baseline R² is the explained variance by the mean as fixed effect. Δ R² is the incremental explained variance by the additional summary statistic, compared to the model including only the mean as predictor. ¹ Due to missing data on the last day or last week for some persons, these models used data from only 115 participants in Study 1 (for models with the last day) and from 475/485 participants in Study 2 (for models with the last day/last week); the baseline R² differs slightly on this data. Bold values of the additional summary statistics indicate that a model without this variable fits the data significantly worse after controlling the false discovery rate at α = 5% (two-tailed) for all model comparisons. The predictors and the criterion in the models are z-standardized. The 10th quantile represents an especially negative relationship evaluation for all items (as annoyance is reverse coded); the 90th quantile represents an especially positive relationship evaluation for all items. Please note that mean and median very highly correlated, leading to Variance Inflation Factors (VIFs) between 5 and 23; all other VIFs were < 10.

What Bias Occurs in Retrospection? (RQ2)

Given that the mean was the best measure for predicting retrospection, for investigating mean-level bias and tracking accuracy, we regressed the retrospective assessment on the mean of relationship satisfaction states. Table 5 shows the results for the different items, including a meta-analytical p -value for the relationship mood item (calculated with the *metap* package, Dewey, 2018), to synthesize the results of both studies.

There was no significant mean-level bias for the two positively framed items (relationship mood and need satisfaction). However, for the negatively framed annoyance item and for the scale out of all three items, a negative mean-level bias emerged.¹¹ It is important to note that the annoyance item was reverse coded, therefore the negative coefficient of the mean-level bias indicates that individuals on average *overestimate* the amount of them having been annoyed by their partner during the study.¹² This bias is still present when computing the scale that includes annoyance next to relationship mood and need satisfaction. In consequence, individuals' overall relationship satisfaction score is lower in retrospection than the average ESM report, driven by a higher level of remembered annoyance.¹³

¹¹We also explored the results for regressing the retrospective assessment on the median, the mean of the last week, of the last day, and of the first week, controlling the FDR for the according number of tests. The reported mean-level bias for the annoyance item and the scale replicated for all of these summary statistics, and even extended to the other two items in some cases.

¹²Therefore, this result could have also been labeled as a positive mean-level bias of annoyance, in the sense of an overestimating of the variable of interest. However, to avoid confusion and to consistently refer to “negative mean-level bias” as assessing the relationship in retrospection worse than what was indicated by the ESM reports, we label the difference that occurred in retrospection of (reverse-scored) annoyance as negative mean-level bias as well (see Fletcher & Kerr, 2010, who also use these terms accordingly).

¹³As all relationship satisfaction variables had skewed distributions, all of our models had an overall poor fit. We reran the analyses of RQ2 as Bayesian MLMs in the *brms* package (Bürkner, 2017) with default priors, but specifying skewed normal distributions with an inverse and a log link. These alternative models fitted better, although still not good in case of annoyance. The results were consistent with those reported here. When specifying a log link with the skewed normal distribution, additionally a negative mean-level bias for the relationship mood and the need satisfaction item emerged (with the 95%-HDI of the intercepts excluding zero).

Table 5
Prediction of Retrospective Assessment by Mean of ESM Relationship Satisfaction States (with Common Zero)

	Relationship mood (S1)			Relationship mood (S2)			Relationship mood (S1+S2)		
	β	95% CI	<i>p</i>	β	95% CI	<i>p</i>	β	95% CI	meta <i>p</i>
Intercept (Mean-level Bias)	-0.19	[-0.49,0.12]	.237	-0.01	[-0.15,0.13]	.855			.736
Gender	0.04	[-0.19,0.26]	.737	-0.05	[-0.16,0.07]	.411			.532
ESM Mean (Tracking Accuracy)	0.86	[0.62,1.10]	.247	0.93	[0.83,1.03]	.175			.127
R ² _{GLMM(m)}	.302			.435					
	Annoyance (reverse) (S2)			Need satisfaction (S2)			Scale (S2)		
	β	95% CI	<i>p</i>	β	95% CI	<i>p</i>	β	95% CI	<i>p</i>
Intercept (Mean-level Bias)	-0.87	[-1.03,-0.71]	< .001	0.01	[-0.12,0.14]	.864	-0.29	[-0.40,-0.18]	< .001
Gender	0.13	[-0.02,0.28]	.087	-0.09	[-0.19,0.02]	.114	-0.01	[-0.10,0.09]	.900
ESM Mean (Tracking Accuracy)	1.39	[1.22,1.55]	< .001	1.12	[1.03,1.21]	.011	1.20	[1.10,1.29]	< .001
R ² _{GLMM(m)}	.369			.566			.567		

Note. S1 = Study 1, S2 = Study 2, Gender = Contrast variable with -1 = women and 1 = men, CI = Confidence Interval. N (Study 1) = 118, N (Study 2) = 486. Retrospective assessment and mean of states were centered on the grand-mean of the mean of states. The intercept of the models indicate whether mean-level bias is present, the slope of the ESM mean state indicates whether the tracking accuracy differs from 1 (likewise, we tested whether the slope differs from 1, i.e., the *p*-value corresponds to the H0: $\beta = 1$). All significant *p*-values remain significant after controlling the false discovery rate at $\alpha = 5\%$ (two-tailed).

Further, the results showed a tracking accuracy of greater than one for the annoyance and need satisfaction item and for the scale. This indicates that experienced annoyance captured by the ESM assessments is amplified during retrospection: High levels of being annoyed are perceived as having been even higher, reinforcing the negative mean-level bias, and leading to an overall more diverging perception. For low annoyance, this effect counterbalances the mean-level bias and results in an overall more similar perception (see Figure 2).¹⁴

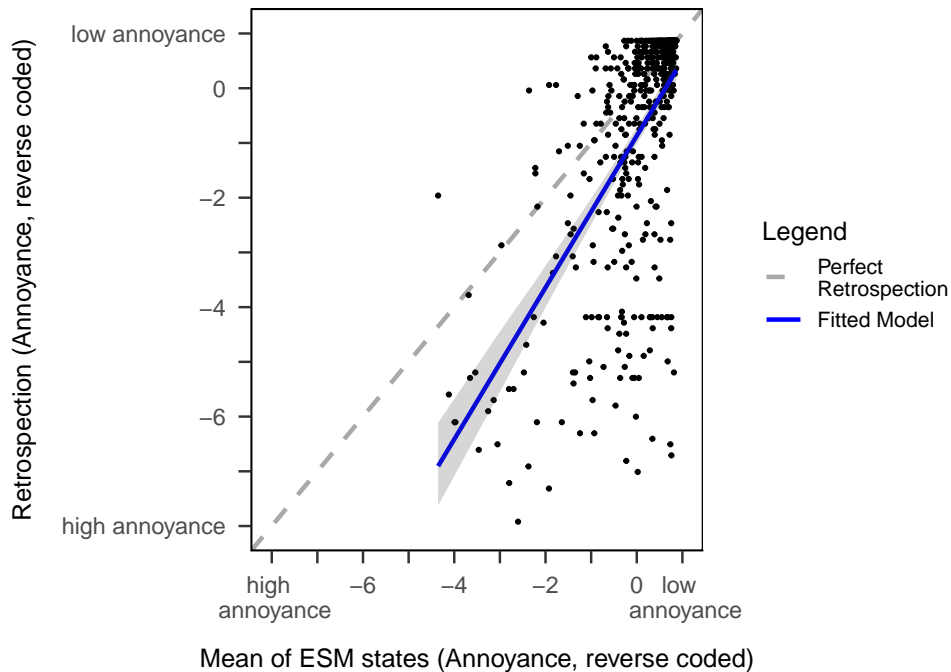


Figure 2. Prediction of retrospective assessment by mean of ESM relationship satisfaction states for the reverse coded annoyance item (with common zero). High values indicate low annoyance. Uncertainty band was calculated with the *merTools* package (Knowles & Frederick, 2018). Figure created with the *ggplot2* package (Wickham, 2016), available at <https://osf.io/sq7mw/>, under a CC-BY4.0 license.

¹⁴When not excluding low quality responses (see Sample) the tracking accuracy of the scale is no longer significantly different from one.

What Moderates Mean-Level Bias? (RQ3)

We added moderators of mean-level bias and tracking accuracy to the models of RQ2, so that retrospection was predicted by an intercept (indicating potential mean-level bias), a main effect of the mean ESM state (indicating potential tracking accuracy), a main effect of a moderator (indicating a potential moderation of the mean-level bias) and the interaction between mean ESM state and the moderator (indicating a potential moderation of the tracking accuracy). We report the results of those moderators that had a significant main effect for at least one item or the scale after controlling the FDR.

Figure 3 illustrates the pattern of main effects for global relationship satisfaction as a moderator: Independent of the item being considered, global relationship satisfaction concurrently assessed with retrospection turned out to be a central moderator of the mean-level bias in both studies, irrelevant of the measure being the CSI or the more specific PNRQ scales. The coefficients indicate that individuals who are globally more satisfied with their relationship during retrospection tend to less strongly underestimate or even overestimate their relationship satisfaction as reported during ESM. In case of annoyance, due to the reverse coding, the coefficients indicate that globally satisfied individuals less strongly overestimate their level of annoyance. Even though the overall mean-level bias for the relationship mood and need satisfaction items was not significantly different from zero (see RQ2 and “Intercept” column in Figure 3), the models with these items still showed the moderating effect by the global measure.

Global relationship satisfaction assessed before the evaluated ESM period had similar, but considerably lower and more inconsistent effects: The aforementioned moderation was present for all items except need satisfaction when looking at the CSI; the moderation by the PRQ was only significant for the annoyance and the need satisfaction item; and there was no significant moderation by the NRQ.

Results of models with different moderators, separately for different items

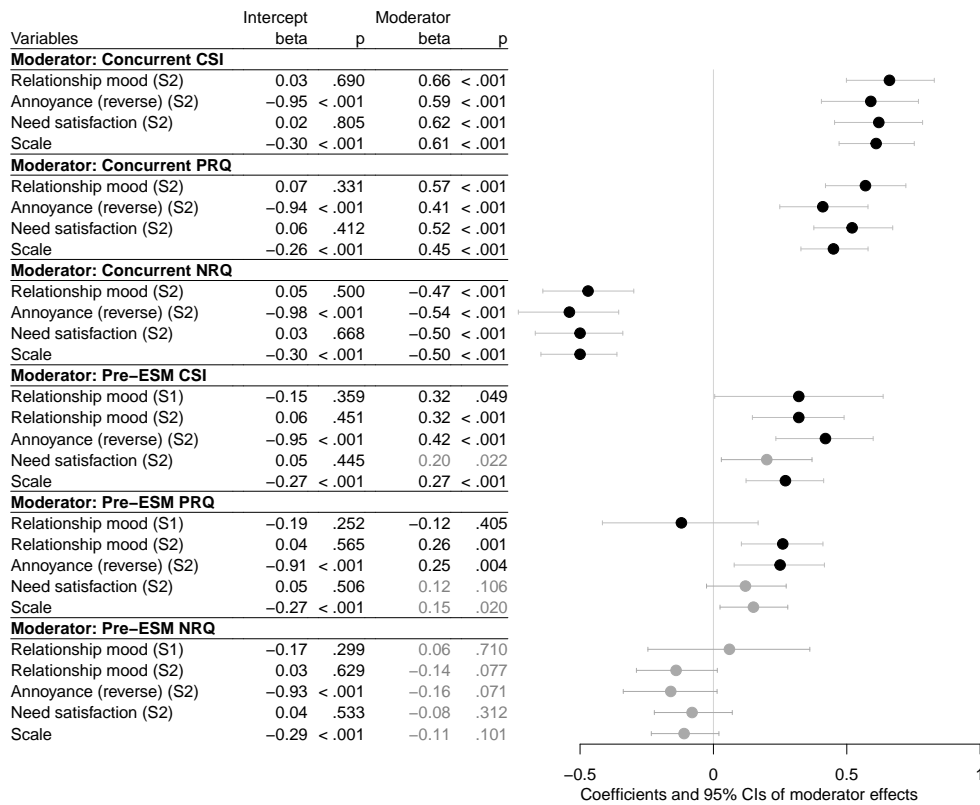


Figure 3. Moderation of mean-level bias by global relationship satisfaction (i.e., main effects of global relationship satisfaction concurrently assessed and assessed “pre-esm” = before the experience sampling study) for different relationship satisfaction items. The interaction between moderator and mean relationship satisfaction states (i.e., the moderation of tracking accuracy) is included in the models, but not reported here. S1 = Study 1, S2 = Study 2. N (Study 1) = 118, N (Study 2) = 486. Moderator effects that were significant after controlling the false discovery rate at $\alpha = 5\%$ (two-tailed) are displayed in black (for relationship mood based on a meta p -value of both studies), all other moderator effects are displayed in grey. Figure created with the *forestplot* package (Gordon & Lumley, 2017), available at <https://osf.io/sq7mw/>, under a CC-BY4.0 license.

As shown in Figure 4, life satisfaction had likewise a positive moderating effect for all items, indicating that individuals who are globally happy with their life show less of an overall underestimation of their relationship satisfaction, resulting from a less strongly overestimation of annoyance and some overestimation of relationship mood and need satisfaction. In contrast, anxious and avoidant attachment, neuroticism, and the explicit desire for being alone had negative moderating effects on some items. Individuals with a high expression of these traits underestimate their relationship satisfaction in some aspects even stronger.

There were some other moderators that only influenced the bias of specifically the annoyance item: The explicit desire for closeness, perceived intimacy, and conscientiousness all had positive effects, counterbalancing the overall neg-

ative bias in the evaluation of annoyance (i.e., resulting in a less strongly overestimation for those scoring high on these traits; see Figure 4).¹⁵

The result pattern suggests that all moderators with positive valence show a positive moderating effect, and those with negative valence a negative effect. Consequently, these findings could result from an overall latent factor reflecting positive compared to negative views about oneself / one's life / one's relationship or more generally a methodological artefact of social desirability. As a first approach to this alternative explanation, we fitted a bifactor model (see e.g., Biderman, Nguyen, Cunningham, & Ghorbani, 2011; Reise, 2012) with structural equation modeling (using *lavaan*, Rosseel, 2012) on all self-report items assessed during the pre-ESM questionnaire in Study 2: In this model all items load on their respective scales (with correlated latent factors of all these scales), as well as on a general factor (orthogonal to the other latent factors). The general factor that resulted from this analysis seems to capture indeed a general positivity or negativity in answering the items (i.e., all items from constructs mirroring positive feelings or experiences loaded positively, irrespective of them being reverse scored or not; items from constructs reflecting negative feelings or experiences loaded negatively; model fit and all factor loadings are presented in the Supplemental Materials). In a second step, we extracted regression factor scores on this latent factor for each person, and added them as additional manifest moderator variable to our analyses (see Figure 4): The results show that this factor moderates the mean-level bias of relationship mood, annoyance, and the scale, but not of need satisfaction.

To assess whether the specific moderators explain variance beyond this general positivity factor, we repeated all analyses with this factor included as covariate (as main effect and in interaction with the averaged ESM states). Robust to adding this control variable were the moderation effects of all relationship satisfaction measures concurrently assessed; of the CSI assessed before the ESM study period; of life satisfaction on all but the relationship mood item; of anxious attachment and conscientiousness on the annoyance item (uncorrected p -values of these moderators $< .05$). Not robust were the effects of the PRQ measured before the study period; of life satisfaction on the relationship mood item; of anxious attachment on the scale; of avoidant attachment and neuroticism; and of intimacy, the explicit desires for closeness and for being

¹⁵When not excluding low quality responses (see Sample) the moderation for the annoyance item by life satisfaction, avoidant attachment, neuroticism, and conscientiousness was no longer significant. PRQ assessed before ESM and anxious attachment are no longer significant moderators for any item. Instead, a significant moderation by gender for the need satisfaction item indicates an underestimation by men and an overestimation by women, and self-reflection shows a significant positive moderation for the annoyance item.

alone on the annoyance item.

The tracking accuracy was moderated only by the intimacy in the relationship and concurrent negative relationship quality for some items (see Supplemental Materials).

Results of models with different moderators, separately for different items

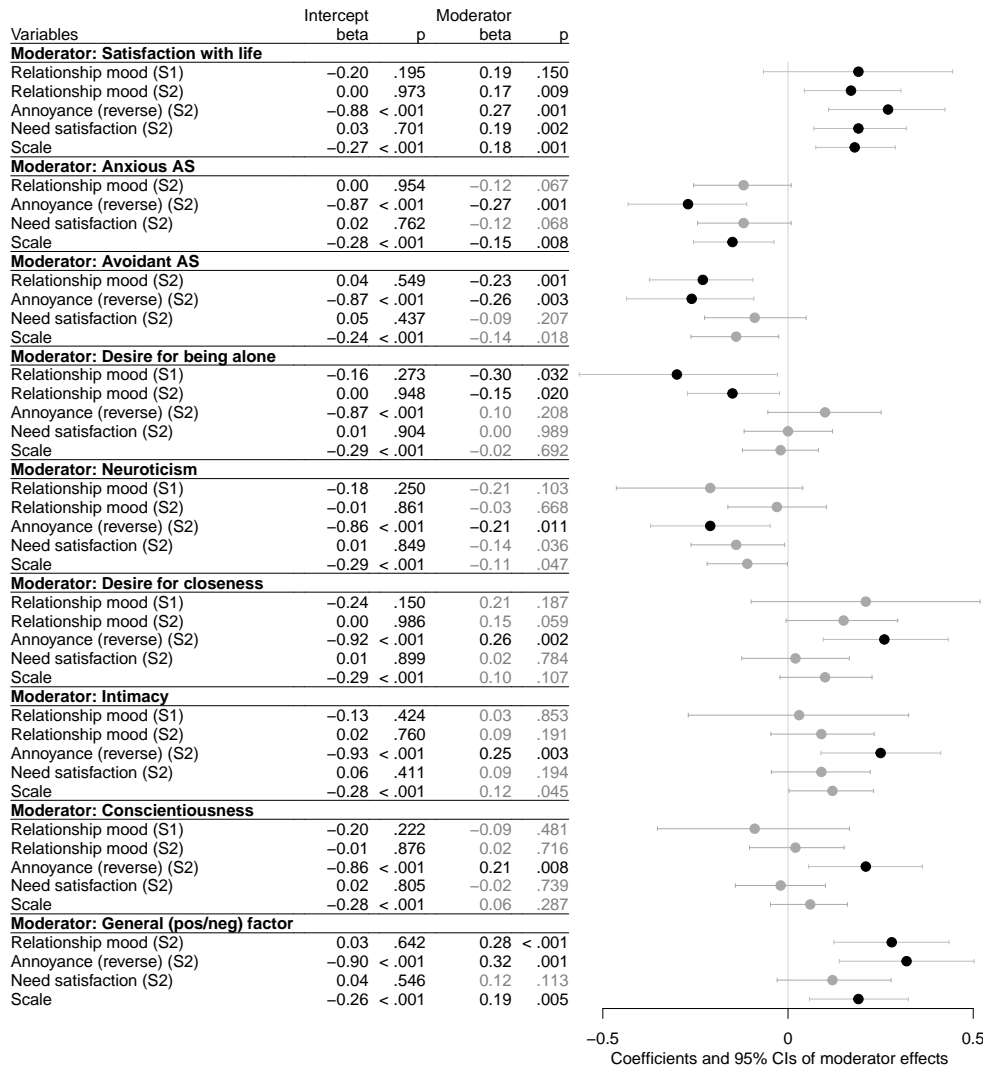


Figure 4. Moderation of mean-level bias by different moderators (i.e., main effects of these moderators) for different relationship satisfaction items. The interaction between moderator and mean relationship satisfaction states (i.e., the moderation of tracking accuracy) is included in the models but not reported here. S1 = Study 1, S2 = Study 2. N (Study 1) = 118, N (Study 2) = 486, AS = Attachment Style. Moderator effects that were significant after controlling the false discovery rate at $\alpha = 5\%$ (two-tailed) are displayed in black (for relationship mood based on a meta p -value of both studies), all other moderator effects are displayed in grey. Figure created with the *forestplot* package (Gordon & Lumley, 2017), available at <https://osf.io/sq7mw/>, under a CC-BY4.0 license.

What Level of Aggregation is Sufficient to Approach a Reliable Measurement of the Global Index? (RQ4)

For RQ4 we only report the results for Study 2 in the main text, because in this study four instead of only two weeks of sampling were available. The respective results for Study 1 can be found in the Supplemental Materials. Figure 5 shows the association between different numbers and schedules of ESM assessments and the CSI as global relationship satisfaction measure assessed before the ESM study period. Using all five assessments of the day for all four weeks that were sampled, the association between the aggregated ESM state relationship satisfaction scale and the CSI was $\beta = .59$ (see Table 3). The size of the association was already nearly achieved after one ($\beta = .55$) or two weeks of sampling ($\beta = .57$).

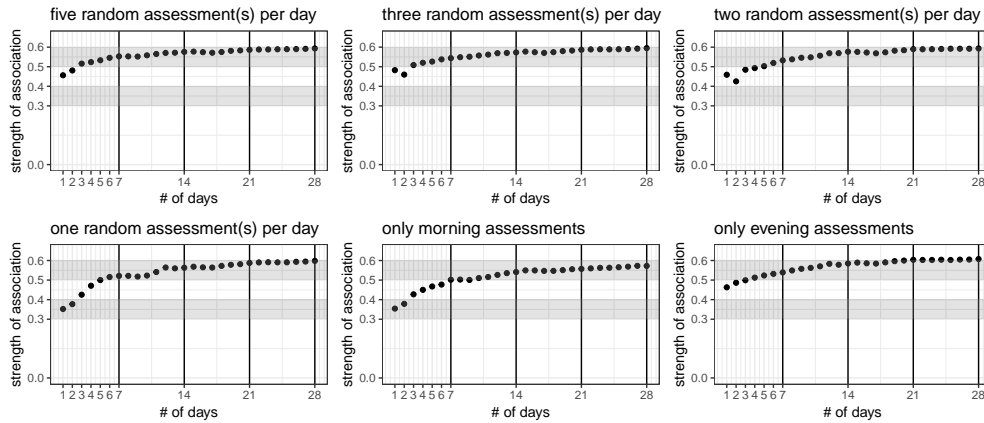


Figure 5. Association between (aggregated) state relationship satisfaction and global relationship satisfaction for different number of assessments and schedules in Study 2.

Looking at different numbers of assessments per day with a random sampling plan shows in both studies that a higher number of assessments matters only for the first few days. Afterwards, a higher sampling rate does not increase the effect size of the association meaningfully faster or stronger than fewer assessments.

Comparing evening assessments with morning and single random assessments shows in Study 2 that the evening assessments descriptively reach peak associations slightly sooner than the other sampling plans. However, we could not observe similar differences between the sampling plans in Study 1.

Discussion

The present studies tapped into different aspects of assessing relationship satisfaction, comparing state assessments with retrospective assessments and global evaluations. To understand the relationship between states, global and retrospective evaluations, different summary statistics of the state assessments were evaluated in their ability to predict the other assessment modes. Averaging the state assessments showed the highest association with the other two measures in both studies, but most other summary statistics performed similarly well or provided small incremental information. When individuals try to recap their experiences in their relationship, they might remember some occurrences better than other ones. We therefore compared the retrospective assessments with the averaged state reports to assess tracking accuracy and to uncover a potential mean-level bias of the sample when recalling the study weeks. As expected, the resulting tracking accuracy was positive, confirming that individuals' retrospective assessments converge to a large extent with what they on average report to have experienced on a momentary basis; however, the estimation differed significantly from a perfect tracking accuracy of one for all but the relationship mood item, indicating also the presence of systematic deviations. We further found a negative mean-level bias during retrospection for the scale of all items in Study 2, driven by individuals reporting a stronger intensity of them having been annoyed in their relationship compared to the average of what they indicated on a momentary basis.

We explored several moderators of this mean-level bias, and found the strongest to be global relationship satisfaction concurrently assessed with the retrospection: Individuals who are globally more satisfied with their relationship when they recall their study weeks, tend to less strongly overestimate their level of annoyance, and also tend to indicate retrospectively better relationship mood and need satisfaction in the relationship. This moderating effect was also observed for global relationship satisfaction assessed before the study period, albeit less strongly and not for all measures, as well as for individuals who report higher levels of life satisfaction, intimacy in their relationship, desire for closeness, and conscientiousness. Individuals who showed higher levels of dysfunctional attachment styles, and those high in neuroticism or with a strong desire for being alone overestimated the level of annoyance even more than the average, or underestimated their relationship mood and need satisfaction. Additionally, in Study 2 we examined the effects of factor scores extracted for a latent factor representing general positivity in trait measures. Individuals who scored high on this factor showed less of an overestimating of

annoyance, but overestimated their relationship mood.

Finally, our results show that when assessing state relationship satisfaction for more than a few days, the amount of surveys per day seems not to play a crucial role with regard to capturing states representative for the global evaluation of relationship satisfaction. It takes however approximately two weeks to maximize the informational value of the state assessments.

Global and Retrospective Assessments of Relationship Satisfaction are Best Represented by the Mean of States (RQ1)

Our data suggests that when individuals globally or retrospectively evaluate their relationship, they provide information that is foremost reflected by the mean, but also by other summaries of their daily relationship satisfaction states. In contrast to what is described by the peak-and-end rule (Fredrickson, 2000), the 90th quantiles of the state distribution (i.e., positive peaks) and the states reported during the last day explained the lowest amount of variance in retrospective evaluations. Still, recency and peaks represented by the mean of the last week and 10th quantiles (i.e., negative peaks), as well as the median reflected the retrospection only a little bit worse than the mean. Further, descriptively compared, the mean of the first week had lower effects than the mean of the last week; this could support the interpretation of a recency effect during retrospection of relationship satisfaction; but it could also point to individuals developing a certain response pattern over the course of the ESM study, which they draw upon when retrospectively assessing the study period. The development of such a response pattern is supported by the fact that in our longer Study 2 the standard deviation of answers during the first week is significantly higher for all relationship satisfaction items than the standard deviation during the last week (all $ps < .001$). That is, individuals seem to develop a more stable response to the questions, which would undermine the goal of ESM studies to capture state experiences instead of more general beliefs about the relationship. Both interpretations, a recency effect and a more stable response pattern over the course of the ESM study, are possible given the current analyses, and might also both be valid simultaneously.

Our varying results for the different conceptualizations of recency effects (last day, last week) and peaks (highs, lows) are consistent with earlier research: For general daily affect which was retrospectively evaluated on the next day the peak-and-end rule was also not the best explanation, whereas the average of affective states proved to be a good indicator (Miron-Shatz, 2009). The author argues that the end of a day is not special in a sense that

some outcome is reached, which was the case for studies that demonstrated the peak-and-end rule. In the same way were the last days of our study periods not distinctively meaningful for the relationship of our participants. Feldman Barrett (1997) further discusses that the peak-and-end rule was shown for retrospective evaluations that were made immediately after an experience, which was also not the case in our studies (e.g., the mean delay was two days in Study 2).

Regarding incremental effects of other summary statistics beyond the mean, previous research showed for general affect that the lowest (i.e., most negative) affect during a day incrementally explained the retrospective evaluation, whereas the highest (i.e., most positive) affect did not or less so (Ganzach & Yaor, 2018; Miron-Shatz, 2009). This additional effect of intense lows but not highs is plausibly attributed to the general phenomenon of negative experiences weighing more than positive ones (see Baumeister, Bratslavsky, Finkenauer, & Vohs, 2001; Vaish, Grossmann, & Woodward, 2008 for reviews). Consistent to this, in Study 2, we found that 10th quantiles (i.e., especially negative relationship evaluations) had incremental value to the prediction of retrospection above the effect of the mean of states, for all but the need satisfaction item, whereas the 90th quantiles of the states had an incremental effect only for the retrospection of annoyance (i.e., when individuals were not annoyed at all by their partner). We propose an additional explanation for 10th quantiles providing more information than 90th quantiles: The distribution of relationship satisfaction was skewed in the direction of positive evaluations (most strongly for the annoyance item, mean skew in Study 2 = -3.67). In consequence, 90th quantiles were highly similar to mean values (thereby reducing the informational value compared to 10th quantiles) and had low variance across the sample because of a ceiling effect. Thus, the predictive value the 90th quantiles could provide was limited from the start.

The fact that they still improved the prediction significantly in case of annoyance, might be explained with the observed negative coefficient: The 90th quantile seemingly corrects the error the skewedness introduced to the effect of the mean state. This kind of correction seems to also be provided by the median, as it had also a negative coefficient, being significant for the relationship mood item and the scale of all items. Therefore, characteristics of the distributions of the constructs that are studied must be considered as they might influence which summary statistic improves the prediction.

Finally, even when the mean across all states was already entered in the regression, the average state of the last week and of the last day did still

provide significant incremental information for the prediction of the (positively framed) retrospective relationship mood and need satisfaction items, but not for the (negatively framed) annoyance item. Consistent to this result, end evaluations seem to matter more for positive affect than for negative affect (Ganzach & Yaor, 2018).

In sum, our results suggest that the use of the mean as a summary statistic of individuals' relationship satisfaction states is a valid option when the goal is to represent what is captured by retrospective or global evaluations. Vice versa, such global evaluations primarily indicate individuals' average experiences. Still, our data show that especially negative relationship evaluations (e.g., captured by the 10th quantile of a distribution) provide additional information. Exceptionally positive evaluations as indicated by the 90th quantile, or the median might only be incrementally relevant when encountering skewed distributions. Averages of states that are more proximal to the time of retrospection provide in our study an incremental effect for positively framed items. All of these incremental effects may have a functional basis, and may cause a single retrospective assessment to be especially influenced by salient events (see also Lay et al., 2017).

Individuals Overestimate their Level of Annoyance in Retrospection (RQ2), which is Moderated by Global Evaluations of the Relationship and Person Characteristics (RQ3)

Overall mean-level bias. When comparing the retrospective relationship satisfaction with the average state during the study period, our data showed significantly different evaluations of the annoyance item, but not of the relationship mood and need satisfaction items. Specifically, individuals overestimated the amount of them having been annoyed by their partner, which results in a lower relationship satisfaction score in retrospection compared to the averaged states (i.e., a negative mean-level bias), if annoyance is included in a scale of relationship satisfaction.

This result cannot be explained by the initial elevation bias found for subjective reports (Shrout et al., 2017), as individuals report an elevated level of annoyance by their partner *after* repeated assessment. It also contrasts the general trend for a positive mean-level bias found in the meta-analysis of Fletcher and Kerr (2010) across judgment categories (“positive” in the sense of evaluating the relationship and the partner better than the relationship or the partner actually is, not in the sense of a general overestimation in retrospection). However, depending on the target of the evaluation, the meta-analysis

showed variance in the direction of biases, which is reflected in our results. Previous research which focused on retrospection of relationship experiences found that individuals overestimate their (positively framed) relationship satisfaction, but also their own and their partner's daily positive *and* negative behaviors (Oishi & Sullivan, 2006). This might point to a general pattern of overestimating the occurrence or intensity of specific experiences, independent of the target of evaluation. Miron-Shatz et al. (2009) found such an overestimation trend for general affect (see also Thomas & Diener, 1990; Mitchell, Thompson, Peterson, & Cronk, 1997), but it was stronger for negative affect (see also a recent study by Neubauer et al., 2020 that also shows an overestimation of negative affect in retrospection, but less so for positive affect). It is therefore noteworthy that a) despite referring to our result as a negative mean-level bias (because the relationship quality is described worse in retrospection compared to the averaged state), we observed an *overestimation* in retrospection, b) this overestimation occurred for the negatively framed domain of annoyance. Negative information dominates positive ones in various domains (see Baumeister et al., 2001; Vaish et al., 2008 for reviews). Lay et al. (2017) argue that the arousal that accompanies an affective reaction is an important factor for the relevance of an experience. Following these ideas, individuals might remember instances of them having been annoyed more profoundly, because these situations were accompanied with negative *and* aroused affect, in contrast to the average positive, not especially aroused daily relationship mood and need satisfaction in healthy relationships.

Moderation of mean-level bias by global relationship satisfaction. This line of reasoning is further supported by the fact that global relationship satisfaction showed a clear pattern of moderating the mean-level bias for every item: The unhappier individuals were globally with their relationship, the lower they rated their relationship mood and need satisfaction during the study period (which then was probably more often accompanied with negative emotions), and the higher they rated their level of annoyance in retrospection. Accordingly, the globally happier individuals indicated to be, the closer was their retrospective assessment to the average ESM reports, eventually showing the trend of overestimating the relationship satisfaction in comparison. This result extends findings highlighting the role of global relationship satisfaction for retrospective relationship reports (e.g., Halford, Keefer, & Osgarby, 2002), and its moderating role of bias and accuracy across a range of other judgment categories (Fletcher & Kerr, 2010). Research by Galak and Meyvis (2011) shows that an overestimation of aversive experiences is especially pronounced when individuals expect such experiences in the future. Being annoyed and

having one's needs frustrated can be considered aversive experiences. Individuals who are globally unhappy in their relationship have a good reason to expect similar experiences in the future, under the assumption that relationships do not break up easily. From a coping perspective, a study by Luong, Wrzus, Wagner, and Riediger (2016) indicates that valuing negative affect may even be functional with regard to psychosocial and physical functioning. It may therefore be adaptive to focus on negative experiences when remembering the past, to brace for and adapt to similar future relationship episodes.

Compared to an assessment before the ESM study period, global relationship satisfaction concurrently assessed with the retrospection showed the strongest moderating effect. Thus, the recall process seems to be strongly affected by individuals' momentary evaluations, as suggested by Ross (1989), thereby replicating early findings (Holmberg & Holmes, 1994; Karney & Coombs, 2000; McFarland & Ross, 1987). It is important to emphasize that although global relationship satisfaction was quite stable across the four weeks ($r_{CSI} = .82$ for women and $r_{CSI} = .79$ for men), the *concurrent* assessments of global relationship satisfaction showed the strongest and most robust effects. That is, the concurrent evaluation of the relationship seems to capture information beyond the stable variance of global relationship satisfaction, which could be interpreted as state variance that is shared with and relied upon during retrospective evaluations (the correlation between retrospection as a scale and the concurrent CSI was $r = .70$ for women and men). However, studies examining the processes involved when individuals evaluate their global *life* satisfaction find little evidence of experientially induced mood on individuals' evaluations (Yap et al., 2016). Future studies should therefore examine the effect of experientially induced momentary relationship feelings on the recall and global evaluation of relationship satisfaction.

Moderation of mean-level bias by other person characteristics.

Additional moderating variables support the idea that individuals draw on stable identity-related and situation-specific beliefs when they report on experiences retrospectively (Robinson & Clore, 2002b): Satisfaction with life, which encompasses the belief that one's life is good, had a positive moderating effect (see also Diener et al., 1984), whereas avoidant and anxious attachment styles, which capture negative situation-specific expectations, had negative moderating effects (see also Overall et al., 2015; Pietromonaco & Feldman Barrett, 1997). Similarly, neuroticism moderated the negative mean-level bias of the more affective annoyance item, showing that individuals high in neuroticism overestimate their level of annoyance even stronger. This result mirrors the

finding that individuals high in neuroticism overestimate their negative affect in retrospection (Feldman Barrett, 1997), and suggests that this effect generalizes to relationship-specific evaluations as well. Additionally, the explicit desire for closeness had a positive moderating effect on the assessment of the annoyance item, whereas individuals' explicit desire for being alone had a negative moderating effect on the relationship mood item. Previous research already shows that motivational variables influence the recall of autobiographical events (e.g., what experiences are remembered, Woike, 1995; or how the partner behaved, Pusch et al., 2020). It is assumed that during memory retrieval individuals' explicit motives modulate which experiences they capitalize on, namely events that support or were key in changing their self-concept of their goals (Woike, 2008). In this line of reasoning it is sensible that individuals with a strong explicit desire for closeness do not overestimate the level of annoyance as much, as these experiences work against reaching their goal of feeling close to their partner, and are hindering in maintaining a coherent fit between one's goals and one's experiences. In contrast, capitalizing on one's relationship mood when it was bad helps reaffirming the self-concept for individuals who have a strong explicit desire for being alone, that is for individuals who indicate that they regularly need distance from their partner and time for themselves. It is however unclear why only specific items of relationship satisfaction were moderated by the desires, but not others. In sum, rather than giving each experience in their relationship equal meaning during retrospection, individuals seem to capitalize on certain experiences based on their expectations about the relationship, their impression of themselves and their self-ascribed desires.

As the evaluation of the annoyance item was the main reason for the mean-level bias, and therefore apparently especially susceptible to distortion, we found further moderators that only affected the assessment of this item: In line with the previous moderators, intimacy in the relationship (an indicator of a satisfying relationship with regard to closeness, Laurenceau, Barrett, & Rovine, 2005) had a positive moderating effect for the retrospection of annoyance, reducing the difference between these assessment modalities towards a more similar perception. Surprisingly, the personality factor of conscientiousness turned also out to be a positive moderator. It might be related to a more thorough process when answering the questions, and therefore a more balanced retrospective evaluation as result.

Moderation of mean-level bias by a global positivity factor. Given that we found positive moderating effects for constructs that might

be perceived as positive (e.g., relationship / life satisfaction), and negative moderating effects for those that might be perceived as negative (e.g., dysfunctional attachment, neuroticism), our results might not be driven by the specific constructs we examined, but alternatively reflect a more general positivity effect or a response style. We considered this possibility by examining a single factor across all self-report items as additional moderator in Study 2: The item loadings suggest that such a factor could be interpreted as a more global identity-related positive self-view about oneself, one's life and one's relationship. Alternatively, it might also reflect a response style characterized by social desirability. This factor indeed moderates the mean-level bias of the annoyance and of the relationship mood item. Hence, depending on the interpretation of the factor, differences between retrospection and the averaged ESM reports seem to be also explained by individuals' global positivity or negativity, or the degree to which they are prone to social desirable responding.

When examining the aforementioned specific moderators simultaneously with this general factor, some moderator effects disappeared, but some other were robust to this control analysis: This suggests that we can confidently interpret some constructs as being relevant as specific moderators of mean level bias. For example, all effects of the relationship satisfaction concurrently assessed with retrospection remained significant, as well as most effects of life satisfaction and relationship satisfaction assessed before the study period. Hence, beyond a general positive assessment of self-report scales, these constructs capture unique variance in satisfaction with specific domains at specific time-points, which explain mean-level differences between retrospection and averaged ESM reports. This robustness was also the case for conscientiousness and anxious attachment as moderators of the annoyance assessment.

The effects of the other moderators (e.g., of avoidant attachment, neuroticism, intimacy, and explicit desires) seem to be more readily explained to be driven by a general positivity/negativity effect. Therefore, our prior interpretations regarding the processes that might cause these specific constructs to moderate the observed differences might be confounded with the effects of a general positive or negative attitude, and should be treated with caution.

Summary of moderating effects. In sum, our results suggest that when individuals globally indicate to be unhappy, on average the retrospective reports will suggest a higher occurrence of negative experiences in the relationship as what would be derived from the average of momentary reports. This difference is more pronounced the globally unhappier the individuals are, and is also influenced by aspects of individuals' attachment styles, personality, and

global positivity during self-report assessments.

We did neither find effects of gender, as it was found for other judgment domains (Fletcher & Kerr, 2010), nor for delay of retrospection, as would be derived from the accessibility model (Robinson & Clore, 2002a, although we did not systematically vary different delay periods, see Supplemental Materials for estimates of the respective models).

Origination of the bias: Retrospection or ESM reports? In our analyses, we treated the mean ESM measure as truth criterion, with deviations from it during retrospection as bias. This modeling choice has consequences for our interpretation, which have to be carefully considered. First, this assumes that averaging the states is the correct way of summarizing the multiple moments of (dis-)satisfaction an individual experienced during the study, rather than giving the satisfaction during certain situations more weight than other situations (e.g., when spending time with the partner or during a conflict). Second, this modeling of ESM states as the reference criterion might be suggestive of these assessments being not or at least less biased than retrospective assessments. However, while ESM reports might produce fewer recall errors than retrospection, they might be equally or more strongly affected by other response biases, such as those generated by one's self-concept (see Finnigan & Vazire, 2017 for a discussion of such "self-biases" for ESM reports). In fact, we could have modeled the retrospection as truth criterion, with deviations of the aggregated ESM states as bias: This would have led to the interpretation that aggregated ESM reports underestimate the amount of annoyance that "actually" (according to retrospection) occurred in the relationship.

We would like to emphasize that our decision to model the ESM reports as truth criterion impacts the way we interpret our results (i.e., as the retrospective assessment being biased in the sense of an over- or underestimation), but that this choice could reasonably be made differently by other researchers. Importantly, our goal was not to present the ESM reports as the objective gold standard (which was rather a side effect of a modeling decision we had to make), but to uncover any differences between retrospection and aggregated ESM reports. The fact that these two measures deviate from each other, may be due to different measurements models being applied for representing the relationship satisfaction during the study period, and may lead to the practical implication that the different measurements produce reports with differential validity, which may be useful for different purposes. For example, one could speculate that for couple therapy the retrospective assessment may be more suited to indicate dysfunctional recall biases, and the need of inter-

ventions aimed at cognitive reframing, while the aggregation of momentary assessments may draw attention to the influence of situations which might be otherwise less salient.

A Saturation Effect is Visible after Assessing Relationship Satisfaction States for Two Weeks (RQ4)

We also investigated what informational value different sampling schemes of ESM assessments provide with regard to capturing a global assessment of relationship satisfaction. We examined two factors that can be manipulated when designing an ESM study: The number and the scheduling of the assessments.

The number of assessments can be influenced in two ways: By increasing the number of assessments per day, or by increasing the overall length of the study. Both ways of collecting more experiences have pros and cons (e.g., capturing short-term dynamics vs. enhancing participant burden) and must be decided depending on the research question at hand (see Bolger & Laurenceau, 2013). The decisions are however not independent, as a less intensive sampling per day may invoke the need for a longer study period to achieve representative information. In our data it takes about five days to achieve a similar overall level of association with global relationship satisfaction, regardless of whether only one random sample per day is considered or five semi-random samples per day. After five days, the increase in association strength is similar steady across different numbers of assessments per day, maxing at around $\beta = .60$ (but see Schönbrodt et al., 2019 demonstrating high within-day variance of state relationship satisfaction, which raises the need to sample multiple times per day to capture the dynamics occurring within a day). Further, we see a saturation effect after approximately two weeks, meaning that after this study period more ESM data does not provide much more incremental information for predicting global relationship satisfaction – independently of the number of assessments per day. This complements the findings of Epstein (1979), who also found two weeks to be necessary for achieving a representative sample of individual’s behaviors.

Regarding the timing of the assessments, we examined three common strategies: Assessing in the evening, in the morning, or at a random time during the day. While we descriptively found in our larger Study 2 that evening assessments seem to be more valid for representing global relationship satisfaction, because both the initial association strength was higher and the maximum association strength was reached sooner, this did not replicate in our smaller

Study 1. Hence, further research is needed to assess the robustness of the differences between sampling plans when only sampling once.

Limitations

Several potential limitations have to be considered when interpreting the results of our studies. First, a necessary condition for the investigation of bias and accuracy (RQ2 and RQ3) is the commensurability of the measures that are being compared, in our case of the retrospection and the state assessment. In principle, this is given in the current studies, as the same content is evaluated in both measures (leading to “nominal equivalence”)¹⁶ on the same scale (transformed to the same metric, leading to “scale equivalence”; see Edwards & Shipp, 2007 for the use of these terms). However, slightly different assessment characteristics for ESM and retrospection, especially visual differences in the presentation of the sliders used, could pose a threat to commensurability: The retrospective assessment was answered in a browser on the participants’ personal computers, and in Study 2 the three relationship satisfaction items were presented in a block. The ESM assessment, in contrast, was completed on the smartphone and the items were presented at different positions in the ESM survey (but see Wells, Bailey, & Link, 2014, finding little psychometric differences between web and smartphone presentation of items). Further, slightly different slider characteristics might have biased the answers (see Matejka, Glueck, Grossman, & Fitzmaurice, 2016). First, a missing “neutral” label in the retrospective assessment could have removed an anchor effect that might have been present in ESM. However, the largest biases were found for the annoyance item, which also in the ESM assessment did not have a neutral label (see Figure 1). Second, the slider having a start position during retrospection, whereas in ESM no start value was preselected, could have evoked

¹⁶A study by Winkielman, Knäuper, and Schwarz (1998) suggests that when referring to different time frames in questionnaires, the interpretation of the phenomenon that is being assessed changes. Specifically, the study provides evidence that a reference to longer time frames (e.g. “during the last month”) prompt individuals to report less frequent, but more intense events, compared to a reference to shorter time frames (e.g. “during the last week”). The authors explain this with the ambiguity of the phenomena that are studied, and note that an explicit definition of the phenomenon resolves this problem; importantly, they also show that the interpretation elicited by a reference to a shorter time frame carries over when subsequently a longer time frame is assessed (although this did not completely eliminate the effect of the time frame, at least not for frequency reports). Such a carry-over effect is to be expected in our study, as individuals could internalize the meaning of the different relationship satisfaction items multiple times per day for several weeks. Although we cannot rule out that their interpretation of the relationship satisfaction items changed when they were asked to assess them retrospectively for the study period right after the study, we do not find it plausible that they did not recognize the questions and interpreted the item content differently as during the multiple instances they assessed it during the prior weeks.

another anchoring effect. As the start position was in the middle of the scale, this might have canceled out the missing “neutral” option for the relationship mood and need satisfaction items. For the annoyance item this might actually have introduced a biased anchoring point, although it is unclear why this would produce an overestimation of annoyance: Participants rather seem to choose preselected options less often (Funke, 2016), that is, the preselection seems to evoke the need to move the slider further away; given that on an absolute level the amount of annoyance reported was low (mean of retrospection of not reverse scored annoyance = 1.72 on a scale from 0 to 10) and the labeled end of the scale “*not at all (annoyed)*” might attract answers, these kind of biasing design effects should have rather led to an underestimation of annoyance, rather than the observed overestimation. Finally, although we transformed all measures to the same metric (0-10), the ESM answers on the slider items were initially saved in a higher resolution (on scales from 1-7 and 0-10 with answers saved with multiple positions after the decimal point) than the retrospective evaluations (on scales from 0-100 and 1-100 with answers rounded to whole numbers). To assess the magnitude of error these different resolutions might have added to our results, we adjusted the resolution of the ESM answers in Study 2 to the answers during retrospection by transforming them to a 1-100 scale, rounding them to whole numbers, and transforming them back to a 0-10 scale. All of the results replicate when running the analyses with these scales, with changes in the estimates only on the third or fourth decimal place after the comma.

Further, our analyses showed that a mean-level bias primarily occurs for the retrospection of experienced annoyance, therefore biasing the whole relationship satisfaction scale in retrospection when this item is included in scale calculation. Therefore, our results may not generalize for other relationship satisfaction scales that do not include annoyance, or maybe more generally those scales that do not contain items pertaining to negative affect in the relationship. We would argue, however, that simply removing the annoyance item, or more generally avoiding the assessment of negative affectivity in relationships is no solution. As also discussed in Schönbrodt et al. (2019), the annoyance item contributes to a more heterogeneous index of relationship satisfaction, taking into account the impact of negative experiences for relationship evaluation (as other scales also do, e.g., the global measures applied in our studies, Funk & Rogge, 2007; Rogge et al., 2016). Depending on the research question, this broader assessment of relationship satisfaction is necessary to achieve a complete picture of individuals’ relationship evaluation and may be more suited to differentiate couples in generally happy relationships.

Moreover, our analyses concerning the required number of ESM surveys and the optimal sampling procedure to reach satisfactory associations with a global evaluation were not based on an experimental design: All participants answered the same amount of five surveys with a semi-random schedule, but for our analyses we selected different subsets of surveys as predictors of global relationship satisfaction. In consequence, the effects we found might differ if individuals would actually only answer one survey (or fewer than five surveys) per day (in the morning or in the evening), as the ESM procedure we applied could have induced reactivity such as a heightened sensitivity for participant's relationship feelings. If this would be the case, then our effects might be exaggerated, and a lower number of surveys for instance might take longer than the reported five days to reach a similar association strength as a higher number of surveys. Future work should compare the effects we found in our study with effects from an experimental study which randomly assigns participants to different ESM designs.

Finally, despite the fact that we preregistered some hypotheses for RQ2 , the presented results should mainly be regarded as exploratory, as we were inconsistent in the preregistration regarding which items we will use as a measure of state and retrospective relationship satisfaction. For maximal transparency and given the exploratory nature of the other research questions, we reported the results for all available items, and controlled the false discovery rate at $\alpha = 5\%$.

Conclusion

The present studies provide insight into various domains related to the assessment of relationship satisfaction. First, our studies showed that global and retrospective evaluations best capture the average of relationship satisfaction states, with other summary statistics providing incremental information. Second, the retrospective overestimation of negative affect found in prior research also holds for a relationship-specific negative evaluation of annoyance. Third, this difference between retrospective and aggregated ESM assessments is especially pronounced for individuals who globally report low relationship and life satisfaction, with other person characteristics being further relevant. Last, our results show that approximately two weeks are necessary to sample a representative amount of relationship satisfaction states. The current research uncovers differences of various assessment modalities of relationship satisfaction that ought to be considered when applying them: Retrospective assessments and in extension also global evaluations might provide notably

different information than aggregated ESM reports when targeting negative experiences in a relationship, especially for individuals who globally report to be unhappy. Depending on the research question or the aim of assessment in a practitioner setting, it has to be carefully decided whether one is interested in the average of the experiences that were reported to happen in the relationship, with each of these momentary reports probably having their own biases; or whether the idiosyncratic capitalization individuals make for specific experiences is of special interest, which is provided by retrospective or global measures.

Contributions

Contributed to conception and design: CZ-H, FS

Contributed to acquisition of data: CZ-H

Contributed to analysis and interpretation of data: CZ-H, FS

Drafted and/or revised the article: CZ-H

Approved the submitted version for publication: CZ-H, FS

Acknowledgements

The authors thank Birk Hagemeyer, Sebastian Pusch, Paula Fehrmann, Nicole Horn, and Lara Lietge for the PACT scoring, Helen Baumann for assistance during data collection of Study 1, as well as Tobias Kächele, Lukas Müller and Ludwig Zellner for the app development.

Funding information

This research was funded by grants from the German Research Foundation to Felix Schönbrodt (SCHO 1334/5-1) and Birk Hagemeyer (HA 6884/2-1).

Competing interests

The authors have no competing interests to declare.

Data accessibility statement

The data of both studies are available as a scientific use file (Zygar et al., 2018b for Study 1; Zygar-Hoffmann, Hagemeyer, Pusch, & Schönbrodt, 2020 for Study 2).

References

- Arslan, R. C., Walther, M. P., & Tata, C. S. (2020). Formr: A study framework allowing for automated feedback generation and complex longitudinal experience-sampling studies using R. *Behavior Research Methods*, 52, 376–387. doi:10.3758/s13428-019-01236-y
- Arslan, R., & Tata, C. (2016). *formr.org survey software (version v0.16.12)*. Retrieved from <https://formr.org/>
- Arslan, R., & Tata, C. (2017). *formr.org survey software (version v0.16.13-version v0.17.16)*. Retrieved from <https://formr.org/>
- Aust, F., & Barth, M. (2018). *papaja: Create APA manuscripts with R Markdown*. Retrieved from <https://github.com/crsh/papaja>
- Barton, K. (2018). *MuMIn: Multi-model inference*. Retrieved from <https://CRAN.R-project.org/package=MuMIn>
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1), 1–48. doi:10.18637/jss.v067.i01
- Baumeister, R. F., Bratslavsky, E., Finkenauer, C., & Vohs, K. D. (2001). Bad is stronger than good. *Review of General Psychology*, 5(4). doi:10.1037/1089-2680.5.4.323
- Baumert, A., Schmitt, M., Perugini, M., Johnson, W., Blum, G., Borke-nau, P., ... Wrzus, C. (2017). Integrating personality structure, personality process, and personality development. *European Journal of Personality*, 31(5), 503–528. doi:10.1002/per.2115
- Bell, K. J. (2008). *Intimate partner violence on campus: A Test of social learning theory* (PhD thesis). Retrieved from <https://pdfs.semanticscholar.org/730b/49e22394259ac4a1c96692f9ec8c280ccff4.pdf>
- Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, 57(1), 289–300. doi:10.1111/j.2517-6161.1995.tb02031.x
- Biderman, M. D., Nguyen, N. T., Cunningham, C. J., & Ghorbani, N. (2011). The ubiquity of common method variance: The case of the big five. *Journal of Research in Personality*, 45(5), 417–429. doi:10.1016/j.jrp.2011.05.001

Blood, R., & Wolfe, D. (1960). *Husbands and wives*. Glencoe, IL: Free Press.

Bolger, N., & Laurenceau, J.-P. (2013). *Intensive longitudinal methods*. New York, NY: Guilford.

Bürkner, P.-C. (2017). brms: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software*, *80*(1), 1–28. doi:10.18637/jss.v080.i01

Chang, V. T., Overall, N. C., Madden, H., & Low, R. S. (2018). Expressive suppression tendencies, projection bias in memory of negative emotions, and well-being. *Emotion*, *18*(7), 925–941. doi:10.1037/emo0000405

Christensen, T. C., Wood, J. V., & Feldman Barrett, L. (2003). Remembering everyday experience through the prism of self-esteem. *Personality and Social Psychology Bulletin*, *29*(1), 51–62. doi:10.1177/0146167202238371

Conner, T. S., & Feldman Barrett, L. (2012). Trends in ambulatory self-report: The role of momentary experience in psychosomatic medicine. *Psychosomatic Medicine*, *74*, 327–337. doi:10.1097/PSY.0b013e3182546f18

Csikszentmihalyi, M., & Larson, R. (1987). Validity and reliability of the Experience-Sampling Method. *The Journal of Nervous and Mental Disease*, *175*(9), 526–536. doi:10.1097/00005053-198709000-00004

Dewey, M. (2018). *metap: Meta-analysis of significance values*. Retrieved from <https://cran.r-project.org/package=metap>

Diener, E., Emmons, R. A., Larsen, R. J., & Griffin, S. (1985). The Satisfaction With Life Scale. *Journal of Personality Assessment*, *49*(1), 71–75. doi:10.1016/j.amepre.2007.09.001

Diener, E., Larsen, R. J., & Emmons, R. A. (1984). *Bias in mood recall in happy and unhappy persons*. Paper presented at the 92nd Annual Convention of the American Psychological Association, Toronto, Ontario, Canada.

Edwards, J. R., & Shipp, A. J. (2007). The relationship between person-environment fit and outcomes: An integrative theoretical framework. In C. Ostroff & T. A. Judge (Eds.), *Perspectives on organizational fit* (pp. 209–258). San Francisco: Jossey-Bass. doi:10.4324/9780203810026

Ehrental, J. C., Dinger, U., Lamla, A., Funken, B., & Schauenburg, H. (2009). Evaluation der deutschsprachigen Version des Bindungsfragebogens "Experiences in Close Relationships - Revised" (ECR-RD) [Evaluation of the German Version of the Attachment Questionnaire "Experiences in Close Rela-

tionships - Revised" (ECR-RD)]. *PPmP-Psychotherapie Psychosomatik Medizinische Psychologie*, 59(06), 215–223. doi:10.1055/s-2008-1067425

Epstein, S. (1979). The stability of behavior: I. On predicting most of the people much of the time. *Journal of Personality and Social Psychology*, 37(July). doi:10.1037/0022-3514.37.7.1097

Fallis, E. E., Rehman, U. S., Woody, E. Z., & Purdon, C. (2016). The longitudinal association of relationship satisfaction and sexual satisfaction in long-term relationships. *Journal of Family Psychology*, 30(7), 822–831. doi:10.1037/fam0000205

Feldman Barrett, L. (1997). The relationship among momentary emotional experiences, personality descriptions, and retrospective ratings of emotion. *Personality and Social Psychology Bulletin*, 23, 1100–1110. doi:10.1177/01461672972310010

Fincham, F. D., & Rogge, R. D. (2010). Understanding relationship quality: Theoretical challenges and new tools for assessment. *Journal of Family Theory & Review*, 2(4), 227–242. doi:10.1111/j.1756-2589.2010.00059.x

Finnigan, K. M., & Vazire, S. (2017). The incremental validity of average state self-reports over global self-reports of personality. *Journal of Personality and Social Psychology*. doi:10.1037/pspp0000136

Fleeson, W. (2001). Toward a structure-and process-integrated view of personality: Traits as density distributions of states. *Journal of Personality and Social Psychology*, 80(6), 1011–1027. doi:10.1037/0022-3514.80.6.1011

Fleeson, W., & Gallagher, P. (2009). The implications of Big Five standing for the distribution of trait manifestation in behavior: Fifteen experience-sampling studies and a meta-analysis. *Journal of Personality and Social Psychology*, 97(6), 1097–1114. doi:10.1037/a0016786

Fletcher, G. J., & Kerr, P. S. (2010). Through the eyes of love: Reality and illusion in intimate relationships. *Psychological Bulletin*, 136(4), 627–658. doi:10.1037/a0019792

Forbes, E. E., Stepp, S. D., Dahl, R. E., Ryan, N. D., Whalen, D., Axelson, D. A., . . . Silk, J. S. (2012). Real-world affect and social context as predictors of treatment response in child and adolescent depression and anxiety: An ecological momentary assessment study. *Journal of Child and Adolescent Psychopharmacology*, 22(1), 37–47. doi:10.1089/cap.2011.0085

Fredrickson, B. L. (2000). Extracting meaning from past affective experiences: The importance of peaks, ends, and specific emotions. *Cognition and*

Emotion, 14(4), 577–606. doi:10.1080/026999300402808

Funk, J. L., & Rogge, R. D. (2007). Testing the ruler with item response theory: Increasing precision of measurement for relationship satisfaction with the Couples Satisfaction Index. *Journal of Family Psychology*, 21(4), 572–583. doi:10.1037/0893-3200.21.4.572

Funke, F. (2016). A web experiment showing negative effects of slider scales compared to visual analogue scales and radio button scales. *Social Science Computer Review*, 34(2), 244–254. doi:10.1177/0894439315575477

Galak, J., & Meyvis, T. (2011). The pain was greater if it will happen again: The effect of anticipated continuation on retrospective discomfort. *Journal of Experimental Psychology: General*, 140(1), 63–75. doi:10.1037/a0021447

Ganzach, Y., & Yaor, E. (2018). The retrospective evaluation of positive and negative affect. *Personality and Social Psychology Bulletin*, 45(1), 93–104. doi:10.1177/0146167218780695

Gerlitz, J.-Y., & Schupp, J. (2005). Zur Erhebung der Big-Five-basierten Persönlichkeitsmerkmale im SOEP. [Collection of the Big-Five personality characteristics in SOEP]. *DIW Research Notes*, 4, 1–36. doi:10.1016/j.jsis.2005.07.003

Glaesmer, H., Grande, G., Braehler, E., & Roth, M. (2011). The German version of the Satisfaction With Life Scale (SWLS). *European Journal of Psychological Assessment*, 27(2), 127–132. doi:10.1027/1015-5759/a000058

Gordon, M., & Lumley, T. (2017). *Forestplot: Advanced forest plot using 'grid' graphics*. Retrieved from <https://CRAN.R-project.org/package=forestplot>

Grant, A., Franklin, J., & Langford, P. (2002). The Self-Reflection and Insight Scale: A new measure of private self-consciousness. *Social Behavior and Personality: An International Journal*, 30(8). doi:10.2224/sbp.2002.30.8.821

Greischel, H., & Johnson, M. (n.d.). *Measurement invariance of a German translation of the Couples Satisfaction Index*.

Hagemeyer, B., & Neyer, F. J. (2012). Assessing implicit motivational orientations in couple relationships: The Partner-Related Agency and Communion Test (PACT). *Psychological Assessment*, 24(1), 114–28. doi:10.1037/a0024822

Hagemeyer, B., Neyer, F. J., Neberich, W., & Asendorpf, J. B. (2013). The ABC of Social Desires: Affiliation, being alone, and closeness to partner. *European Journal of Personality, 27*(5), 442–457. doi:10.1002/per.1857

Halford, W. K., Keefer, E., & Osgarby, S. M. (2002). "How has the week been for you two?" Relationship satisfaction and hindsight memory biases in couples' reports of relationship events. *Cognitive Therapy and Research, 26*(6), 759–773. doi:10.1023/A:1021289400436

Hedges, S. M., Jandorf, L., & Stone, A. A. (1985). Meaning of daily mood assessments. *Journal of Personality and Social Psychology, 48*(2), 428–434. doi:10.1037/0022-3514.48.2.428

Hofmann, W., Finkel, E. J., & Fitzsimons, G. M. (2015). Close relationships and self-regulation: How relationship satisfaction facilitates momentary goal pursuit. *Journal of Personality and Social Psychology, 109*(3), 434–452. doi:10.1037/pspi0000020

Holmberg, D., & Holmes, J. G. (1994). Reconstruction of relationship memories: A mental models approach. In N. Schwarz & S. Sudman (Eds.), *Autobiographical memory and the validity of retrospective reports* (pp. 267–288). doi:10.1007/978-1-4612-2624-6_18

Johnson, P. C. D. (2014). Extension of nakagawa and schielzeth's r2glmm to random slopes models. *Methods in Ecology and Evolution, 5*(9), 944–946. doi:10.1111/2041-210X.12225

Kahneman, D., Krueger, A. B., Schkade, D. A., Schwarz, N., & Stone, A. A. (2014). A survey method for characterizing daily life experience: The Day Reconstruction Method. *Science, 306*(5702), 1776–1780.

Karney, B. R., & Coombs, R. H. (2000). Memory bias in long term close relationships: Consistency or improvement? *Personality and Social Psychology Bulletin, 26*(8), 959–970. doi:10.1177/01461672002610006

Karney, B. R., & Frye, N. E. (2002). "But we've been getting better lately": Comparing prospective and retrospective views of relationship development. *Journal of Personality and Social Psychology, 82*(2), 222–238. doi:10.1037/0022-3514.82.2.222

Knowles, J. E., & Frederick, C. (2018). *MerTools: Tools for analyzing mixed effect regression models*. Retrieved from <https://CRAN.R-project.org/package=merTools>

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical*

Software, 82(13), 1–26. doi:10.18637/jss.v082.i13

Laurenceau, J.-P., Barrett, L. F., & Rovine, M. J. (2005). The interpersonal process model of intimacy in marriage: A daily-diary and multilevel modeling approach. *Journal of Family Psychology*, 19(2), 314–323. doi:10.1037/0893-3200.19.2.314

Lay, J. C., Gerstorf, D., Scott, S. B., Pauly, T., & Hoppmann, C. A. (2017). Neuroticism and extraversion magnify discrepancies between retrospective and concurrent affect reports. *Journal of Personality*, 85(6), 817–829. doi:10.1111/jopy.12290

LimesurveyGmbH. (2017). *LimeSurvey: An open source survey tool*. LimeSurvey GmbH, Hamburg, Germany. Retrieved from <http://www.limesurvey.org>

Lucas, R. E., Wallsworth, C., Anusic, I., & Donnellan, M. B. (2019). *A direct comparison of the Day Reconstruction Method and the Experience Sampling Method*. doi:10.31234/osf.io/cv73u

Luchies, L. B., Rusbult, C. E., Eastwick, P. W., Wieselquist, J., Kumashiro, M., Coolsen, M. K., & Finkel, E. J. (2013). Trust and biased memory of transgressions in romantic relationships. *Journal of Personality and Social Psychology*, 104(4), 673–694. doi:10.1037/a0031054

Luong, G., Wrzus, C., Wagner, G. G., & Riediger, M. (2016). When bad moods may not be so bad: Valuing negative affect is associated with weakened affect-health links. *Emotion*, 16(3), 387–401. doi:10.1037/emo0000132

Matejka, J., Glueck, M., Grossman, T., & Fitzmaurice, G. (2016). The effect of visual appearance on the performance of continuous sliders and visual analogue scales. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 5421–5432. doi:10.1145/2858036.2858063

McCrae, R. R., Bond, M. H., Yik, M. S., Trapnell, P. D., & Paulhus, D. L. (1998). Interpreting personality profiles across cultures: Bilingual, acculturation, and peer rating studies of Chinese undergraduates. *Journal of Personality and Social Psychology*, 74(4), 1041–1055. doi:10.1037/0022-3514.74.4.1041

McFarland, C., & Ross, M. (1987). The relation between current impressions and memories of self and dating partners. *Personality and Social Psychology Bulletin*, 13(2), 228–238. doi:10.1177/0146167287132008

Mill, A., Realo, A., & Allik, J. (2016). Retrospective ratings of emotions: The effects of age, daily tiredness, and personality. *Frontiers in Psychology*, 6(JAN). doi:10.3389/fpsyg.2015.02020

- Miron-Shatz, T. (2009). Evaluating multiepisode events: Boundary conditions for the peak-end rule. *Emotion, 9*(2), 206–213. doi:10.1037/a0015295
- Miron-Shatz, T., Stone, A., & Kahneman, D. (2009). Memories of yesterday's emotions: Does the valence of experience affect the memory-experience gap? *Emotion, 9*(6), 885–891. doi:10.1037/a0017823
- Mitchell, T. R., Thompson, L., Peterson, E., & Cronk, R. (1997). Temporal adjustments in the evaluation of events: The "Rosy View". *Journal of Experimental Social Psychology, 33*(4), 421–448. doi:10.1006/jesp.1997.1333
- Nakagawa, S., & Schielzeth, H. (2013). A general and simple method for obtaining R^2 from generalized linear mixed-effects models. *Methods in Ecology and Evolution, 4*(2), 133–142. doi:10.1111/j.2041-210x.2012.00261.x
- Neubauer, A. B., Scott, S. B., Sliwinski, M. J., & Smyth, J. M. (2020). How was your day? Convergence of aggregated momentary and retrospective end-of-day affect ratings across the adult life span. *Journal of Personality and Social Psychology, 119*(1), 185–203. doi:10.1037/pspp0000248
- Oishi, S., & Sullivan, H. W. (2006). The predictive value of daily vs. retrospective well-being judgments in relationship stability. *Journal of Experimental Social Psychology, 42*(4), 460–470. doi:10.1016/j.jesp.2005.07.001
- Overall, N. C., Fletcher, G. J., Simpson, J. A., & Fillo, J. (2015). Attachment insecurity, biased perceptions of romantic partners' negative emotions, and hostile relationship behavior. *Journal of Personality and Social Psychology, 108*(5), 730–749. doi:10.1037/a0038987
- Parkinson, B., Briner, R. B., Reynolds, S., & Totterdell, P. (1995). Time frames for mood: Relations between momentary and generalized ratings of affect. *Personality and Social Psychology Bulletin, 21*(4), 331–339. doi:10.1177/0146167295214003
- Pietromonaco, P. R., & Feldman Barrett, L. (1997). Working models of attachment and daily social interactions. *Journal of Personality and Social Psychology, 73*(6), 1409–1423. doi:10.1037/0022-3514.73.6.1409
- Pusch, S., Schönbrodt, Felix D., Zygar-Hoffmann, C., & Hagemeyer, B. (2020). Truth and wishful thinking: How inter-individual differences in communal motives manifest in momentary partner perceptions. *European Journal of Personality, 34*(1), 115–134. doi:10.1002/per.2227
- Rammstedt, B., & John, O. P. (2007). Measuring personality in one minute or less: A 10-item short version of the Big Five Inventory in English and German. *Journal of Research in Personality, 41*(1), 203–212.

doi:10.1016/j.jrp.2006.02.001

R Core Team. (2018). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from <https://www.R-project.org/>

Reis, H. T., & Gable, S. L. (2000). Event Sampling and other methods for studying everyday experience. In H. T. Reis & C. M. Judd (Eds.), *Handbook of research methods in personality psychology* (pp. 190–222). Cambridge, UK: Cambridge University Press.

Reise, S. P. (2012). The rediscovery of bifactor measurement models. *Multivariate Behavioral Research*, 47(5), 667–696. doi:10.1080/00273171.2012.715555

Robinson, M. D., & Clore, G. L. (2002a). Belief and feeling: Evidence for an accessibility model of emotional self-report. *Psychological Bulletin*, 128(6), 934–960. doi:10.1037/0033-2909.128.6.934

Robinson, M. D., & Clore, G. L. (2002b). Episodic and semantic knowledge in emotional self-report: Evidence for two judgment processes. *Journal of Personality and Social Psychology*, 83(1), 198–215. doi:10.1037/0022-3514.83.1.198

Robinson, M. D., Johnson, J. T., & Shields, S. A. (1998). The gender heuristic and the database: Factors affecting the perception of gender-related differences in the experience and display of emotions. *Basic and Applied Social Psychology*, 20(3), 206–219. doi:10.1207/s15324834basp2003_3

Rogge, R. D., Fincham, F. D., Crasta, D., & Maniaci, M. R. (2016). Positive and negative evaluation of relationships: Development and validation of the Positive – Negative Relationship Quality (PN-RQ) Scale. *Psychological Assessment*, 29(8), 1028–1043. doi:10.1037/pas0000392

Rosenberg, E. L. (1998). Levels of analysis and the organization of affect. *Review of General Psychology*, 2(3), 247–270. doi:10.1037/1089-2680.2.3.247

Ross, M. (1989). Relation of implicit theories to the construction of personal histories. *Psychological Review*, 96(2), 341–357. doi:10.1037/0033-295X.96.2.341

Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, 48(2), 1–36. Retrieved from <http://www.jstatsoft.org/v48/i02/>

Schimmack, U., & Hartmann, K. (1997). Individual differences in the memory representation of emotional episodes: Exploring the cognitive pro-

cesses in repression. *Journal of Personality and Social Psychology*, 73(5), 1064–1079. doi:10.1037/0022-3514.73.5.1064

Schönbrodt, F. D., & Gerstenberg, F. X. R. (2012). An IRT analysis of motive questionnaires: The Unified Motive Scales. *Journal of Research in Personality*, 46, 725–742. doi:10.1016/j.jrp.2012.08.010

Schönbrodt, F. D., Zygar-Hoffmann, C., Nestler, S., Pusch, S., & Hagemeyer, B. (2019). *Measuring motivational relationship processes in experience sampling: A reliability model for moments, days, and persons nested in couples*. doi:10.31234/osf.io/6mq7t

Shrout, P. E., Stadler, G., Lane, S. P., McClure, M. J., Jackson, G. L., Clavél, F. D., ... Bolger, N. (2017). Initial elevation bias in subjective reports. *Proceedings of the National Academy of Sciences*, 115(1), E15–E23. doi:10.1073/pnas.1712277115

Sprecher, S. (1999). "I love you more today than yesterday": Romantic partners' perceptions of changes in love and related affect over time. *Journal of Personality and Social Psychology*, 76(1), 46–53. doi:10.1037/0022-3514.76.1.46

Thomas, D. L., & Diener, E. (1990). Memory accuracy in the recall of emotions. *Journal of Personality and Social Psychology*, 59(2), 291–297. doi:10.1037/0022-3514.59.2.291

Vaish, A., Grossmann, T., & Woodward, A. (2008). Not all emotions are created equal: The negativity bias in social-emotional development. *Psychological Bulletin*, 134(3), 383–403. doi:10.1037/0033-2909.134.3.383

Walentynowicz, M., Schneider, S., & Stone, A. A. (2018). The effects of time frames on self-report. *PLOS ONE*, 13(8), 1–18. doi:10.1371/journal.pone.0201655

Watson, D., Clark, L., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal of Personality and Social Psychology*, 54(6), 1063–1070. doi:10.1037/0022-3514.54.6.1063

Wells, T., Bailey, J. T., & Link, M. W. (2014). Comparison of smartphone and online computer survey administration. *Social Science Computer Review*, 32(2), 238–255. doi:10.1177/0894439313505829

West, T. V., & Kenny, D. A. (2011). The Truth and Bias Model of Judgment. *Psychological Review*, 118(2), 357–378. doi:10.1037/a0022936

Wickham, H., François, R., Henry, L., & Müller, K. (2018). *Dplyr: A grammar of data manipulation*. Retrieved from <https://CRAN.R-project.org/package=dplyr>

Winkielman, P., Knäuper, B., & Schwarz, N. (1998). Looking back at anger: Reference periods change the interpretation of emotion frequency questions. *Journal of Personality and Social Psychology*, *75*(3), 719–28. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/9781408>

Woike, B. A. (1995). Most-memorable experiences: Evidence for a link between implicit and explicit motives and social cognitive processes in everyday life. *Journal of Personality and Social Psychology*, *68*(6), 1081–1091.

Woike, B. A. (2008). A functional framework for the influence of implicit and explicit motives on autobiographical memory. *Personality and Social Psychology Review*, *12*(2), 99–117. doi:10.1177/1088868308315701

Yap, S. C. Y., Wortman, J., Anusic, I., Baker, S. G., Scherer, L. D., Donnellan, M. B., & Lucas, R. E. (2016). The effect of mood on judgments of subjective well-being: Nine tests of the judgment model. *Journal of Personality and Social Psychology*, *113*(6), 939–961. doi:10.1037/pspp0000115

Zygar, C., Hagemeyer, B., Pusch, S., & Schönbrodt, F. D. (2018a). From motive dispositions to states to outcomes: An intensive experience sampling study on communal motivational dynamics in couples. *European Journal of Personality*, *32*(3), 306–324. doi:10.1002/per.2145

Zygar, C., Hagemeyer, B., Pusch, S., & Schönbrodt, F. D. (2018b). *From motive dispositions to states to outcomes: Research data of an intensive experience sampling study on communal motivational dynamics in couples [Translated title] (Version 2.0.0)* [data and documentation]. Trier: Center for Research Data in Psychology: PsychData of the Leibniz Institute for Psychology Information ZPID. doi:10.5160/psychdata.zrce16dy99_v20000

Zygar-Hoffmann, C., Hagemeyer, B., Pusch, S., & Schönbrodt, F. D. (2020). *A large longitudinal study on motivation, behavior and satisfaction in couples: Research data from a four-week experience sampling study with a pre-, post-, and one-year follow-up-assessment [Translated title] (Version 1.0.0)* [data and documentation]. Trier: Center for Research Data in Psychology PsychData of the Leibniz Institute for Psychology Information ZPID. doi:10.5160/psychdata.zrce18mo99

Chapter 7

General Discussion

This dissertation includes three empirical papers that address the overall question of how states can foster our understanding of couple relationships. The papers rely on two datasets obtained from dyads who participated in Experience Sampling Method (ESM) studies spanning two and four weeks. The conducted research takes advantage of several benefits offered by this dyadic ESM design, such as a large-scale study of intra- and interpersonal processes in everyday life, the possibility to compare dispositions and states, as well as the reduction of retrospective errors and the assessment of partner-reports.

The goal of Paper 1 was to investigate what processes might explain the association between motive dispositions and global relationship satisfaction in the literature, while this effect itself was subject to a replication attempt. The theoretically derived proposed sequence of effects was termed the Dynamics of Motive Satisfaction (DynaMoS) process model and empirically tested for the domain of communion motives on the first data set.

In Paper 2, a first goal was to examine one path of the DynaMoS model more thoroughly: the effect of motivation on subsequent behavior. This was realized by a replication attempt in the second data set and by testing the generalizability of the effect for other motive domains and partner-reports. A second goal of Paper 2 was to illustrate how motivational states compare to motive dispositions in their ability to predict behavior reports.

Finally, Paper 3 zoomed in on the construct of relationship satisfaction as focal outcome in the DynaMoS model and in relationship research more general. The overall goal of Paper 3 was to better understand different measures of relationship

satisfaction. This was achieved by illustrating how different ways of summarizing state relationship satisfaction compares to retrospective and global relationship satisfaction; as well as by demonstrating the retrospective bias occurring during retrospection and the inter-individual influences on this bias.

7.1 Summary of Results

7.1.1 The DynaMoS Model (Paper 1 & Paper 2)

A starting point for the DynaMoS model was the finding in the literature that motive dispositions are related to relationship outcomes, such as global relationship satisfaction (Hagemeyer & Neyer, 2012). For implicit communion motives, Paper 1 presents confirming evidence for this between-person effect, replicating it for three different operationalizations of global relationship satisfaction as outcomes: Individuals with a higher implicit communion motive report a higher global satisfaction with their relationship. It was therefore meaningful to further investigate which processes might drive this effect. Paper 1 reports evidence for all steps of the DynaMoS model that constitute potential explanations for why individuals with a higher implicit communal motive strength are globally more satisfied with their relationship: First, a high implicit communion motive is positively related to the average experience of (explicit) motivational communal states in everyday life (a between-person effect). In other words, individuals with a dispositional high implicit communal motive strength report on average more often to experience the desire to establish closeness with their partner. Second, this desire is positively related to subsequent instrumental behavior to attain this goal (a within-person effect): When individuals experience a stronger communal motivation than what is typical for them, they subsequently report more communal behavior towards their partner (e.g., affection). Paper 2 demonstrates the robustness of this result, by presenting a direct replication of this effect, and illustrating that it holds up when considering the behavior as reported by their partner (in contrast to self-reports of their own behavior). This effect further conceptually replicates for other motive domains: Paper 2 shows that when individuals are more motivated for independence or power, they also report subsequently more corresponding behavior (e.g., spending time for their interests in case of independence motivation, influencing their partner in case of power motivation). As an intermediate con-

clusion, the results illustrate that a higher implicit communal motive disposition is positively related to communal motivation, and that motivation is followed by corresponding behavior; the next crucial step is to establish how this might affect relationship satisfaction.

The DynaMoS model postulates the following “affect amplifying nature” of motivation (a within-couple effect): When individuals are highly motivated and actually do have a relationship experience with their partner that fit their motivation, this boosts the satisfaction they gain from this experience compared to situations when they were less motivated. Again for the communion domain, there was evidence for this effect when considering individuals’ state relationship satisfaction as outcome (but not regarding their mood). In consequence, individuals with a strong implicit communal motive disposition might experience momentary satisfaction with their relationship more often, because a) they are more frequently motivated to experience closeness; b) they more often have close relationship experiences by initiating them frequently, and c) as a result, they more often benefit from the combination of a high motivation and a corresponding fulfillment of their desire in the relationship. Finally, Paper 1 also provides evidence for an association between average state relationship satisfaction and global relationship satisfaction, bridging the gap between state relationship satisfaction as outcome of the affect amplifying effect of motivation and global relationship satisfaction as outcome of the initially considered between-person effect of motive dispositions.

7.1.2 Comparison of Dispositions and States (Paper 2 & Paper 3)

Paper 2 and Paper 3 both address the comparison of constructs assessed at different abstraction levels. In Paper 2, the association between motivational variables and behavior reports is considered not only on a within-person level (as reported above), but also on a between-person level, comparing the predictive value of aggregated motivational states, explicit and implicit motive dispositions. The results show that aggregated motivational states had the most consistent effects across different motive domains, across different classes of reported behavior, and across the two sources of behavior report, additionally showing the largest effect sizes compared to explicit and implicit motive dispositions. Further, the predictive effects of aggregated motivational states on behavior reports were often incremental

to the other two investigated motive dispositions; this was also often the case for explicit motives, but only in one case for implicit motives. The results illustrate the unique variance aggregated motivational states and explicit motives capture for explaining between-person differences in behavior. The additional variance explained by these between-person effects were compared to the additional variance explained by the within-person fixed effects of motivational states predicting behavior: The results show that the amounts of total variance explained by the within-person effects were overall lower than those of the between-person effect of aggregated states (although there was considerable variance of the outcomes on the within-person level).

In contrast to these comparisons in Paper 2 concerning the predictive value of motivational dispositions and states, Paper 3 addressed comparisons between the different assessment modalities of relationship satisfaction. Analyses of different ways of summarizing relationship satisfaction states (e.g., computing the mean, the median, certain quantiles) revealed that the average state relationship satisfaction corresponds best to a global assessment of relationship satisfaction, as well as to a retrospective assessment. Further, comparing the mean-level of retrospection with the averaged states in the sample indicated that on average in retrospection individuals reported to have experienced a higher amount of annoyance by their partner than their average report during the study period. As annoyance is part of the relationship satisfaction scale, this results in a lower total score of relationship satisfaction in retrospection than the average of the relationship satisfaction states reported during the study period. Paper 3 further identifies several moderators of this bias, the most robust and strongest effects stemming from the global level of relationship satisfaction during retrospection: The less satisfied individuals are globally, the stronger is their over-reporting of annoyance in retrospection, and the stronger they additionally under-report other, positive aspects of relationship satisfaction. Other moderators were also identified, which might be in some cases subsumed by a general factor reflecting general positivity or social desirability in responding. Finally, the results also show that after sampling relationship satisfaction states for two weeks, the association with global relationship satisfaction reaches a saturation effect. This points to a sampling of two weeks being necessary for momentary assessments of relationship satisfaction to maximize their representativeness for global relationship satisfaction.

7.2 General Limitations

Next to the limitations specific to each paper, some limitations are inherently based on the study design and therefore concern all papers.

7.2.1 Challenges of the Experience Sampling Method

As outlined in the general introduction, the ESM provides several advantages and benefits for researchers, but is naturally not free of drawbacks, two of which are particularly relevant for the current research. First, a potential problem is reactivity induced by the measurement procedure, a topic on which there is mixed evidence in case of the ESM (see e.g., Conner & Reid, 2012; De Vuyst, Dejonckheere, Van der Gucht, & Kuppens, 2019; Larson & Sbarra, 2015; Merrilees, Goeke-Morey, & Mark Cummings, 2008; Reynolds, Robles, & Repetti, 2016): When individuals reflect on their motivational states, behaviors and relationship satisfaction multiple times per day, this might not only constitute a passive observation of their states, but could have the character of an intervention on any of these variables. When the ESM actually changes the phenomena one is interested in passively observing, this could influence the effects found in the current analyses. For example, the amount of shown behavior might change just because individuals are prompted to reflect on it repeatedly (e.g., show more affection and less criticism towards the partner; but see Merrilees et al., 2008). This would introduce or reduce variance in the occurrence of behavior, changing the variance that would be naturally observed, and could distort the effects found for the prediction of behavior reports by motivational states. The current analyses only control for participated time in the study as a linear main effect on the outcomes (“detrending”, Wang & Maxwell, 2015), which does not consider non-linear reactivity effects, or those not related to the passage of time. Similarly, reactivity concerning the predictor variables and its consequences are neglected, for example, if enhanced attention to one’s motivational states causally produces subsequent behavior just because attention was drawn to the motivation, but the motivation itself would not have caused behavior otherwise. More experimental research investigating the constructs of the current studies is needed to assess the presence and magnitude of reactivity effects, and to consider their consequences for the current results.

Second, the ESM places a high demand on the participants, who are required to have their smartphones available and to answer several questions, several times per day, within a specific time span (45 minutes in the current studies). This can be experienced as a large burden, and might therefore reduce the motivation to respond to the questions carefully across the whole study period. While having some advantages compared to traditional psychological measures, ESM assessments remain reports on questionnaires, which are susceptible to all kinds of measurement errors, with fatigue effects being particularly plausible (see e.g., Reynolds et al., 2016). The consequences might be a reduced variance in the answers due to responding in a routinely manner (losing partly a crucial characteristic of state measures, see Paper 3), and an increased missing of answers in situations in which it is particularly bothersome to respond (e.g., when being around others, during conflicts; but see Silvia et al., 2013). A reduced representativeness of answers could be especially relevant for the comparison of ESM and retrospective assessments, as for example conflicts are situations which plausibly involve higher amounts of annoyance by the partner – having fewer amount of reports in such situations could therefore be the cause of the effect that was now labeled as retrospective bias in Paper 3. Planned missing data designs are proposed in the literature as one way to circumvent some fatigue effects (Silvia, Kwapil, Walsh, & Myin-Germeys, 2014), but they tackle not all associated problems.

7.2.2 Constraints on Generalizability

Simons, Shoda, & Lindsay (2017) proposed that all papers should include a statement characterizing the generalizability of the findings for different participant samples, materials, and contexts. The included papers of this dissertation made such statements less explicitly; hence, the current section provides a short summary of the constraints on generalizability for the presented effects.

The current dissertation includes data from two studies that contain highly similar materials, but differ in the samples the data was obtained from: The participants of Study 1 were mainly unmarried students in their twenties without children, who were committed to their relationships for a short or medium period at the time of participation. In contrast, a third of the participants in Study 2 were married, and one fifth of them had children. Although the educational level was still quite high, most were not students (anymore). Overall, this sample was more

heterogeneous, as there was more variation in age and relationship duration, with the averages being considerably higher than in Study 1 (see Table 2 in Paper 3). In sum, the two samples vary in their representation of age groups, occupations, and certain relationship characteristics (status, duration, and children). However, both samples include only German-speaking, monogamous, heterosexual, overall highly educated couples, so they both represent to some extent this population of couples. Although the samples were obviously not drawn randomly from this population, effects that replicate across both samples show a certain robustness for heterogeneity in demographic characteristics, and lend confidence for the generalizability of the results for this population of couples (who could be in many aspects labeled as WEIRD couples, with the acronym standing for Western, Educated, Industrialized, Rich, and Democratic; Henrich, Heine, & Norenzayan, 2010).

In principle, the DynaMoS model summarizes general assumptions about the functionality of motives and motivation. Hence, the results are not expected to differ for homosexual couples, or less WEIRD individuals. In consequence, the expected target population for the model are individuals in couple relationships in general.¹ However, for some paths of the DynaMoS model, situational influences and temporal considerations were discussed to be theoretically meaningful (see Paper 1), but not directly tested in the presented papers. For example, for the effect of motivation on behavior, it is highly relevant if situational circumstances leave the freedom, and provide the incentives to show different kinds of behaviors. Some couple characteristics (e.g., the presence of children, living arrangements, long-distance relationships) strongly influence everyday situations, constraining the own scope of action and the available incentives. It would therefore be expected for the results of the DynaMoS model to be moderated by sample characteristics that involve influences on such situational characteristics. Similarly, the results may not generalize to other time spans than the ones investigated in the current studies (i.e., the temporal dynamics of the emergence and persistence of the investigated effects are mostly unclear).

¹A recent ManyLabs study (Klein et al., 2018) shows in exploratory analyses that the heterogeneity of effects for classical findings in psychology is mostly not attributable due to the degree in which samples could be described as WEIRD. Although this was no systematic investigation of the effect of WEIRD samples, this tentatively provides confidence to expect psychological effects to generalize beyond the commonly investigated WEIRD samples (as it was the case in the current studies), unless a theoretical explanation would predict that demographic or cultural characteristics influence the investigated effects. In contrast, Yarkoni (2019) recently argued that researchers should be much more conservative about making inferences beyond the investigated data.

In Paper 3 an explorative analysis was conducted to show which person characteristics moderate retrospective biases, with the demographics of age, gender and relationship duration turning out not to be significant. Based on these results, the retrospective biases would be expected to generalize to the represented population of individuals. As there is work on retrospective biases of other constructs being moderated by age (Neubauer, Scott, Sliwinski, & Smyth, 2019), or relationship duration (in combination with psychological characteristics; Fletcher & Kerr, 2010), the present results might not generalize to older participants in even longer relationships than those sampled in Study 2. For all other results of Paper 3, there is no known reason to expect them to be influenced by demographic sample characteristics, so the target population would be partnered adults in general.

Regarding the materials, a constraint on generalizability is indicated for Paper 3: Here, the main effect of a retrospective mean-level bias only occurred for the negatively framed annoyance item. Empirical evidence on memory biases in emotion reports suggests that it might be due to the negative affectivity captured in this item, so the results might only generalize for relationship satisfaction scales that include items with negative affect. Apart from that, the materials used in the studies are quite interchangeable with similar materials as long as they represent reliable and valid measures of the constructs that were described in the papers. This logic was applied throughout all papers by often using multiple operationalizations of the same construct, and testing the effects for all of them (while controlling the false discovery rate). In most of these cases, the preregistrations explicitly state that effects are expected to occur for all operationalizations of the construct (e.g., different measures of global relationship satisfaction). In general, the effects are expected to be stronger for relationship-specific, or outcome-specific measures compared to more global measures (e.g., in the case of general versus partner-related motives).

Currently, there is no reason to believe that the results depend on other characteristics of the participants, materials, or context. Further, the DynaMoS model aims to be a model that is valid for different motive domains, even though mostly only communion motives were tested.

7.2.3 Replicability of Effects

Although two data sets are available, not all hypotheses were tested on both data sets; only Paper 3 integrated both data sources in the analyses (thereby replicating the relationship between mean relationship states and global relationship satisfaction of the DynaMoS model discovered in the first data set in Paper 1 for the second data set). Only one other path of the DynaMoS model was also investigated in both data sets and published so far (separately in Papers 1 and 2): the main effect of motivation on behavior. This was simultaneously the only path related to motivation that was tested for the communion as well as the agency motive domain. High confidence can therefore be placed in the robustness of this effect across motive domains, and the generalizability of the result for the population, which the participants of both samples represent.

However, the results of the other two paths of the DynaMoS model (the relationship between motive dispositions and motivation, and the affect-amplifying effect of motivation) were only published based on analyses on the first data set, and only for the domain of communion motives. Even though this is not published yet, both paths were preliminarily analyzed for agency motives in the first data set, and for communion motives in the second data set. Crucially, the results of these analyses are not always as expected. First, while the relationship between motive dispositions and motivation replicates in the second data set for the domain of communion (although with a considerably smaller effect size), the affect amplification effect could not be directly replicated for the same operationalization of satisfying experiences as in the first dataset (see conference presentation, Zygar et al., 2018c).² Moreover, for the domain of agency, preliminary analyses on the first data set showed neither consistent effects for the path from motive dispositions to motivation, nor for the affect amplification path (see conference presentation, Zygar et al., 2018d).³ As the analyses for agency motives were not yet run on the second data set (which has a considerably larger sample size), a conclusion about a lacking generalizability of the DynaMos model to agency motives would be premature. Still, overall, these analyses demonstrate that not all of the results presented in the dissertation replicate across both sources of available data or extent to agency motives, and should therefore await further replication

²New analyses show however a conceptual replication for a different operationalization.

³These analyses do however show a replication of the between-person effect of implicit agency motives predicting global relationship satisfaction.

before they can be treated as robust effects.

7.3 Contributions and Future Directions

The current work makes unique contributions for basic motivational research, for applied psychological relationship research, and for the understanding of psychological measures. Beyond the advances in these specific psychological research domains, a general contribution is made for the field of psychology: In the era of the replication crisis (Open Science Collaboration, 2015), psychology requires transparent research that can be verified and readily reproduced, that provides necessary information to replicate and build upon this work, and that enables access to high quality data that can be used for secondary analyses. This dissertation adds to the growing amount of research that follows these criteria, by presenting (partly) preregistered work that can be easily reproduced with the provided scientific use-files and R scripts that accompany each paper; by providing the study materials necessary to repeat and advance this research; by transparently describing null-results; and by providing thoroughly documented data sets, that contain a manifold of assessed variables, being useful for research questions beyond the ones that were envisioned a priori. Additional to these contributions for research, this dissertation provides insights that can be considered by therapists and counselors who work with couples. Both from a research and a practical perspective, there are fruitful avenues for future research based on the results presented in this dissertation.

7.3.1 Motivational Psychology

For the domain of motivational psychology, the current dissertation builds upon the rich theoretical work in that field and empirically contributes to it by a) replicating prior research, thereby showing the robustness of these results in the literature, b) demonstrating that although implicit motives are considered to be unconsciously represented, they seem to partly manifest in explicit motivational states, c) testing basic assumptions about the functionality of motives translated to the within-person level, d) systematically comparing the effects of different motivational variables for the prediction of behavior reports.

First, the successful replication of the link between implicit communal motive

dispositions and relationship satisfaction underlines the relevance of motivational psychology for understanding couple relationships. The motivational predispositions individuals bring to a relationship can be one important source of explanation for how some intra- and interpersonal processes unfold in everyday couple life.

Second, the fact that our results show that implicit communion motives are associated with the average motivational communal motivational state reported during ESM is important from several perspectives. In a first step, it confirms the correspondence that is assumed to be at place between motive dispositions and motivational states: Although situational influences were not considered in our analyses (which should modulate whether motivational appetite or aversion arises), theory would predict that across situations individuals with a higher motive disposition experience motivational appetite faster and more often. Moreover, the result strengthens the confidence in the implicit measure that was applied: Implicit motives and motivation were associated, although the assessment modalities were completely different (content coded imagined stories vs. momentary self-reports); this could be considered as evidence for the validity of the implicit motive measure. It is important to note that it is not self-evident that implicit measures translate to explicit motivational states. For example, Bischof (2008) argues that motivational states only become consciously represented when situational constraints interfere with an implementation of the motivation in corresponding behavior. Otherwise, motivational states are assumed to operate outside of consciousness. The current results do not falsify this assumption, because situational barriers were not considered, which constitutes an important avenue for future research. They do however challenge the notion that motivational states operate mostly outside of conscious awareness; maybe individuals just do not reflect on them unless they are prompted to do so (as during an ESM study). Whether it is desirable for individuals to deliberately attend to their motivation more often is a subsequent question, which relates to research on self-attention (see e.g., Harrington & Lofredo, 2011), and could be potentially relevant for interventions (see Korotitsch & Nelson-Gray, 1999; e.g., for mindfulness training, Atkinson, 2013).

Third, motivational theories are rich of functional explanations of motives (e.g., McClelland, 1987). In the present research, their proposed influence on behavior and their affect amplifying effect were investigated. Crucially, the current research tested them on the within-person level, the level on which they are originally formulated and on which they plausibly occur as mechanisms. For motivational

psychology to further advance, it is necessary that all theoretical predictions that refer to mechanisms come under scrutiny on their corresponding analysis level. Otherwise, the field misses out on taking full advantage of the theoretical explanations about processes, as between-person analyses only give first hints about their validity.

Last, comparing different motivational variables on the between-person level reveals which measures are especially useful and provide incremental value for explaining between-person differences in the occurrence of behavior. The use of partner-reports as they were applied in the current work demonstrates the robustness of results beyond biases that are uniquely associated with self-reports (see e.g., Backer-Fulghum & Sanford, 2015). The results of Paper 2 show that aggregated states and explicit motives provided in many cases unique contributions to the prediction of behavior reports, but the incremental validity of implicit motives was not indicated or negligible.⁴ This raises the question whether the costly method of assessing implicit measures is justified (especially, because explicit motives are related to relationship outcomes as well, see e.g., Hagemeyer et al., 2013b). However, there is evidence that the congruence of explicit and implicit motives matters for relationship outcomes and emotional well-being (Hagemeyer et al., 2013a; Schüler, Job, Fröhlich, & Brandstätter, 2008), and proponents of implicit motives would argue that they are more relevant for regulation processes or nonverbal, uncontrolled behavior than for the behavior that was assessed in the current studies (see e.g., Schultheiss, 2008; Schultheiss & Brunstein, 2010; Schultheiss et al., 2010 for overviews; or Hagemeyer, Dufner, & Denissen, 2016; McAdams, Jackson, & Kirshnit, 1984 for empirical work). The subsequent question would then be whether these subtler behavioral outcomes are incrementally relevant for processes that lead to satisfaction in relationships, and how big their corresponding effect sizes are in comparison. In any case, motive dispositions provide an important, additional perspective on relationships in general, but the jury is still out on the added value of implicit motives compared to explicit motives for the topics investigated in the presented papers.

⁴This was also the case for the prediction of aggregated motivation beyond explicit motives in Paper 1, as well as in other research conducted on this data that looked at the effects of explicit and implicit motives simultaneously (Pusch, Schönbrodt, Zygar-Hoffmann, & Hagemeyer, 2020).

7.3.2 Relationship Research and Psychological Methods

The literature on relationships yields many insights about the behaviors, interactions, and experiences that are supposed to foster or impair relationship satisfaction (Bradbury et al., 2000; Finkel et al., 2017). If shown to be causal, this offers for instance valuable information about which interpersonal dynamics are a source of dissatisfaction and thus harmful for relationships, and with what behavioral habits they can be substituted to promote or re-establish satisfaction. This perspective illustrates how important behaviors look like, but not why they occur in the first place, or when they are especially beneficial. The motivational approach provided by the current work to the study of interpersonal processes in couple relationships helps in understanding the origin of behavior (Winter et al., 1998). By combining this perspective with data from the ESM, this dissertation complements prior work in the field with results from a large-scale investigation of motivational sources of diverse behaviors, pinpointing the motivational circumstances in which these behaviors are particularly relevant for feelings of satisfaction.

The thorough empirical investigation of the construct of relationship satisfaction in the current dissertation also contributes to the field of relationship research. In particular, the presented research provides evidence for the perspective that global and retrospective evaluations of relationship satisfaction are similar, but still distinct to the average of reported relationship satisfaction states (conceptualized as feelings reflecting individuals' relationship mood, need satisfaction, and annoyance). Hence, these different assessments may have different purposes for different research questions (and can therefore also differ in their predictive validities, see Oishi & Sullivan, 2006). With retrospective assessments of annoyance being particularly far away from the reported momentary experiences, this may have implications for conceptualizing what aspects of relationship satisfaction are more strongly rooted in everyday experiences and what aspects are additionally strongly globally influenced by other things like general beliefs about relationships.⁵ As the current work identified which relationship and person characteristics contribute to or counterbalance a mean-level bias in retrospective evaluations, researchers can use this information to investigate the potential (dys-)functional

⁵While having the constraint in mind that individuals might not respond to ESM questions in a state of high annoyance by their partner, thereby capturing these experiences with retrospective evaluations while they are missed by the ESM.

foundation of the uncovered bias and its moderators (see e.g., Fletcher, 2015 for the functionality of other perceptual biases). More generally, the illustrated exploratory analyses may be a first step to derive theoretical assumptions about how relationship satisfaction states interact with relationship, person and situation characteristics to result in a global evaluation of what is called relationship satisfaction, so that future research could test them in a confirmatory fashion.

From an assessment perspective, it is crucial to understand what biases occur when individuals complete self-report measures and whether they can be taken into account (see e.g., Gnambs, 2015 for a classification of measurement biases in self-reports of the Big Five, and how they relate to reliability estimation). The current work provides evidence for the potential presence of a retrospective bias in the evaluation of relationship satisfaction, which underlines the value of assessing relationship satisfaction in the moment. It further equips researchers with guidelines on when relationship satisfaction states get representative of global evaluations, making a practical contribution to the design of ESM relationship research. However, this raises the question of the reliability of the ESM measures themselves (see a discussion by Schimmack, 2003 on how the aggregation of state measures can lead to accumulated systematic measurement errors)⁶. Another perspective is to not consider the difference between retrospection and averaged states as measurement error: Both measures could represent different conceptualizations of relationship satisfaction, or at least assume different measurement models, as retrospective evaluations could comprise a more complex integration of the available information than taking the average and thus giving each experience equal weight.⁷ To that end, as a status diagnostic, it has to be acknowledged that differences in the compared measures are present, and that they could be interpreted either as measurement error, or as a desired characteristic contributing to the measures' validity. This has implications for work that aims to reduce these differences by taking them into account in a measurement model, or correct for them in subsequent analyses, similar to response styles (see e.g., Diener, Smith, & Fujita, 1995 for an investigation of the association between positive and negative affect after accounting for memory and other method biases, by integrating infor-

⁶Reliability estimates on different levels for the motivation and state relationship satisfaction ESM measures used in the current work are the focus of another paper that is not part of this dissertation (Schönbrodt, Zygar-Hoffmann, Nestler, Pusch, & Hagemeyer, 2019).

⁷However, the retrospective bias replicated also for the median of relationship satisfaction states, as well as for the mean over the last week or last day, so at least these alternatives of summarizing the states are also not valid with regard to what happens during retrospection.

mation assessed with retrospective self-reports, ESM assessments, and informant reports).

7.3.3 Practitioner Settings

Before implications for practice can be drawn from the DynaMoS model, the replicability and the causality of the proposed processes has to be shown first (see also Stanley, Bradbury, & Markman, 2000, for raising caution when trying to translate results from basic research to interventions for couples). Schmiedek & Neubauer (2019) recently introduced the within-person encouragement design, which describes how to go about interventions focused on the intra-individual level with experimental manipulations, random encouragements, and appropriate ways to analyze the resulting data. Such an approach provides a promising way to experimentally study the DynaMoS model on the within-person level in future research. So far, the longitudinal data provides only first, tentative evidence on how motivational states in the domain of communion might influence the joy individuals derive from certain experiences in their relationship.

If shown to be robust and causal, potential practical implications from the model could be envisioned. Interventions could aim at modulating individuals' motivational experiences, but as these are supposed to inherently stem from their stable motive dispositions, other avenues might be more fruitful: For instance, therapists could help individuals to draw more attention to their motivation, to effectively communicate their desires, and to identify the idiosyncratic behavioral implementations that are especially effective in resulting in satisfying experiences with the partner; or accordingly those behaviors that are ineffective or result in experiences that are highly frustrating (see e.g., Benson, McGinn, & Christensen, 2012). Further, identifying the motives that are important for both partners in a relationship might help in guiding couples to the domains in which they should be particularly sensitive to their own and each other's desires. Deliberately arranging corresponding behavioral routines might be one way to help satisfying these motives on a regular basis (e.g., scheduling (more) quality time together for those with strong communion motives, but scheduling (more) individual time for those with strong independence motives, bearing in mind that although the current research only considered one motive domain at a time, individuals have several motives that have to be considered simultaneously, see e.g. Kluwer, Karremans, Riedijk, & Knee, 2019). Such approaches should further take into account the

motivation of both partners: The current work only looked at the motivation of one individual, but couple relationships are a system consisting of two individuals with potentially conflicting desires (see Righetti & Impett, 2017 for a review). How individuals negotiate their desires and the influence of different strategies on relationship satisfaction is an important area of adjacent research (see Zygar et al., 2018b for a first approach to study how such situations are typically resolved in the case of conflicting communal motivation; and Impett, Muise, & Harasymchuk, 2019 for detrimental effects of neglecting one's own sexual desires for pleasing the partner's sexual desires).

Moreover, the work presented on retrospective errors in relationship satisfaction evaluations yields valuable information about the recollections couple therapists will encounter on a regular basis (e.g., when asking about relationship satisfaction since the last therapy session for example to track changes, see Benson et al., 2012; Halford et al., 2012; Halford, Keefer, & Osgarby, 2002). As the current work suggests that such recollections include a special (over-)emphasis on experiences of annoyance particularly for globally unhappy individuals, therapists might aim at putting them more into perspective, for instance with cognitive techniques (Dattilio, 1993), which are already a central part of several therapy types (e.g., "cognitive restructuring" in cognitive-behavior couple therapy, Dattilio & Epstein, 2005; Epstein & Zheng, 2017; "exception questions" in solution-focused brief therapy, Yu, 2018).

It should be emphasized that the current research was not designed to empirically test the effect of any interventions on relationship outcomes. One potential avenue for future research would be to translate the provided insights into clinical or counseling settings, and to investigate their effectiveness (see Stanley et al., 2019 for a recent review of best practices and aspects to consider when evaluating relationship interventions). Further, the current research singled out specific aspects of relationship processes for single motivational domains, focusing on the intra-individual processes, which are only one important aspect for therapy (see e.g., Heatherington, Friedlander, & Greenberg, 2005). Individuals, their relationships, and their situational circumstances are far more complex than that and must be considered from a more holistic perspective when trying to find effective ways to promote individuals' satisfaction in a relationship (a central element of systemic thinking, see Stanton, 2009; Stanton & Welsh, 2012).

7.4 Conclusion

The generated insights about everyday motivational processes in the current dissertation were made possible by the repeated state assessments in individuals' natural environments, realized with an intensive longitudinal ESM study design. Establishing the empirical evidence for such processes is fundamental to arrive at explanations on what drives different relationship outcomes. In sum, what have we learned from these state measures about couple relationships, after two studies, three papers, and a lot of empirical results? We gained insights about the empirical validity of some of the theoretical foundations in motivational psychology that lay the groundwork for many applied research areas, and know more about the psychological and situational circumstances that give rise to momentary and global relationship satisfaction. We also gained a better understanding of the assessment modalities that are commonly used in motivational and relationship research. This dissertation is therefore a building block in illustrating the role of psychological characteristics for the functioning of couple relationships. In a broader sense, the transparency and openness of the current research makes it possible for researchers to re-analyze this work from different perspectives and to eventually draw own, possibly different conclusions – building the basis for fruitful and constructive scientific discourse.

Chapter 8

References

- Anderson, J. R., Van Ryzin, M. J., & Doherty, W. J. (2010). Developmental trajectories of marital happiness in continuously married individuals: A group-based modeling approach. *Journal of Family Psychology, 24*(5), 587–596. <https://doi.org/10.1037/a0020928>
- Atkinson, B. J. (2013). Mindfulness training and the cultivation of secure, satisfying couple relationships. *Couple and Family Psychology: Research and Practice, 2*(2), 73–94. <https://doi.org/10.1037/cfp0000002>
- Atkinson, J. W. (1957). Motivational determinants of risk-taking behavior. *Psychological Review, 64*(6), 359–372.
- Back, M. D., & Vazire, S. (2015). The social consequences of personality: Six suggestions for future research. *European Journal of Personality, 29*(2), 296–307. <https://doi.org/10.1002/per.1998>
- Backer-Fulghum, L. M., & Sanford, K. (2015). The validity of retrospectively reported conflict interactions in couples. *Journal of Family Psychology, 29*(2), 253–262. <https://doi.org/10.1037/fam0000057>
- Baird, B. M., Lucas, R. E., & Donnellan, M. B. (2017). The role of response styles in the assessment of intraindividual personality variability. *Journal of Research in Personality, 69*, 170–179. <https://doi.org/10.1016/j.jrp.2016.06.015>
- Baucom, B. R., Dickenson, J. A., Atkins, D. C., Baucom, D. H., Fischer, M. S., Weusthoff, S., ... Zimmermann, T. (2015). The interpersonal process model of demand/withdraw behavior. *Journal of Family Psychology, 29*(1), 80–90.

<https://doi.org/10.1037/fam0000044>

- Benson, L. A., McGinn, M. M., & Christensen, A. (2012). Common principles of couple therapy. *Behavior Therapy, 43*(1), 25–35. <https://doi.org/10.1016/j.beth.2010.12.009>
- Bescheid, E., & Lopes, J. (1997). A temporal model of relationship satisfaction and stability. In R. J. Sternberg & M. Hojjat (Eds.), *Satisfaction in close relationships* (pp. 129–159). Guildford Press.
- Bischof, N. (1975). A systems approach toward the functional connections of attachment and fear. *Child Development, 46*(4), 801–817. <https://doi.org/10.1111/j.1467-8624.1975.tb04024.x>
- Bischof, N. (1995). *Struktur und Bedeutung: Eine Einführung in die Systemtheorie*. Bern: Huber.
- Bischof, N. (2008). *Psychologie: Ein Grundkurs für Anspruchsvolle*. Stuttgart: Kohlhammer.
- Bolger, N., Davis, A., & Rafaeli, E. (2003). Diary methods: Capturing life as it is lived. *Annual Review of Psychology, 54*, 579–616. <https://doi.org/10.1146/annurev.psych.54.101601.145030>
- Bolger, N., & Laurenceau, J.-P. (2013). *Intensive longitudinal methods*. New York, NY: Guilford.
- Bolger, N., Zee, K. S., Rossignac-Milon, M., & Hassin, R. R. (2019). Causal processes in psychology are heterogeneous. *Journal of Experimental Psychology: General, 148*(4), 601–618. <https://doi.org/10.1037/xge0000558>
- Bradbury, T. N., Fincham, F. D., & Beach, S. R. (2000). Research on the nature and determinants of marital satisfaction: A decade in review. *Journal of Marriage and Family, 62*(4), 964–980. <https://doi.org/10.1111/j.1741-3737.2000.00964.x>
- Bühler, J. L., Weidmann, R., Nikitin, J., & Grob, A. (2019). A closer look at life goals across adulthood: Applying a developmental perspective to content, dynamics, and outcomes of goal importance and goal attainability. *European Journal of Personality, 33*(3), 359–384. <https://doi.org/10.1002/per.2194>
- Caughlin, J. P., & Vangelisti, A. L. (2000). An individual difference explanation

- of why married couples engage in the demand/withdraw pattern of conflict. *Journal of Social and Personal Relationships*, *17*(4-5), 523–551. <https://doi.org/10.1177/0265407500174004>
- Conner, T. S., & Reid, K. A. (2012). Effects of intensive mobile happiness reporting in daily life. *Social Psychological and Personality Science*, *3*(3), 315–323. <https://doi.org/10.1177/1948550611419677>
- Cooper, A. N., Totenhagen, C. J., McDaniel, B. T., & Curran, M. A. (2018). Volatility in daily relationship quality: The roles of attachment and gender. *Journal of Social and Personal Relationships*, *35*(3), 348–371. <https://doi.org/10.1177/0265407517690038>
- Cui, M., Lorenz, F. O., Conger, R. D., Melby, J. N., & Bryant, C. M. (2005). Observer, self-, and partner reports of hostile behaviors in romantic relationships. *Journal of Marriage and Family*, *67*(5), 1169–1181. <https://doi.org/10.1111/j.1741-3737.2005.00208.x>
- Dattilio, F. M. (1993). Cognitive techniques with couples and families. *The Family Journal: Counseling and Therapy for Couples and Families*, *1*(1), 51–65. <https://doi.org/10.1177/106648079300100108>
- Dattilio, F. M., & Epstein, N. B. (2005). Introduction to the special section: The role of cognitive-behavioral interventions in couple and family therapy. *Journal of Marital and Family Therapy*, *31*(1), 7–13. <https://doi.org/10.1111/j.1752-0606.2005.tb01539.x>
- Deci, E., & Ryan, R. (2000). The "what" and "why" of goal pursuits: Human needs and the self-determination of behavior. *Psychological Inquiry*, *11*(4), 227–268. https://doi.org/10.1207/S15327965PLI1104_01
- De Vuyst, H. J., Dejonckheere, E., Van der Gucht, K., & Kuppens, P. (2019). Does repeatedly reporting positive or negative emotions in daily life have an impact on the level of emotional experiences and depressive symptoms over time? *PLoS ONE*, *14*(6), 1–12. <https://doi.org/10.1371/journal.pone.0219121>
- Diener, E., Smith, H., & Fujita, F. (1995). The personality structure of affect. *Journal of Personality and Social Psychology*, *69*(1), 130–141. <https://doi.org/10.1037/0022-3514.69.1.130>
- Eastwick, P. W., Finkel, E. J., & Simpson, J. A. (2019). Relationship trajectory

- ries: A meta-theoretical framework and theoretical applications. *Psychological Inquiry*, 30(1), 1–28. <https://doi.org/10.1080/1047840X.2019.1577072>
- Epstein, N. B., & Zheng, L. (2017). Cognitive-behavioral couple therapy. *Current Opinion in Psychology*, 13, 142–147. <https://doi.org/10.1016/j.copsyc.2016.09.004>
- Eurostat. (2019). *Crude marriage and divorce rates, eu-28, 1965-2016*. Retrieved from https://ec.europa.eu/eurostat/statistics-explained/index.php/Marriage_and_divorce_statistics#Fewer_marriages.2C_more_divorces
- Fincham, F. D., Rogge, R., & Beach, S. R. H. (2018). Relationship satisfaction. In A. L. Vangelisti & D. Perlman (Eds.), *The cambridge handbook of personal relationships* (pp. 422–436). <https://doi.org/10.1017/9781316417867.033>
- Finkel, E. J., Simpson, J. A., & Eastwick, P. W. (2017). *The psychology of close relationships: Fourteen core principles*. 383–411. <https://doi.org/10.1146/annurev-psych-010416-044038>
- Fisher, A. J., Medaglia, J. D., & Jeronimus, B. F. (2018). Lack of group-to-individual generalizability is a threat to human subjects research. *Proceedings of the National Academy of Sciences*, 115(27), E6106–E6115. <https://doi.org/10.1073/pnas.1711978115>
- Fleeson, W. (2001). Toward a structure-and process-integrated view of personality: Traits as density distributions of states. *Journal of Personality and Social Psychology*, 80(6), 1011–1027. <https://doi.org/10.1037/0022-3514.80.6.1011>
- Fleeson, W., & Gallagher, P. (2009). The implications of Big Five standing for the distribution of trait manifestation in behavior: Fifteen experience-sampling studies and a meta-analysis. *Journal of Personality and Social Psychology*, 97(6), 1097–1114. <https://doi.org/10.1037/a0016786>
- Fletcher, G. J. (2015). Accuracy and bias of judgments in romantic relationships. *Current Directions in Psychological Science*, 24(4), 292–297. <https://doi.org/10.1177/0963721415571664>
- Fletcher, G. J., & Kerr, P. S. (2010). Through the eyes of love: Reality and illusion in intimate relationships. *Psychological Bulletin*, 136(4), 627–658. <https://doi.org/10.1037/a0019792>
- Funk, J. L., & Rogge, R. D. (2007). Testing the ruler with item response the-

- ory: Increasing precision of measurement for relationship satisfaction with the Couples Satisfaction Index. *Journal of Family Psychology*, *21*(4), 572–583. <https://doi.org/10.1037/0893-3200.21.4.572>
- Geiser, C., Götz, T., Preckel, F., & Freund, P. A. (2017). States and traits: Theories, models, and assessment. *European Journal of Psychological Assessment*, *33*(4), 219–223. <https://doi.org/10.1027/1015-5759/a000413>
- Geukes, K., Nestler, S., Hutteman, R., Küfner, A. C. P., & Back, M. D. (2017). Trait personality and state variability: Predicting individual differences in within- and cross-context fluctuations in affect, self-evaluations, and behavior in everyday life. *Journal of Research in Personality*, *69*, 124–138. <https://doi.org/10.1016/j.jrp.2016.06.003>
- Gillath, O., Hart, J., Nofhle, E. E., & Stockdale, G. D. (2009). Development and validation of a state adult attachment measure (SAAM). *Journal of Research in Personality*, *43*(3), 362–373. <https://doi.org/10.1016/j.jrp.2008.12.009>
- Gnambs, T. (2015). Facets of measurement error for scores of the Big Five: Three reliability generalizations. *Personality and Individual Differences*, *84*, 84–89. <https://doi.org/10.1016/j.paid.2014.08.019>
- Gottman, J. M. (1998). Psychology and the study of marital processes. *Annual Review of Psychology*, *49*, 169–197. <https://doi.org/10.1146/annurev.psych.49.1.169>
- Gottman, J. M., & Krokoff, L. J. (1989). Marital interaction and satisfaction: A longitudinal view. *Journal of Consulting and Clinical Psychology*, *57*(1), 47–52. <https://doi.org/10.1037/0022-006x.57.1.47>
- Gottman, J. M., & Notarius, C. I. (2000). Decade review: Observing marital interaction. *Journal of Marriage and Family*, *62*(4), 927–947. <https://doi.org/10.1111/j.1741-3737.2000.00927.x>
- Gómez-López, M., Viejo, C., & Ortega-Ruiz, R. (2019). Well-being and romantic relationships: A systematic review in adolescence and emerging adulthood. *International Journal of Environmental Research and Public Health*, *16*(2145), 1–31. <https://doi.org/10.3390/ijerph16132415>
- Graham, J. W., & Donaldson, S. I. (1993). Evaluating interventions with differential attrition: The importance of nonresponse mechanisms and use of follow-up

- data. *Journal of Applied Psychology*, 78(1), 119–128. <https://doi.org/10.1037/0021-9010.78.1.119>
- Haak, E. A., Keller, P. S., & DeWall, C. N. (2017). Daily variations in attachment anxiety and avoidance: A density distributions approach. *Journal of Research in Personality*, 69, 218–224. <https://doi.org/10.1016/j.jrp.2016.08.002>
- Haefffel, G. J., & Howard, G. S. (2010). Self-report: Psychology's four-letter word. *American Journal of Psychology*, 123(2), 181–188. <https://doi.org/10.5406/amerjpsyc.123.2.0181>
- Hagemeyer, B., Dufner, M., & Denissen, J. J. A. (2016). Double dissociation between implicit and explicit affiliative motives: A closer look at socializing behavior in dyadic interactions. *Journal of Research in Personality*, 65, 89–93. <https://doi.org/10.1016/j.jrp.2016.08.003>
- Hagemeyer, B., Neberich, W., Asendorpf, J. B., & Neyer, F. J. (2013a). (In)Congruence of implicit and explicit communal motives predicts the quality and stability of couple relationships. *Journal of Personality*, 81(4), 390–402. <https://doi.org/10.1111/jopy.12016>
- Hagemeyer, B., & Neyer, F. J. (2012). Assessing implicit motivational orientations in couple relationships: The Partner-Related Agency and Communion Test (PACT). *Psychological Assessment*, 24(1), 114–128. <https://doi.org/10.1037/a0024822>
- Hagemeyer, B., Neyer, F. J., Neberich, W., & Asendorpf, J. B. (2013b). The ABC of Social Desires: Affiliation, being alone, and closeness to partner. *European Journal of Personality*, 27(5), 442–457. <https://doi.org/10.1002/per.1857>
- Halford, W. K., Hayes, S., Christensen, A., Lambert, M., Baucom, D. H., & Atkins, D. C. (2012). Toward making progress feedback an effective common factor in couple therapy. *Behavior Therapy*, 43(1), 49–60. <https://doi.org/10.1016/j.beth.2011.03.005>
- Halford, W. K., Keefer, E., & Osgarby, S. M. (2002). "How has the week been for you two" Relationship satisfaction and hindsight memory biases in couples' reports of relationship events. *Cognitive Therapy and Research*, 26(6), 759–773. <https://doi.org/10.1023/A:1021289400436>
- Halford, W. K., Rhoades, G., & Morris, M. (2018). Effects of the parents'

- relationship on children. In M. R. Sanders & A. Morawska (Eds.), *Handbook of parenting and child development across the lifespan* (pp. 97–120). https://doi.org/10.1007/978-3-319-94598-9_5
- Harrington, R., & Loffredo, D. A. (2011). Insight, rumination, and self-reflection as predictors of well-being. *Journal of Psychology: Interdisciplinary and Applied*, *145*(1), 39–57. <https://doi.org/10.1080/00223980.2010.528072>
- Heatherington, L., Friedlander, M. L., & Greenberg, L. (2005). Change process research in couple and family therapy: Methodological challenges and opportunities. *Journal of Family Psychology*, *19*(1), 18–27. <https://doi.org/10.1037/0893-3200.19.1.18>
- Hektner, J. M., Schmidt, J. A., & Csikszentmihalyi, M. (2007). *Experience sampling method: Measuring the quality of everyday life*. Sage Publications, Inc.
- Henrich, J., Heine, S. J., & Norenzayan, A. (2010). Most people are not WEIRD. *Nature*, *466*(7302), 29.
- Hoffman, L., & Stawski, R. S. (2009). Persons as contexts: Evaluating between-person and within-person effects in longitudinal analysis. *Research in Human Development*, *6*(2-3), 97–120. <https://doi.org/10.1080/15427600902911189>
- Hofmann, W., Finkel, E. J., & Fitzsimons, G. M. (2015). Close relationships and self-regulation: How relationship satisfaction facilitates momentary goal pursuit. *Journal of Personality and Social Psychology*, *109*(3), 434–452. <https://doi.org/10.1037/pspi0000020>
- Impett, E. A., Muise, A., & Harasymchuk, C. (2019). Giving in the bedroom: The costs and benefits of responding to a partner’s sexual needs in daily life. *Journal of Social and Personal Relationships*, *36*(8), 2455–2473. <https://doi.org/10.1177/0265407518787349>
- Karney, B. R., & Bradbury, T. N. (1995). The longitudinal course of marital quality and stability: A review of theory, method and research. *Psychological Bulletin*, *118*(1), 3–34.
- Kievit, R. A., Frankenhuis, W. E., Waldorp, L. J., & Borsboom, D. (2013). Simpson’s paradox in psychological science: A practical guide. *Frontiers in Psychology*, *4*(513), 1–14. <https://doi.org/10.3389/fpsyg.2013.00513>
- Klein, R. A., Vianello, M., Hasselman, F., Adams, B. G., Adams, R. B., Alper, S.,

- ... Nosek, B. A. (2018). Many Labs 2: Investigating variation in replicability across samples and settings. *Advances in Methods and Practices in Psychological Science*, *1*(4), 443–490. <https://doi.org/10.1177/2515245918810225>
- Kluwer, E. S., Karremans, J. C., Riedijk, L., & Knee, C. R. (2019). Autonomy in relatedness: How need fulfillment interacts in close relationships. *Personality and Social Psychology Bulletin*. <https://doi.org/10.1177/0146167219867964>
- Korotitsch, W. J., & Nelson-Gray, R. O. (1999). An overview of self-monitoring research in assessment and treatment. *Psychological Assessment*, *11*(4), 415–425. <https://doi.org/10.1037/1040-3590.11.4.415>
- Kritzler, S., Krasko, J., & Luhmann, M. (2019). *Inside the happy personality: Personality states, situation experience, and state affect mediate the relation between personality and affect*. <https://doi.org/10.31234/osf.io/nrydp>
- Larson, G. M., & Sbarra, D. A. (2015). Participating in research on romantic breakups promotes emotional recovery via changes in self-concept clarity. *Social Psychological and Personality Science*, *6*(4), 399–406. <https://doi.org/10.1177/1948550614563085>
- Laurenceau, J.-P., Barrett, L. F., & Pietromonaco, P. R. (1998). Intimacy as an interpersonal process: The importance of self-disclosure, partner disclosure, and perceived partner responsiveness in interpersonal exchanges. *Journal of Personality and Social Psychology*, *74*(5), 1238–1251. <https://doi.org/10.1037/0022-3514.74.5.1238>
- Laurenceau, J.-P., Barrett, L. F., & Rovine, M. J. (2005a). The interpersonal process model of intimacy in marriage: A daily-diary and multilevel modeling approach. *Journal of Family Psychology*, *19*(2), 314–323. <https://doi.org/10.1037/0893-3200.19.2.314>
- Laurenceau, J.-P., & Bolger, N. (2005). Using diary methods to study marital and family processes. *Journal of Family Psychology*, *19*(1), 86–97. <https://doi.org/10.1037/0893-3200.19.1.86>
- Laurenceau, J.-P., Troy, A. B., & Carver, C. S. (2005b). Two distinct emotional experiences in romantic relationships: Effects of perceptions regarding approach of intimacy and avoidance of conflict. *Personality and Social Psychology Bulletin*, *31*(8), 1123–1133. <https://doi.org/10.1177/0146167205274447>

- Lavner, J. A., Bradbury, T. N., & Karney, B. R. (2012). Incremental change or initial differences? Testing two models of marital deterioration. *Journal of Family Psychology, 26*(4), 606–616. <https://doi.org/10.1037/a0029052>
- Linville, D., Chronister, K., Dishion, T., Todahl, J., Miller, J., Shaw, D., ... Wilson, M. (2010). A longitudinal analysis of parenting practices, couple satisfaction, and child behavior problems. *Journal of Marital and Family Therapy, 36*(2), 244–255. <https://doi.org/10.1111/j.1752-0606.2009.00168.x>
- McAdams, D. P., Jackson, R. J., & Kirshnit, C. (1984). Looking, laughing, and smiling in dyads as a function of intimacy motivation and reciprocity. *Journal of Personality, 52*(3), 261–273. <https://doi.org/10.1111/j.1467-6494.1984.tb00881.x>
- McClelland, D. C. (1987). *Human Motivation*. New York: Cambridge University Press.
- McClelland, D. C., Koestner, R., & Weinberger, J. (1989). How do self-attributed and implicit motives differ? *Psychological Review, 96*(4), 690–702. <https://doi.org/10.1037/0033-295X.96.4.690>
- McLean, D. C., Nakamura, J., & Csikszentmihalyi, M. (2017). Explaining system missing: Missing data and experience sampling method. *Social Psychological and Personality Science, 8*(4), 434–441. <https://doi.org/10.1177/1948550617708015>
- Merrilees, C. E., Goeke-Morey, M., & Mark Cummings, E. (2008). Do event-contingent diaries about marital conflict change marital interactions? *Behaviour Research and Therapy, 46*(2), 253–262. <https://doi.org/10.1016/j.brat.2007.11.007>
- Miller, G. (2012). The smartphone psychology manifesto. *Perspectives on Psychological Science, 7*(3), 221–237. <https://doi.org/10.1177/1745691612441215>
- Mueller, S., Ram, N., Conroy, D. E., Pincus, A. L., Gerstorf, D., & Wagner, J. (2019). Happy like a fish in water? The role of personality–situation fit for momentary happiness in social interactions across the adult lifespan. *European Journal of Personality, 33*(3), 298–316. <https://doi.org/10.1002/per.2198>
- Neubauer, A. B., Scott, S. B., Sliwinski, M. J., & Smyth, J. M. (2019). How was your day? Convergence of aggregated momentary and retrospective end-of-day

- affect ratings across the adult life span. *Journal of Personality and Social Psychology*. <https://doi.org/10.1037/pspp0000248>
- Oishi, S., & Sullivan, H. W. (2006). The predictive value of daily vs. retrospective well-being judgments in relationship stability. *Journal of Experimental Social Psychology*, *42*(4), 460–470. <https://doi.org/10.1016/j.jesp.2005.07.001>
- Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science*, *349*(6251). <https://doi.org/10.1126/science.aac4716>
- Paunonen, S. V., & O’Neill, T. A. (2010). Self-reports, peer ratings and construct validity. *European Journal of Personality*, *24*(3), 189–206. <https://doi.org/10.1002/per.751>
- Proulx, C. M., Helms, H. M., & Buehler, C. (2007). Marital quality and personal well-being: A meta-analysis. *Journal of Marriage and Family*, *69*(3), 576–593. <https://doi.org/10.1111/j.1741-3737.2007.00393.x>
- Pusch, S., Schönbrodt, F. D., Zygar-Hoffmann, C., & Hagemeyer, B. (2020). Truth and wishful thinking: How inter-individual differences in communal motives manifest in momentary partner perceptions. *European Journal of Personality*, *34*(1), 115–134. <https://doi.org/10.1002/per.2227>
- Rammstedt, B., & Farmer, R. F. (2013). The impact of acquiescence on the evaluation of personality structure. *Psychological Assessment*, *25*(4), 1137–1145. <https://doi.org/10.1037/a0033323>
- R Core Team. (2018). *R: A language and environment for statistical computing*. Retrieved from <https://www.R-project.org/>
- Reis, H., & Shaver, P. (1988). Intimacy as an interpersonal process. In S. Duck, D. F. Hay, S. E. Hobfoll, W. Ickes, & B. M. Montgomery (Eds.), *Handbook of personal relationships* (pp. 367–389). Retrieved from <http://psycnet.apa.org/psycinfo/1988-97881-020>
- Reynolds, B. M., Robles, T. F., & Repetti, R. L. (2016). Measurement reactivity and fatigue effects in daily diary research with families. *Developmental Psychology*, *52*(3), 442–456. <https://doi.org/10.1037/dev0000081>
- Rheinberg, F. (2002). *Motivation*. Stuttgart: Kohlhammer.
- Righetti, F., & Impett, E. (2017). Sacrifice in close relationships: Motives, emo-

- tions, and relationship outcomes. *Social and Personality Psychology Compass*, *11*(10), 1–11. <https://doi.org/10.1111/spc3.12342>
- Robinson, M. D., & Clore, G. L. (2002). Belief and feeling: Evidence for an accessibility model of emotional self-report. *Psychological Bulletin*, *128*(6), 934–960. <https://doi.org/10.1037/0033-2909.128.6.934>
- Robles, T. F., Slatcher, R. B., Trombello, J. M., & McGinn, M. M. (2014). Marital quality and health: A meta-analytic review. *Psychological Bulletin*, *140*(1), 140–187. <https://doi.org/10.1037/a0031859>
- Rogge, R. D., Fincham, F. D., Crasta, D., & Maniaci, M. R. (2017). Positive and negative evaluation of relationships: Development and validation of the Positive-Negative Relationship Quality (PN-RQ) Scale. *Psychological Assessment*, *29*(8), 1028–1043. <https://doi.org/10.1037/pas0000392>
- Rosenkranz, T., Takano, K., Watkins, E. R., & Ehring, T. (2019). Assessing repetitive negative thinking in daily life: Development of an ecological momentary assessment paradigm. *PsyArXiv*. <https://doi.org/10.31234/osf.io/nj73c>
- Schimmack, U. (2003). Affect measurement in experience sampling research. *Journal of Happiness Studies*, *4*(1), 79–106. <https://doi.org/10.1023/A:1023661322862>
- Schmiedek, F., & Neubauer, A. B. (2019). Experiments in the wild: Introducing the within-person encouragement design. *Multivariate Behavioral Research*, *0*(0), 1–21. <https://doi.org/10.1080/00273171.2019.1627660>
- Schönbrodt, F. D., Zygar-Hoffmann, C., Nestler, S., Pusch, S., & Hagemeyer, B. (2019). *Measuring motivational relationship processes in experience sampling: A reliability model for moments, days, and persons nested in couples*. <https://doi.org/10.31234/osf.io/6mq7t>
- Schultheiss, O. C. (2008). Implicit motives. In O. P. John, R. W. Robins, & L. A. Pervin (Eds.), *Handbook of personality: Theory and research* (3rd Edition, pp. 603–633). Retrieved from <http://media.matthewsbooks.com.s3.amazonaws.com/documents/tocwork/019/9780195335156.pdf>
- Schultheiss, O. C., & Brunstein, J. (2010). *Implicit motives*. <https://doi.org/10.1093/acprof:oso/9780195335156.001.0001>
- Schultheiss, O. C., Rösch, A. G., Rawolle, M., Kordik, A., & Graham, S. (2010).

- Implicit motives: Current topics and future directions. In T. C. Urdan & S. A. Karabenick (Eds.), *Advances in motivation and achievement* (Vol. 16, pp. 199–233). [https://doi.org/10.1108/S0749-7423\(2010\)000016A009](https://doi.org/10.1108/S0749-7423(2010)000016A009)
- Schultheiss, O. C., Strasser, A., Rösch, A., Kordik, A., & Graham, S. (2012). Motivation. In V. Ramachandran (Ed.), *Encyclopedia of human behavior* (2nd Edition, pp. 650–656). <https://doi.org/10.1016/B978-0-12-375000-6.00238-X>
- Schüler, J., Job, V., Fröhlich, S. M., & Brandstätter, V. (2008). A high implicit affiliation motive does not always make you happy: A corresponding explicit motive and corresponding behavior are further needed. *Motivation and Emotion*, *32*(3), 231–242. <https://doi.org/10.1007/s11031-008-9096-y>
- Silvia, P. J., Kwapil, T. R., Eddington, K. M., & Brown, L. H. (2013). Missed beeps and missing data: Dispositional and situational predictors of nonresponse in experience sampling research. *Social Science Computer Review*, *31*(4), 471–481. <https://doi.org/10.1177/0894439313479902>
- Silvia, P. J., Kwapil, T. R., Walsh, M. A., & Myin-Germeys, I. (2014). Planned missing-data designs in experience-sampling research: Monte Carlo simulations of efficient designs for assessing within-person constructs. *Behavior Research Methods*, *46*(1), 41–54. <https://doi.org/10.3758/s13428-013-0353-y>
- Simons, D. J., Shoda, Y., & Lindsay, D. S. (2017). Constraints on Generality (COG): A proposed addition to all empirical papers. *Perspectives on Psychological Science*, *12*(6), 1123–1128. <https://doi.org/10.1177/1745691617708630>
- Slatcher, R. B. (2010). Marital functioning and physical health: Implications for social and personality psychology. *Social and Personality Psychology Compass*, *4*(7), 455–469. <https://doi.org/10.1111/j.1751-9004.2010.00273.x>
- Stanley, S. M., Bradbury, T. N., & Markman, H. J. (2000). Structural flaws in the bridge from basic research on marriage to interventions for couples. *Journal of Marriage and Family*, *62*(1), 256–264. <https://doi.org/10.1111/j.1741-3737.2000.00256.x>
- Stanley, S. M., Carlson, R. G., Rhoades, G. K., Markman, H. J., Ritchie, L. L., & Hawkins, A. J. (2019). Best practices in relationship education focused on intimate relationships. *Family Relations*. <https://doi.org/10.1111/fare.12419>
- Stanton, M. (2009). The systemic epistemology of the specialty of family psy-

- chology. In J. H. Brey & M. Stanton (Eds.), *The wiley-blackwell handbook of family psychology* (pp. 5–20). <https://doi.org/10.1002/9781444310238.ch1>
- Stanton, M., & Welsh, R. (2012). Systemic thinking in couple and family psychology research and practice. *Couple and Family Psychology: Research and Practice, 1*(1), 14–30. <https://doi.org/10.1037/a0027461>
- Steyer, R., Ferring, D., & Schmitt, M. (1992). States and traits in psychological assessment. *European Journal of Psychological Assessment, 8*(2), 79–98.
- Sun, J., Rhemtulla, M., & Vazire, S. (2019). *Eavesdropping on missing data: What are people doing when they miss experience sampling reports?* <https://doi.org/10.31234/osf.io/5tcwd>
- Twenge, J. M., Campbell, W. K., & Freeman, E. C. (2012). Generational differences in young adults' life goals, concern for others, and civic orientation, 1966-2009. *Journal of Personality and Social Psychology, 102*(5), 1045–1062. <https://doi.org/10.1037/a0027408>
- Vaez, E., Indran, R., Abdollahi, A., Juhari, R., & Mansor, M. (2015). How marital relations affect child behavior: Review of recent research. *Vulnerable Children and Youth Studies, 10*(4), 321–336. <https://doi.org/10.1080/17450128.2015.1112454>
- Wang, L. P., & Maxwell, S. E. (2015). On disaggregating between-person and within-person effects with longitudinal data using multilevel models. *Psychological Methods, 20*(1), 63–83. <https://doi.org/10.1037/met0000030>
- Wilt, J., & Revelle, W. (2019). The Big Five, everyday contexts and activities, and affective experience. *Personality and Individual Differences, 136*(March 2017), 140–147. <https://doi.org/10.1016/j.paid.2017.12.032>
- Winter, D. G., John, O. P., Stewart, A. J., Klohnen, E. C., & Duncan, L. E. (1998). Traits and motives: Toward an integration of two traditions in personality research. *Psychological Review, 105*(2), 230–250. <https://doi.org/10.1037/0033-295X.105.2.230>
- Yarkoni, T. (2019). The generalizability crisis. *PsyArXiv*. <https://doi.org/10.31234/osf.io/jqw35>
- Yu, F. (2018). Exception question in couple and family therapy. In J. Lebow, A. Chambers, & D. Breunlin (Eds.), *Encyclopedia of couple and family therapy*

(pp. 1–3). https://doi.org/10.1007/978-3-319-15877-8_1071-1

- Zhang, Z., Zyphur, M. J., & Preacher, K. J. (2009). Testing multilevel mediation using hierarchical linear models: Problems and solutions. *Organizational Research Methods, 12*(4), 695–719. <https://doi.org/10.1177/1094428108327450>
- Zygar, C., Hagemeyer, B., Pusch, S., & Schönbrodt, F. D. (2018a). *From motive dispositions to states to outcomes: Research data of an intensive experience sampling study on communal motivational dynamics in couples [translated title] (version 2.0.0) [data and documentation]*. Trier, Germany: Center for Research Data in Psychology PsychData of the Leibniz Institute for Psychology Information ZPID. https://doi.org/10.5160/psychdata.zrce16dy99_v2000.
- Zygar, C., Hagemeyer, B., Pusch, S., & Schönbrodt, F. D. (2018b). *Motivationale Konflikte in Partnerschaften Synchronisation oder Dominanz?* <https://doi.org/10.17605/osf.io/qjr97>
- Zygar, C., Hagemeyer, B., Pusch, S., & Schönbrodt, F. D. (2018c). *Paarforschung mit impliziten Motiven: Mit Experience-Sampling die motivationalen Prozesse verstehen*. <https://doi.org/10.17605/osf.io/nqzgm>
- Zygar, C., Hagemeyer, B., Pusch, S., & Schönbrodt, F. D. (2018d). *Understanding the relationship between motives and relationship satisfaction: A differentiated look at intra-individual processes of agentic motivation*. <https://doi.org/10.17605/osf.io/2mxpt>
- Zygar-Hoffmann, C., Hagemeyer, B., Pusch, S., & Schönbrodt, F. D. (2020). *A large longitudinal study on motivation, behavior and satisfaction in couples: Research data from a four-week experience sampling study with a pre-, post-, and one-year follow-up-assessment [translated title] (version 1.0.0) [data and documentation]*. Trier, Germany: Center for Research Data in Psychology PsychData of the Leibniz Institute for Psychology Information ZPID. <https://doi.org/10.5160/psychdata.zrce18mo99>.