

Bewegungssynchronie zwischen Patienten mit
Sozialer Angststörung und ihren
Psychotherapeuten

Testung, Weiterentwicklung und Anwendung linearer zeitreihenanalytischer
Verfahren

Dissertation
Zur Erlangung des akademischen Grades
doctor philosophiae (Dr. phil.)

vorgelegt dem Rat der Fakultät für Sozial- und Verhaltenswissenschaften
der Friedrich-Schiller-Universität Jena
von Silvia Désirée Schönherr geb. Thielemann (M.Sc. Psychologie)
geboren am 27.04.1990 in Altdöbern

Gutachter

- 1. Prof. Dr. Bernhard Strauß (Jena)**
- 2. Prof. Dr. Peter Noack (Jena)**
- 3. Prof. Dr. Wolfgang Tschacher (Bern)**

Tag der mündlichen Prüfung: 08.05.2019

Inhaltsverzeichnis

| | |
|---|----|
| Abkürzungsverzeichnis | 7 |
| Zusammenfassung | 9 |
| Abstract | 11 |
| 1 Einleitung | 13 |
| 2 Ziel der vorliegenden Arbeit | 15 |
| 3 Theoretische Einbettung in den bisherigen Forschungsstand | 16 |
| 3.1 Soziale Angststörung..... | 16 |
| 3.1.1 Kognitives Erklärungsmodell der Sozialen Angststörung..... | 18 |
| 3.1.2 Nonverbale Charakteristika..... | 19 |
| 3.1.3 Therapie Sozialer Angststörungen..... | 20 |
| Symptomverbesserung..... | 20 |
| Therapieabbruch..... | 20 |
| 3.2 Nonverbale Synchronie..... | 22 |
| 3.2.1 Nonverbale Synchronie im Zeitverlauf..... | 24 |
| 3.2.2 Einflussfaktoren auf Synchronie..... | 25 |
| Im klinischen/psychotherapeutischen Setting..... | 29 |
| Interpersonelle Synchronie bei sozialen Ängsten..... | 29 |
| 3.2.3 Konsequenzen von Synchronie..... | 31 |
| Nonverbale Synchronie und Therapieoutcome..... | 33 |
| Nonverbale Synchronie und Bindung..... | 37 |
| 3.2.4 Biologische Grundlagen interpersoneller Synchronie und Mechanismus..... | 38 |
| 3.2.5 In-Sync Modell der Psychotherapie..... | 39 |
| 3.2.6 Erfassung Synchronie..... | 41 |
| Messung von Merkmalen..... | 41 |
| Vorverarbeitung der Messequenz & statistische Vorüberlegungen..... | 44 |
| Bestimmung der nonverbalen Synchronie..... | 45 |

| | |
|--|-----------|
| Validität der nonverbalen Synchronie..... | 47 |
| Maße nonverbaler Synchronie..... | 53 |
| 4 Datengrundlage & Inhalte der einzelnen Publikationen | 54 |
| 4.1 Erfassung von Körperbewegungen mittels Motion Energy Analysis | 58 |
| 4.2 Untersuchungen innerhalb der Originalarbeiten | 64 |
| 4.2.1 Untersuchung der konvergenten und kriteriumsorientierten Validität zeitreihenbasierter linearer Methoden (Publikation 1)..... | 65 |
| 4.2.2 Identifikation der besten Parameterkonfiguration für die zeitreihenbasierten Methoden (Publikation 2)..... | 66 |
| Zusatzuntersuchungen zur Validität der verwendeten R^2 Schwelle von 0.25 | 68 |
| 4.2.3 Zusammenhang zwischen Therapieabbruch und nonverbaler Synchronie (Publikation 3)..... | 71 |
| Zusatzanalysen – andere Moderatoren | 72 |
| Zusatzanalysen – andere Zeitpunkte | 74 |
| Zusatzanalysen zum Vergleich des Algorithmus mit Anwendung zweier verschiedener maximaler time lags | 75 |
| Zusatzanalyse: Robustheit der Effekte unter Hinzunahme der Variable „therapeutische Allianz“, gemessen durch den Helping Alliance Questionnaire | 76 |
| 4.2.4 Zusammenhang zwischen Bindungsangst und –vermeidung und nonverbaler Synchronie (Publikation 4)..... | 77 |
| 4.2.5 Vorhersage von Veränderungen der interpersonellen Probleme und Depression durch nonverbale Synchronie (Publikation 5)..... | 79 |
| Zusatzanalyse: Veränderung der Synchronie-Häufigkeit und der Leading-Anteile innerhalb der Therapie | 80 |
| 4.2.6 Zusammenhang zwischen Bewegungsparametern und Selbst-Ratings zu interpersonellen Problemen (Zusatzanalyse)..... | 82 |
| 4.2.7 Arbeitsanteil D. Schönherr innerhalb der Publikationen | 86 |
| 5 Publikationsübergreifende Diskussion | 87 |
| 5.1 Bedeutung der methodische Arbeiten | 87 |
| 5.2 Wichtigkeit der inhaltlichen Arbeiten & Erweiterung des In-Sync Modells | 90 |

| | |
|--|------------|
| 5.3 Stärken und Einschränkungen | 98 |
| 5.4 Implikationen für die Psychotherapie | 101 |
| 5.5 Implikationen für die Forschung | 103 |
| Referenzen..... | 105 |
| Abbildungsverzeichnis | 116 |
| Tabellenverzeichnis | 118 |
| Anhang | 119 |
| Publikation 1 | 120 |
| Publikation 2 | 143 |
| Publikation 3 | 167 |
| Publikation 4 | 204 |
| Publikation 5 | 240 |
| Deskriptive Daten Bewegungsanalysen N =86 | 279 |
| Publikationen und Kongressbeiträge | 280 |
| Publikationen – angenommen/veröffentlicht | 280 |
| Publikationen – in Vorbereitung/ eingereicht | 281 |
| Kongressbeiträge – vorgestellt | 281 |
| Kongressbeiträge – beteiligt..... | 283 |
| Ehrenwörtliche Erklärung | 285 |
| Danksagung..... | 286 |

Abkürzungsverzeichnis

| | |
|-----------|--|
| ANOVA | <i>Analysis of Variance</i> |
| BDI | <i>Beck-Depression-Inventory</i> |
| CLC | <i>Cross Lagged-Correlation</i> |
| CLR | <i>Cross Lagged-Regression</i> |
| CRQA | <i>Cross Recurrence Quantification Analysis</i> |
| CSC | Klinisch bedeutsame Veränderung (<i>clinically significant change</i>) |
| DSM | <i>Diagnostic and Statistical Manual of Mental Disorders</i> |
| ES | Effektstärke |
| HAQ | <i>Helping Alliance Questionnaire</i> |
| HLM | <i>Hierarchical Linear Modeling</i> |
| ICD | <i>International Classification of Diseases</i> |
| IIP | Inventar Interpersoneller Probleme |
| IRESTRA | Projekt „Irritationsfreies und emotionssensitives Trainingssystem“ |
| KVT/CBT | Kognitive Verhaltenstherapie |
| MEA | <i>Motion Energy Analysis</i> |
| MSI | <i>Movement synchrony interval</i> |
| PDT | Psychodynamisch orientierte Therapie |
| PKS | Persönlichkeitsstörung |
| ROI | <i>Region-of-Interest</i> |
| RMSEA | <i>Root Mean Square Error of Approximation</i> |
| RR | Rekurrenz-Raten |
| SAD | Soziale Angststörung (<i>social anxiety disorder</i>) |
| SOPHO-NET | Projekt „Forschungsverbund soziale Phobie“ |
| SX | X. Sitzung der Therapie |
| TIMPATHIN | Projekt „Timing nonverbaler Patient-Therapeut-Interaktionen und Therapieerfolg bei sozialen Phobien“ |
| TSAM | <i>Time series analysis methods</i> |
| WCC | <i>Windowed Cross-Correlation</i> |
| WCLC | <i>Windowed Cross-Lagged Correlation</i> |

| | |
|--------------------|--|
| WCLC _F | <i>Windowed Cross-Lagged Correlation</i> mit <i>Peak-Picking</i> Algorithmus (vgl. Altmann, 2013) |
| WCLC _{S1} | <i>Windowed Cross-Lagged Correlation</i> (vgl. Paulick, 2017) |
| WCLC _{S2} | <i>Windowed Cross-Lagged Correlation</i> (vgl. Ramseyer und Tschacher, 2011) |
| WCLR | <i>Windowed Cross-Lagged Regression</i> |
| ZBKT | Zentrales Beziehungs-Konflikt-Thema |

Zusammenfassung

Bewegungssynchronie zwischen Patient und Psychotherapeut beschreibt das Aufeinander-Bezug-nehmen von deren Bewegungen und ist mit Outcome-Maßen der Psychotherapie assoziiert. Diese Zusammenhänge wurden meist in störungsheterogenen Stichproben bei verschiedenen Behandlungskonzepten gefunden. Der Einfluss der Störung und des Behandlungskonzeptes auf Synchronie kann jedoch nicht ausgeschlossen werden. Die Untersuchungen wurden mittels zeitreihenanalytischen Methoden durchgeführt. Dabei wurde die Validität nur partiell untersucht. Aktuell fehlt eine umfassende Prüfung anhand verschiedener Validitätskriterien, um die Ergebnisse abzusichern.

Innerhalb dieser Arbeit wurde die Bewegungssynchronie zu verschiedenen Zeitpunkten dreier verschiedener Kurzzeit-/Langzeittherapien (manualisierte KVT, manualisierte psychodynamisch-orientierte Therapie [PDT] und naturalistische KVT) von Patienten mit Sozialer Angststörung in Zusammenhang zu verschiedenen Outcome-Maßen (Therapieabbruch, Symptomatik, Bindung) untersucht. Dabei wurden die Bewegungszeitreihen automatisiert aus Videosequenzen mittels Motion Energy Analysis gewonnen. Zur Ermittlung der Synchronie für inhaltliche Fragestellungen wurde die $WCLC_F$ verwendet, die sich bei den vorgeschalteten Untersuchungen der Validität (Faktorstruktur, konvergente, konkurrente & prädiktive Validität, Identifikation von optimalen Parametersettings) als beste Methode herausgestellt hat. Bei den Validitätsuntersuchungen stellte sich heraus, dass verschiedene Synchronie-Maße kein eindimensionales Konstrukt sondern verschiedene Facetten von Synchronie messen: die *Häufigkeit der Synchronie* (Häufigkeitsmaße der $WCLC_F$, WCLR, CRQA), die *Stärke der Synchronie in Synchronisationsintervallen* (Stärkemaße der $WCLC_F$, WCLR) und die *Stärke der Synchronie in der Gesamtinteraktion* ($WCLC_{S1}$, $WCLC_{S2}$, CLC). Die Facetten mit übergeordnetem Synchronie-Faktor waren jedoch nur für bestimmte Parametereinstellungen unter Ausschluss einiger Methoden nachweisbar. Die konkurrente Validität mit einer non-linearen Methode kann als gut beschrieben werden. Die prädiktive Validität, gemessen durch den Zusammenhang zwischen interpersonellen Problemen und Synchronie wurde nur für die WCC, $WCLC_F$ und WCLR gefunden. Die Parametereinstellungen waren von zentraler Bedeutung. Divergenzen zwischen bisherigen Forschungsergebnissen können unter anderem durch verschiedene Methoden oder unterschiedliche Parameter der Synchronie-Bestimmung erklärt werden.

Die Lokalisationsfähigkeit der $WCLC_F$ und WCLR wurde anschließend anhand simulierter Zeitreihen und menschlich eingeschätzten Synchronisationsintervallen geprüft. Es

wurden kleine Bandbreiten, keine Glättung aber eine Logarithmus-Transformation der Zeitreihe und ein R^2 -Schwellenwert von 0.25 für Synchronisationsintervalle als optimale Parametereinstellungen identifiziert. Wurden die Synchronisationsintervalle nicht von anderen Bewegungen isoliert, resultierte eine geringe Identifikationsrate. Als Grund für die fehlende Übereinstimmung wird angenommen, dass von Ratern auch andere Qualitäten der Interaktion (z.B. Blickkontakt, Semantik) zur Detektion von Synchronie benutzt werden.

Innerhalb der unterschiedlichen Therapieansätze wurde ein höheres Ausmaß an Bewegungssynchronie in der KVT gefunden. Abbrüche der Therapie konnten bereits durch die Bewegungssynchronie innerhalb der probatorischen Sitzungen vorhergesagt werden. Eine Erhöhung der Synchronie um 1% resultierte in einer um 5% geringeren Abbruchrate. Dies bietet das Potential, Risikopatienten für einen Therapieabbruch bereits zu Beginn der Therapie anhand einer geringen Synchronie zu identifizieren. Der Effekt wurde vor allem für Patienten-geleitete Synchronie und innerhalb der PDT nachgewiesen.

Das Ausmaß der interpersonellen Probleme der Patienten zum Ende der Therapie konnte ebenfalls durch Synchronie vorhergesagt werden (Klinisch signifikante Veränderung sowie Vorhersage interpersoneller Probleme). Dabei war die Symptomatik am geringsten, wenn der Therapeut die Bewegungssynchronie häufiger initiierte als der Patient. Dies deutet darauf hin, dass gerade für die Symptomreduktion eine Anleitung sowie Strukturierung der Interaktion durch den Therapeuten förderlich ist. Dies konnte ebenfalls für die Bindungsangst des Patienten nachgewiesen werden. In Dyaden mit viel Therapeuten-geleiteter Bewegungssynchronie nach der Probatorik wurde eine geringere Bindungsangst des Patienten am Ende der Therapie gefunden. Diese Effekte wurden vermehrt in der manualisierten KVT festgestellt. Dies ist plausibel, da diese Behandlungskonzeption strukturierter und geleiteter vorgeht als die anderen beiden Therapierichtungen. Die Synchronie zu Beginn der Therapie war nicht von der Bindungsangst und –vermeidung des Patienten abhängig. Jedoch kann eine hohe Dauer zwischen synchronen Bewegungen einen Hinweis auf eine hohe Bindungsvermeidung des Patienten geben.

Bewegungssynchronie scheint eine eher proximale Variable zu interpersonellen Konstrukten zu sein. Dies wird dadurch hervorgehoben, dass kein Zusammenhang zwischen Synchronie und Sozialer Angst/Vermeidung sowie nur sehr eingeschränkte Zusammenhänge zu Depression festgestellt werden konnten. Die Ergebnisse deuten auf die Notwendigkeit einer Erweiterung des In-Sync Modells hin, die in der vorliegenden Dissertation vorgeschlagen wird.

Abstract

Movement synchrony between patient and psychotherapist describes the interrelationship of their movements and is associated with outcome measures of psychotherapy. These associations were mostly found in heterogeneous samples in different treatment approaches. However, the influence of the disorder and the treatment approach on synchrony cannot be ruled out. In the social and behavioral sciences, synchrony is determined by time series analysis methods. However, the validity of these methods is only partially investigated. There is no comprehensive examination on the basis of various validity criteria.

Within this work movement synchrony at different sessions of three different short-term/long-term therapies (manual-guided CBT, manual-guided psychodynamic-oriented therapy [PDT] and naturalistic CBT) of patients with social anxiety disorder was investigated in association with different outcome measures (premature termination, symptomatology, attachment). Thereby, movement time series were obtained automatically from video sequences using Motion Energy Analysis. The $WCLC_F$, which was proved to be the best method in the preceding investigations of validity (factor structure, convergent, concurrent & predictive validity, identification of optimal parameter settings) was used to determine synchrony.

The investigations showed that different synchrony measures do not measure a one-dimensional construct but different facets of synchrony: the *frequency of synchrony* (frequency measures of $WCLC_F$, $WCLR$, $CRQA$), the *strength of synchrony in synchronization intervals* (strength measures of $WCLC_F$, $WCLR$) and the *strength of synchrony in the overall interaction* ($WCLC_{S1}$, $WCLC_{S2}$, CLC). However, the facets with second-order synchrony factor were only detectable for certain parameter settings and the exclusion of some methods. The concurrent validity with a non-linear method can be described as good. The predictive validity, measured by the relationship between interpersonal problems and synchrony, was shown found for the WCC , $WCLC_F$ and $WCLR$. The parameter settings were of critical importance. Divergences between previous research results can be explained, among other things, by different methods or the usage of different parameter settings which are not described detailed.

The localization capability of the $WCLC_F$ and $WCLR$ was then tested using simulated time series and human-rated synchronization intervals. Small bandwidths, no smoothing but a logarithmic transformation of the time series and an R^2 cut-off of 0.25 for synchronization intervals were identified as optimal parameter settings. If the synchronization intervals were not isolated from other movements, a low identification rate resulted. The reason for the lack

of agreement might be that human raters also use other qualities of the interaction (e.g., eye contact, semantics) to detect synchrony.

In the comparison of the different treatment approaches, a higher degree of synchrony of movement was found within the CBT approaches. Premature terminations could already be predicted by the synchrony of movement within the probatoric sessions. An increase of the synchrony by 1% resulted in a 5% lower termination rate. This offers the potential to identify risk patients for a premature termination already at the beginning of the therapy on the basis of a low synchrony. The effect has been demonstrated especially for patient-leading synchrony and within PDT.

The extent of the patients' interpersonal problems at the end of therapy could also be predicted by synchrony (clinically significant change and prediction of interpersonal problems). Symptoms were lowest if the therapist initiated synchrony more often than the patient. This indicates that especially for symptom reduction a guidance and structuring of the interaction by the therapist is beneficial. This could also be demonstrated for the attachment-related anxiety of the patient. A lower patient's attachment-related anxiety was found at the end of the therapy in dyads with a lot of therapist-leading movement synchrony after the probatoric sessions. The effects were rather observed in manual-guided CBT. This is plausible, as this treatment concept is more structured and guided than the other two therapeutic approaches. The synchrony at the beginning of the therapy was not dependent on the patient's attachment-related anxiety and avoidance. However, a high duration between synchronous movements can give an indication for a high attachment-related avoidance of the patient.

Movement synchrony seems to be a more proximal variable to interpersonal constructs. This is further emphasized by the fact that no connection between synchrony and social anxiety or avoidance and only very limited relations to depression could be detected. There are also differences in who initiates the synchrony and how it is related to different outcome measures. The results indicate the need for an extension of In-Sync model of psychotherapy, which is proposed in this dissertation.

1 Einleitung

Innerhalb von Psychotherapien ist Kommunikation die wichtigste Basis für das therapeutische Arbeiten. Sie besteht aus verbalen und nonverbalen Elementen, wobei gerade die nonverbalen Aspekte mindestens 2/3 der Kommunikation ausmachen (Burgoon, Guerrero, & Floyd, 2009). Beide Kommunikationssysteme stehen in ständiger Interaktion miteinander. Nonverbale Kommunikation kann dabei die verbale Nachricht wiederholen, ersetzen, komplettieren, unterstreichen oder moderieren. Ebenso ist es möglich, dass die nonverbale Nachricht eine entgegengesetzte Botschaft zur verbalen vermittelt (Ekman, 1965). Tritt ein Widerspruch auf, wird der nonverbalen Information der Vorrang gegeben (Westland, 2015). Gerade deshalb ist es unabdingbar, der nonverbalen Kommunikation im Kontext der Psychotherapie in der Forschung und der Praxis mehr Bedeutung zuzusprechen.

Psychische Störungen manifestieren sich auf beiden Ebenen. Die Soziale Angststörung ist eine der am häufigsten auftretenden Störungen mit starken interpersonellen Einschränkungen (Alden & Taylor, 2004; Wenzel, Graff-Dolezal, Macho, & Brendle, 2005). Sie ist beispielsweise durch Veränderungen der Aufmerksamkeit, Wahrnehmung, Kognition sowie durch spezifische nonverbale Charakteristika gekennzeichnet. Patienten¹ zeigen einen erhöhten Selbstfokus und eine verschobene Aufmerksamkeit hinsichtlich negativer bzw. ambivalenter Reize (Gilboa-Schechtman & Shachar-Lavie, 2013). Auf nonverbaler Ebene definiert das DSM-V eine rigide Körperhaltung, gedämpfte Stimme sowie inadäquaten Blickkontakt als störungsspezifisch. Die Einschränkungen im interpersonellen Kontakt können auch per interpersoneller Synchronie zwischen diesen Patienten und ihren Interaktionspartnern festgestellt werden. Die interpersonelle Synchronie stellt dabei den Sammelbegriff für verschiedene Synchronisationsphänomene dar, die sich dadurch kennzeichnen, dass das Verhalten, die Emotionen oder andere Prozesse zweier Interaktionspartner für einen lokal eingrenzenden Zeitraum (zeitversetzt oder simultan) aufeinander Bezug nehmen. Beispielsweise nickt ein Interaktionspartner dem anderen zu, der seine Aussage gerade mit Gesten untermauert oder die Stimmfrequenzen zweier Interaktionspartner gleichen sich an. Weitere Synchronisationsphänomene sind z.B. das Einnehmen gleicher Körperhaltungen, Spiegeln der Mimik oder die Kopplung neuronaler Verarbeitungsprozesse.

¹ Für die erleichterte Lesbarkeit der Arbeit wird für die Bezeichnung der Patienten/Patientinnen und Therapeuten/Therapeutinnen auf das generische Maskulinum zurückgegriffen. Damit sind innerhalb dieser Arbeit stets Patienten sowie Patientinnen und Therapeuten sowie Therapeutinnen gemeint.

Bei Patienten mit Sozialer Angststörung wurde ein geringeres Level an Synchronie festgestellt. Ein Grund dafür kann die selbstzentrierte Aufmerksamkeit der Patienten sein (Clark & Wells, 1995), die in Folge dessen scheinbar weniger in Kontakt mit dem Interaktionspartner treten können und eine geringere Synchronie zeigen. Da interpersonelle Synchronie auf viele Aspekte des Gesprächs (z.B. Geschmeidigkeit des Gesprächs, Gefühl mit dem Gesprächspartner eine Einheit zu bilden) positiv wirkt, könnte die veränderte Synchronie in Patienten mit Sozialer Angststörung ein wichtiger Ansatzpunkt für die Psychotherapie-Prozessforschung sein.

Innerhalb der Prozessforschung wurden mehrere Aspekte als Wirkfaktoren von Psychotherapie identifiziert. Zentrale Wirkfaktoren sind beispielsweise die therapeutische Allianz, das Problemlösen oder die Ressourcenaktivierung (Pfammatter, Junghan, & Tschacher, 2012). Diese gründen sich auf spezifischen Techniken der Psychotherapie wie beispielsweise dem Verbalisieren oder der positiven Verstärkung. Unklar ist momentan die Stellung nonverbaler Verhaltensweisen und speziell der interpersonellen Synchronie innerhalb dieses Kontextes. Betrachtet man nonverbale Aspekte, könnte die interpersonelle Synchronie zwischen Patient und Therapeut ebenfalls ein Wirkfaktor von Psychotherapien sein, der allen Therapierichtungen gemein ist. Gerade in Bezug auf den Prozess wie Psychotherapie wirkt, ist es wichtig zu verstehen, welcher Wirkfaktor mit welcher Veränderung der Symptomatik des Patienten assoziiert ist.

Eine spezielle Form der interpersonellen Synchronie, die Bewegungssynchronie, wurde bereits mehrfach im psychotherapeutischen Setting untersucht. Sie meint die dynamische Kopplung von Bewegungen, wie beispielsweise das Vorlehnen des Therapeuten mit leicht zeitverzögertem Nicken des Patienten. Die Ergebnisse der Studien zur Bewegungssynchronie an Stichproben verschiedener Störungsbilder sowie Therapiekonzepte weisen darauf hin, dass Bewegungssynchronie zwischen Patient und Therapeut einen positiven Einfluss auf die therapeutische Allianz, Therapieabbrüche und den Outcome hat (Galbusera et al., 2016; Paulick et al., 2017; Ramseyer & Tschacher, 2011). Jedoch sind nicht alle Befunde konsistent, was durch verschiedene Störungsbilder innerhalb einer Stichprobe, unterschiedliche Therapieansätze sowie nicht komplett deckungsgleiche Methoden zur Synchronie-Erfassung erklärt werden könnte. Messmethoden für interpersonelle Synchronie variieren zwischen menschlichen Ratings und Algorithmen oder Hardware-basierten Erhebungsmethoden. Innerhalb der Sozial- und Verhaltenswissenschaften haben lineare Zeitreihenanalysemethoden an Bedeutung gewonnen. Jedoch sind diese Methoden nur eingeschränkt objektiv und meist nur

Augenschein-valid oder genügen lediglich dem Anspruch der prädiktiven Validität. Oft fehlt es an einer umfassend abgesicherten und Methoden-basierten Definition der Synchronie.

Untersuchungen innerhalb von Settings mit einer geringeren Anzahl an Einflussfaktoren auf die Synchronie sind unabdingbar. Da gerade die Soziale Angststörung mit (nonverbalen) intra- sowie interpersonellen Veränderungen assoziiert ist, stellen Patienten mit dieser Störung eine prädestinierte Stichprobe für die Untersuchungen dieser Arbeit dar. Um eine messfehlerfreie Erhebung von Bewegungssynchronie zu gewährleisten, sollten überdies umfassende Untersuchungen der Validität der benutzten Zeitreihenanalysemethoden erfolgen.

2 Ziel der vorliegenden Arbeit

Das primäre Ziel der vorliegenden Arbeit war die Untersuchung des Zusammenhangs von Bewegungssynchronie und Therapieerfolg (operationalisiert als Therapieabbruch in Publikation 3 und Änderung der interpersonellen Probleme und Depression in Publikation 5) in verschiedenen Therapiesettings. Dabei wurde eine störungshomogene Stichprobe von Patienten mit Sozialer Angststörung untersucht, die innerhalb drei verschiedener ambulanter Kurzzeit- bzw. Langzeittherapien behandelt wurden. Weiterhin wurden an einer Teilstichprobe der Zusammenhang zwischen Bindungsangst und –vermeidung und Bewegungssynchronie sowie die Assoziation zwischen Sozialer Angst und Vermeidung und Bewegungssynchronie erforscht (Publikation 4).

Da es in der Literatur keine Studien zur Güte der verwendeten zeitreihenanalytischen Methoden gibt, war es vor Bearbeitung der inhaltlichen Fragestellungen nötig, die anzuwendenden Methoden kritisch auf ihre methodische Qualität zu testen. Dafür wurde eine Testung der konvergenten Validität verschiedener linearer Zeitreihenanalysemethoden sowie eine kriteriums-basierte Validitätsprüfung an Outcomevariablen (prädiktive Validität), einer generalisierten non-linearen Methode (konkurrente Validität) und eine Validierung anhand von simulierten und menschlich-eingeschätzten Synchronisationsintervallen durchgeführt (Publikation 1 + 2). Um die Parameterkombination zu bestimmen, die zu den konvergentesten Ergebnissen führt, wurden eine Vielzahl an verschiedenen Zeitreihenanalyseparameter zweier als valide identifizierten Methoden anhand ihrer Identifikationsraten hinsichtlich simulierter und menschlich-eingeschätzter Synchronisationsintervalle verglichen (Publikation 2). Es wurde die für die untersuchten Daten beste Methode verwendet, um die Synchronie-Werte für die inhaltlichen Analysen zu bestimmen. Durch diese Methodentestung und -weiterentwicklung wurde sichergestellt, dass die inhaltlichen Fragestellungen unter Einhaltung der psychologischen Gütekriterien bearbeitet wurden.

3 Theoretische Einbettung in den bisherigen Forschungsstand

3.1 Soziale Angststörung

Das Störungsbild der Sozialen Angststörung (DSM-V Terminus *social anxiety disorder* (SAD)) ist durch eine anhaltende Angst vor sozialen Situationen gekennzeichnet. Diese Ängste führen oft zu einem ausgeprägten Vermeidungs- und Sicherheitsverhalten. Sozial-ängstliche Patienten haben die Befürchtung sich peinlich zu verhalten, von anderen negativ bewertet oder in sozialen Situationen gedemütigt zu werden. Dabei ist nicht nur die eigentliche Situation, sondern bereits der Gedanke an diese angstausslösend. Mit einer Lebenszeitprävalenz von 12.1²% und einer 1-Jahres Prävalenz von 7.1% zählt die Soziale Angststörung zu den am häufigsten auftretenden psychischen Störungen (Ruscio et al., 2008). Die diagnostischen Kriterien, nach DSM-V und ICD-10, sind in der folgenden Tabelle dargestellt:

Tabelle 1 *Diagnosekriterien einer Sozialen Phobie nach DSM-V und ICD-10*

| DSM-V: Soziale Phobie (300.23) | ICD-10: Soziale Phobie (F40.1) |
|--|--|
| A) Ausgeprägte Furcht oder Angst vor einer oder mehreren sozialen Situationen, in denen die Person von anderen Personen beurteilt werden könnte. Beispiele hierfür sind soziale Interaktionen (z.B. Gespräche mit anderen, Treffen mit unbekanntem Personen), beobachtet zu werden (z. B. beim Essen oder Trinken) und vor anderen Leistungen zu erbringen (z. B. eine Rede halten). | Angst vor prüfender Betrachtung durch andere in relativ kleinen Gruppen. Beginn häufig in der Jugend. Die Angst kann in klar abgegrenzten oder in fast allen sozialen Situationen auftreten. Sie geht in der Regel mit niedrigem Selbstwertgefühl und Angst vor Kritik einher. Die Angstsymptome können sich zu Panikattacken entwickeln. Durch Vermeidung kann es zur sozialen Isolierung kommen. |
| B) Betroffene befürchten, dass sie sich in einer Weise verhalten könnten oder Symptome der Angst offenbaren, die von anderen negativ bewertet werden (d. h. die beschämend oder peinlich sind, zu Zurückweisung führen oder andere Personen kränken). | Leitlinien: 1. psychische, Verhaltens- oder vegetative Symptome, die Ausdruck von Angst sind, und nicht auf Wahn oder Zwangsgedanken beruhen 2. Angst ist beschränkt auf oder überwiegt |

² Für die Konsistenz mit den englisch-sprachigen Publikationen wird auch in vorliegender Rahmenschrift der Punkt als Dezimaltrennzeichen benutzt.

- C) Die sozialen Situationen rufen fast immer eine Furcht- oder Angstreaktion hervor. in bestimmten sozialen Situationen.
3. Vermeidung der gefürchteten Situation,
- D) Die sozialen Situationen werden vermieden oder unter intensiver Furcht oder Angst ertragen. wenn möglich
Dazugehörige Begriffe: Anthropophobie, soziale Neurose
- E) Die Furcht oder Angst geht über das Ausmaß der tatsächlichen Bedrohung durch die soziale Situation hinaus und ist im soziokulturellen Kontext unverhältnismäßig. Differentialdiagnose: Agoraphobie, Depression
- F) Die Furcht, Angst oder Vermeidung ist andauernd, typischerweise über sechs Monate oder länger.
- G) Die Furcht, Angst oder Vermeidung verursacht in klinisch bedeutsamer Weise Leiden oder Beeinträchtigungen in sozialen, beruflichen oder anderen wichtigen Lebensbereichen.
- H) Die Furcht, Angst oder Vermeidung ist nicht Folge der physiologischen Wirkung einer Substanz (z.B. Substanz mit Missbrauchspotential, medikamentöse Wirkstoffe) oder eines medizinischen Krankheitsfaktors.
- I) Die Furcht, Angst oder Vermeidung kann nicht besser durch die Symptome einer anderen psychischen Störung erklärt werden, wie z.B. Panikstörung, Körperdysmorphie Störung oder Autismus-Spektrums-Störung.
- J) Falls ein medizinischer Krankheitsfaktor (z.B. Morbus Parkinson, Adipositas, eine Entstellung durch Verbrennung oder Verletzung) vorliegt, so steht die Furcht, Angst oder Vermeidung nicht damit in
-

Zusammenhang oder geht deutlich darüber

hinaus.

Tabelle abgeändert aus: Hautzinger und Thies (2009), Seite 10; DSM-V, Seite 138 f.

3.1.1 Kognitives Erklärungsmodell der Sozialen Angststörung

Die Prozesse, die auf kognitiver Ebene zum Erhalt der Störung ablaufen, wurden von Clark und Wells (1995) in ihrem kognitiven Modell der Sozialen Phobie zusammengefasst (siehe Abb. 1).

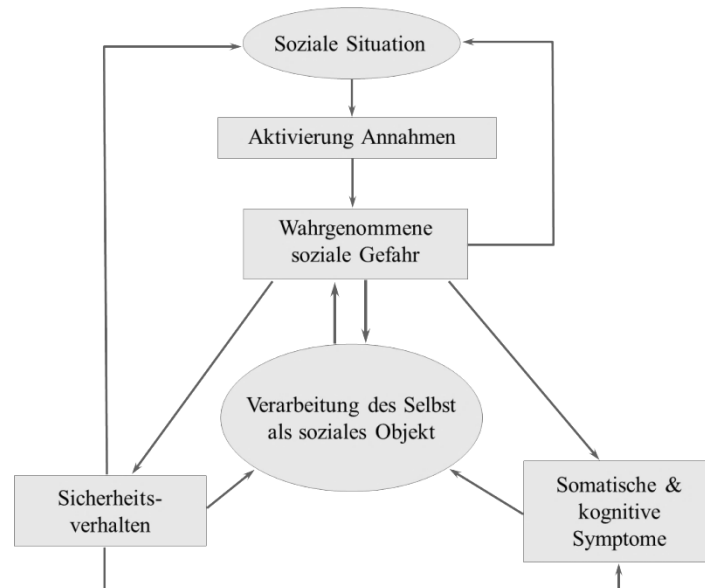


Abbildung 1 Kognitives Modell der Sozialen Angststörung nach Clark und Wells (2001), adaptiert aus Clark (2005)

Patienten mit einer sozialen Phobie entwickeln eine Reihe von Annahmen über soziale Kontexte, die drei Kategorien zuzuordnen sind: 1) übermäßig hohe Standards für das Verhalten in sozialen Situationen, 2) Annahmen über Konsequenzen, die sie nach gewissen Handlungen zu befürchten haben, 3) übermäßig negative Sicht auf das eigene Selbst. Soziale Situationen werden somit als gefährlich eingeschätzt, da der von sich selbst erwartete Standard an übermäßig angepasstem Verhalten nicht erfüllt werden könnte (Clark, 2005). Negative oder mehrdeutige Stimuli werden stets wahrgenommen und negativ interpretiert (Gilboa-Schechtman & Shachar-Lavie, 2013; Hofmann, 2007). Nehmen Patienten mit Sozialer Angststörung solche Stimuli wahr, fokussieren sie ihre Aufmerksamkeit vermehrt auf sich selbst. Damit versuchen sie durch ständiges Monitoring der eigenen Körperempfindungen und Kognitionen zu verhindern, sich unangemessen zu verhalten und somit nicht im Mittelpunkt der Aufmerksamkeit zu stehen, was zu einer vermeintlich negativen Bewertung durch andere führen könnte. Sie wenden eine Reihe verschiedener Sicherheitsverhaltensweisen in der

Situation an (z.B. ständiges Durchgehen der zu sagenden Sätze, dünne Kleidung tragen, um die Körpertemperatur niedrig zu halten und nicht zu erröten), auf die sie eine geglückte Interaktion attribuieren, was zu einer Aufrechterhaltung der Störung führt (Clark, 2005).

3.1.2 Nonverbale Charakteristika

Patienten mit Angststörungen zeigen neben kognitiven und perzeptuellen Verzerrungen eine Reihe von nonverbalen Auffälligkeiten. Für Patienten mit Panikstörung wurden eine erhöhte Geschwindigkeit von Bewegungen, längere Bewegungen, Körperschwankungen (Perna et al., 2001) und unübliche Bewegungen (Jacob, Lilienfeld, Furman, Durrant, & Turner, 1989) festgestellt. Bei der Präsentation von bedrohlichen Bildern, tendieren Patienten mit Panikstörung dazu, ein *Freezing*-Verhalten zu zeigen (Lopes et al., 2009). Die Agoraphobie ist verbunden mit reduzierten Kopfbewegungen und einer Versteifung der Muskulatur (Brandt, Kugler, Schniepp, Wuehr, & Huppert, 2015). In Bezug auf den generellen Prozentsatz von gezeigten Bewegungen, zeigen Patienten mit Angststörungen eine Bewegungshäufigkeit von ca. 25% in einer Therapiesitzung (Ramseyer, 2008).

Speziell bei der Sozialen Angststörung können besondere nonverbale Charakteristika beobachtet werden. Das DSM-V beschreibt als mit der Diagnose assoziierte nonverbale Merkmale eine generell rigide Körperhaltung, inadäquaten Blickkontakt sowie eine gedämpfte Stimme (American Psychiatric Association, 2018). Dies wird von Forschungsergebnissen unterstützt. Patienten zeigen häufig eine Vermeidung von Blickkontakt, motorische Unruhe und Zittern (Clark & Wells, 1995; Del-Monte et al., 2016; Farabee, Ramsey, & Cole, 1993; Gilboa-Schechtman & Shachar-Lavie, 2013; Heerey & Kring, 2007; Horley, Williams, Gonsalvez, & Gordon, 2003; Kang & Gratch, 2012; Schneier, Rodebaugh, Blanco, Lewin, & Liebowitz, 2011). Weiterhin weisen sie seltener expressives Verhalten auf (z.B. Herstellen eines echten Lächelns; Del-Monte et al., 2013). Charakteristisch sind ebenfalls Veränderungen der Stimme wie beispielsweise eine zitternde Stimme, erhöhte Grundfrequenz und häufige Pausen, eine geringe Lautstärke sowie eine eher geringe Sprechgeschwindigkeit (Gilboa-Schechtman & Shachar-Lavie, 2013; Laukka et al., 2008; Scharfstein, Beidel, Sims, & Rendon Finnell, 2011). Es finden sich auch Hinweise darauf, dass Patienten mit Sozialer Angststörung generell ein geringeres Bewegungsniveau haben als gesunde Probanden (Levitan et al., 2012). Demnach stehen sich die Freezing-Hypothese (geringere Bewegung) mit der Unruhe-Hypothese (erhöhte Bewegungen) gegenüber. Weitere Studien zeigen, dass sich die Angst in einem Wechsel aus spezifischen Ruhe- und Bewegungsphasen äußert (Kreyenbrink, Joraschky, Konstantinidis, Neumann, & Lausberg, 2017), was eine Integration beider Hypothesen ermöglicht. Innerhalb

von Bewegungssequenzen wurde besondere motorische Unruhe gezeigt, in Ruhephasen jedoch eine geringere Körperschwankung und eine Muskelversteifung. Es könnte ebenfalls sein, dass die Soziale Angststörung eher mit Unruhe gekennzeichnet ist, die Patienten jedoch durch erhöhte Fokussierung auf die Symptome Bewegungen bewusst unterdrücken, um nicht aufzufallen. Dies könnte die entgegengesetzten Befunde ebenfalls erklären. Generell sind die Bewegungscharakteristika von Patienten mit Sozialer Angststörung jedoch noch unzureichend untersucht, um ein konsistentes Bild zu erhalten.

3.1.3 Therapie Sozialer Angststörungen

Symptomverbesserung

Zur Behandlung einer Sozialen Angststörung werden pharmakotherapeutische und psychotherapeutische Behandlungen durch die S3-Leitlinie empfohlen (Bandelow & Wedekind, 2014; Bandelow et al., 2014). Generell findet sich in Studien eine durchschnittliche Effektstärke von 1.7 für eine Abnahme der Angstsymptomatik von Beginn der Therapie zum Ende der Therapie (Fedoroff & Taylor, 2001). Betrachtet man lediglich randomisiert-kontrollierte Studien, so sind die berichteten Effekte etwas geringer. In Bezug auf Angstmaße wurde eine Verbesserung mit einer Effektstärke von 0.7 und bezüglich anderer Maße ebenfalls vergleichbare Effektstärken (kognitive Maße: $ES = 0.8$, Depression: $ES = 0.7$) gezeigt (Acarturk, Cuijpers, Van Straten, & De Graaf, 2009). Die kognitive Verhaltenstherapie ist der tiefenpsychologisch fundierten Therapie dabei etwas überlegen (Leichsenring et al., 2013). Die Effekte pharmakotherapeutischer Behandlungen liegen bei einer durchschnittlichen Effektstärke von 0.9. Damit sind sie psychotherapeutischer Therapien nicht überlegen. Betrachtet man Langzeiteffekte, so sind psychotherapeutische Therapien der medikamentösen Behandlung deutlich überlegen (Clark et al., 2003; Stangier, Clark, Ginzburg, & Ehlers, 2016).

Therapieabbruch

Unter einem Therapieabbruch versteht man nach Swift, Greenberg, Tompkins, und Parkin (2017) ein Beenden einer Therapie, das unilateral initiiert wird und gegen die Empfehlungen der Therapeuten stattfindet. Dadurch wird die Therapie beendet bevor die Probleme, wegen denen die Therapie begonnen wurde, gelöst sind. Die Möglichkeiten einen Therapieabbruch zu definieren, werden in verschiedene Arten unterteilt: Längenbasierte Definitionen und Qualitäts-/Outcome-basierte Definitionen. Unter die längenbasierten Definitionen fallen z.B. die Definition von Abbruch als Beendigung vor Absolvierung einer bestimmten Anzahl von Therapiestunden oder das fehlende Absolvieren des kompletten Behandlungsprotokolls (Swift & Greenberg, 2012). Zu den Qualitäts-/Outcome-basierten

Definitionen zählen beispielsweise die Definition von Abbruch als Beendigung vor Eintritt einer klinisch bedeutsamen Veränderung (*clinically significant change*, CSC) oder die Abbruchsdefinition nach Einschätzung des Therapeuten (Swift & Greenberg, 2012).

Abbrüche einer Psychotherapie und Pharmakotherapie sind häufig auftretende Phänomene, die in mehrerer Hinsicht problematisch sind. Die Meta-Analyse von Swift et al. (2017) zeigte, dass die Abbruchhäufigkeit im Durchschnitt bei 21.9% liegt. Für Patienten mit Angststörungen liegt die Anzahl der Abbrüche mit 15.1% nur gering niedriger (Hans & Hiller, 2013). Als schädliche Wirkung von Abbrüchen werden die fehlende Verbesserung des psychischen Funktionsniveaus, daraus folgend Auswirkungen auf Familie, Freunde und Kollegen, die Erhöhung des Risikos einer Chronifizierung sowie entstehende Frustration des betreffenden Therapeuten genannt. Außerdem kann ein Abbruch bei medikamentöser Therapie sogar lebensbedrohlich sein (Swift et al., 2017). Demzufolge werden viele Versuche unternommen, Prädiktoren auf Patienten- oder Therapeutenebene zu identifizieren, die die Wahrscheinlichkeit für ein vorzeitiges Beenden einer Therapie erhöhen. Die bisherige Forschung liefert jedoch weder auf Patienten- noch auf Therapeutenebene konsistent zuverlässige Prädiktoren für die Identifikation von Risikopatienten, weshalb in den letzten Jahren häufiger die Untersuchung von Prozess- und dyadischen Variablen (d.h. Variablen der Therapeut-Patient-Interaktion) gefordert wurde (Swift & Greenberg, 2012; Wierzbicki & Pekarik, 1993). Auch für die Soziale Angststörung sind die bisher identifizierten Prädiktoren nicht konsistent. Eskildsen, Hougaard, und Rosenberg (2010) zeigten innerhalb eines Reviews, dass es einige Studien zu demografischen und klinischen Patientenvariablen als Prädiktoren für Abbrüche gibt, diese jedoch nicht zuverlässig sind (Tabelle 2). Gerade für diese Patientengruppe bietet die nonverbale Synchronie als eine dyadische Variable eine vielversprechende Möglichkeit, eventuelle Risikopatienten für einen Therapieabbruch besser zu identifizieren.

Tabelle 2 Anzahl der Studien und Effekte, die Prädiktoren für Therapieabbruch bei Patienten mit Sozialer Angststörung untersuchten (nach Eskildsen, 2010)

| Prädiktor | Anzahl Studien | Positiver Effekt | Negativer Effekt | Kein Effekt |
|-----------------|----------------|------------------|------------------|-------------|
| Depression | 7 | 1 | | 6 |
| Vermeidende PKS | 4 | | | 4 |

| | | | | |
|-------------------------------------|---|---------------------|---|---|
| Andere Persönlichkeitseigenschaften | 2 | | | 2 |
| Angst komorbid | 4 | | | 4 |
| Andere Komorbiditäten | 2 | | | 2 |
| Symptomschwere | 9 | 1 | 1 | 7 |
| Demografische Variablen | 6 | 2 (jung & weiblich) | | 4 |

Anmerkung: PKS = Persönlichkeitsstörung.

3.2 Nonverbale Synchronie

Innerhalb menschlicher Interaktion kann oft beobachtet werden, dass sich das Verhalten zweier Interaktionspartner aufeinander bezieht oder einander angleicht. Diese Abhängigkeit der Interaktion zweier oder mehrerer Personen äußert sich in *Synchronisationsphänomenen*. Solche Synchronisationsphänomene sind beispielsweise aufeinander bezogene Bewegungen, gleiche Sitzhaltungen, aufeinander abgestimmte Stimmfrequenzen oder gleiche emotionale Gesichtsausdrücke.

Neben dem übergeordneten Begriff Synchronisationsphänomene, existieren eine Vielzahl von Begriffen, die dazu benutzt werden, die Abhängigkeit der Interaktion zu beschreiben (z.B. Imitation, Mimicry, Synchronie, Timing, Koordination, usw.) (Altmann, 2013; Condon & Ogston, 1966; Fogel, 1993; Hale & Hamilton, 2016; Tschacher & Ramseyer, 2017). Mehrere theoretische Arbeiten fassen die verschiedenen Synchronie-Begriffe zusammen und versuchen, sie voneinander abzugrenzen. Dabei fällt immer wieder auf, dass Autoren die Begriffe inkonsistent verwenden (Altmann, 2013) und auch teilweise je nach Forschungsgebiet verschiedene Begriffe verwendet werden, um das gleiche Konstrukt zu benennen (Paxton, 2015). Abbildung 2 fasst die Abgrenzung und Zusammengehörigkeit der am häufigsten verwendeten Begriffe in einem Venn-Diagramm zusammen³.

³ Eindruck der Erstautorin der Arbeit

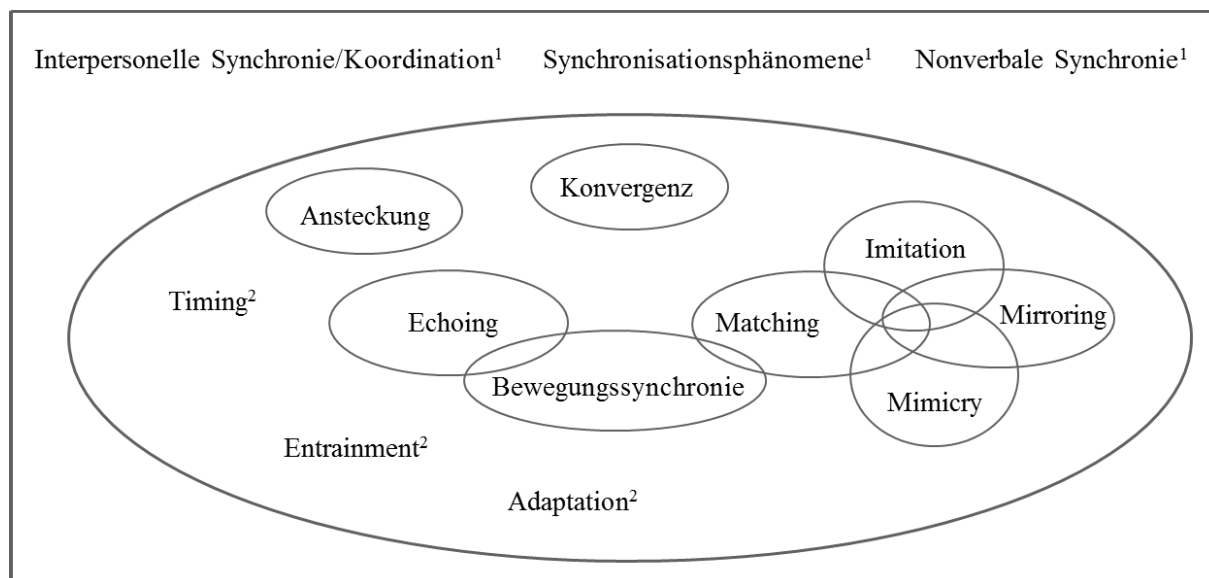


Abbildung 2 Venn-Diagramm mit verschiedenen Begriffen für interpersonelle Synchronie, ¹in dieser Arbeit synonym verwendete Oberbegriffe, ²in der Literatur ebenfalls verwendete Oberbegriffe, die nicht eindeutig abgrenzbar sind

Paxton (2015) grenzte in einer qualitativen Arbeit die Begriffe Akkommodation, Adaptation, Alignment, Mimicry, Ansteckung, Mirroring, Synergie und Synchronie voneinander ab und suchte nach einem geeigneten Oberbegriff für alle Phänomene. Anhand der lexikalischen Analyse mehrerer Abstracts aus dem Forschungsbereich wurde *interpersonelle Synchronie* bzw. interpersonelle Koordination als geeigneter Oberbegriff identifiziert. Dieser Oberbegriff wird auch in vorliegender Arbeit als übergeordneter Begriff für sämtliche Synchronisationsphänomene verwendet. Die wichtigsten Charakteristika interpersoneller Synchronie wurden unter anderem von Altmann (2013) zusammengefasst: 1) in Interaktionen von mindestens zwei Personen, 2) Bezug zu Verhalten oder Emotionen der Interaktionspartner, 3) Voraussetzung: Aktivität Interaktionspartner, und 4) Aufeinander-Bezugnehmen der Verhaltensweisen/Emotionen. Dabei ist die interpersonelle Synchronie keine Eigenschaft eines einzelnen Individuums. Sie realisiert sich stets auf der Interaktionsebene und stellt eine Komplexitätsreduktion dar, die sich zwischen zwei oder mehreren Individuen ereigneten wechselseitigen Umwelt dar (Tschacher & Storch, 2010). Da es sich bei den in dieser Arbeit beschriebenen Synchronisationsphänomenen um Phänomene auf nonverbaler Ebene handelt, wird der Begriff *nonverbale Synchronie* synonym zu interpersoneller Synchronie verwendet.

Hinsichtlich der Charakteristika von nonverbaler Synchronie liegen wichtige Unterscheidungsmerkmale verschiedener Synchronie-Begriffe vor allem auf der Ebene der Zeit (simultane Verhaltensweisen vs. zeitversetzte Verhaltensweisen), der Ebene des imitierten

Verhaltens (ähnliche Verhaltensweisen z.B. Bewegungssynchronie oder genau identische Verhaltensweisen z.B. Mimicry, Mirroring) und der Dynamik des Prozesses (statisch oder dynamisch). In Bezug auf die Messung von Synchronisationsphänomenen kann weiterhin die Verhaltensmodalität (Mimik, Bewegungen, Stimme, etc.), das erfasste Synchronisationsmaß (Stärke vs. Häufigkeit der Synchronie) und die verwendete Methode (lineare/ non-lineare Algorithmen, Ratings) für eine genauere Spezifikation des Konstruktes genutzt werden (z.B. facial mimicry, Häufigkeit von Bewegungssynchronie).

Im Gegensatz zu Mimicry oder Mirroring, versteht man unter *Bewegungssynchronie*, einer Form der nonverbalen Synchronie, den dynamischen Prozess, mit dem sich Verhalten auf einem zeitlichen Level gleicht. Das heißt, wenn eine Bewegung des Therapeuten mit einer Bewegung des Patienten einhergeht, so bezeichnet man dieses Verhalten als bewegungssynchron, auch wenn nicht die gleichen Bewegungen oder Bewegungen der gleichen Körperteile gezeigt wurden. Das wichtige Element dieser Definition ist nicht die Bewegung selbst, sondern die zeitliche Kopplung von Bewegungen (Altmann, 2013; Chetouani, Delaherche, Dumas, & Cohen, 2017; Ramseyer & Tschacher, 2011). Demnach kann Bewegungssynchronie als eine Bezogenheit der Bewegungsdynamik über die Zeit definiert werden. Sie operationalisiert einen Aspekt der Beziehungsqualität und das Miteinander-in-Kontakt-treten (Workshop nonverbaler Synchronisation; Strauss et al., 2018). In dem ersten deutschen Workshop der nonverbaler Synchronisation wurden ebenfalls spezifische Merkmale von Bewegungssynchronie herausgearbeitet: 1) Zeitraum der Etablierung kann variieren (Synchronisationsintervalle vs. Synchronie als über die Zeit entstehender Prozess), 2) findet unbewusst statt (vgl. Oullier, De Guzman, Jantzen, Lagarde, & Scott Kelso, 2008; Van Baaren et al., 2009), 3) ist nicht semantisch, 4) tritt simultan und mit Zeitverzögerung auf (*matching* und *echoing*), 5) findet in verbaler Interaktion statt (Ausschluss: Fragebogen ausfüllen, Progressive Muskelrelaxation, etc.), 6) eine generelle Wahrnehmung der Interaktionspartner ist nötig und 7) ist nicht losgelöst von Psyche, sondern stellt ein Embodiment psychischer Prozesse dar (Strauss et al., 2018).

3.2.1 Nonverbale Synchronie im Zeitverlauf

Die nonverbale Synchronie als Merkmal einer Dyade ist keine konstante Eigenschaft. Codrons, Bernardi, Vandoni, und Bernardi (2014) zeigten, dass Synchronie in Gruppen über die Sitzungen zunimmt. Im Psychotherapie-Kontext wurden hingegen Veränderungen in verschiedene Richtungen beobachtet. Tschacher, Scheier, und Grawe (1998) fanden in Bezug auf aus Stundenbögen bestimmter Synchronie eine stetige signifikante Zunahme über die Dauer

einer Therapie. Abfallende oder umgekehrt u-förmige Zusammenhänge traten jedoch ebenfalls auf. Ramseyer und Tschacher (2008) berichteten eine stetige Zunahme von Synchronie über den Therapieverlauf von Sitzung 1 bis Sitzung 40. Die Ergebnisse von Paulick et al. (2018) deuten darauf hin, dass verschiedene Störungsgruppen verschiedene Synchronie-Verläufe haben. Die Autoren stellten innerhalb einer Stichprobe ängstlicher Patienten eine Abnahme der gesamten Synchronisationsstärke fest. Für Depressive nahm die Synchronisationsstärke innerhalb der Therapie zu. Dabei ist der Verlauf ebenfalls vom Kontext abhängig. Sun, Nijholt, Truong, und Pantic (2011); Sun, Truong, Pantic, und Nijholt (2011) zeigten, dass die stimmliche Synchronie am Anfang einer Diskussion hoch ist, wenn dargestellt werden soll, dass man gleicher Meinung ist. Am Ende der Diskussion, wenn jeder mehr seinen eigenen Standpunkt einnimmt, sinkt das Niveau der Synchronie wieder.

3.2.2 Einflussfaktoren auf Synchronie

Synchronie wird durch mehrere förderliche oder hemmende Faktoren beeinflusst. Sie wird von Charakteristika der eigenen Person auf demografischer, kognitiver sowie sozialer Ebene (z.B. prosoziale Orientierung, Geschlecht, Bindung, Wunsch nach Beziehungsaufbau), Charakteristika des Interaktionspartners (z.B. Attraktivität, Gruppenzugehörigkeit, Stigmatisierung), dyadischen Charakteristika (z.B. Ähnlichkeit, bestehender Rapport) und dem Interaktionsinhalt sowie dem Kontext (z.B. kooperative Aufgabenbewältigung, freundliche vs. konfliktreiche Situation) beeinflusst. Verschiedene Einflussfaktoren, ihre Wirkung und forschungsbasierte Evidenzen sind beispielhaft in Tabelle 3 zusammengefasst.

Tabelle 3 Einfluss verschiedener Faktoren auf interpersonelle Synchronie

| Einflussfaktor | Synchronie fördernde | Synchronie hemmende | Beispielstudien |
|---|----------------------------|---|---|
| | Ausprägung | Ausprägung | |
| | Charakteristika der Person | | |
| Orientierung | Prosozial | Pro-self | Chartrand und Lakin (2013); Duffy und Chartrand (2015), Lumsden, Miles, Richardson, Smith, und Macrae (2012) |
| Soziale Kompetenz | Vorhanden | Nicht vorhanden | Schmidt, Christianson, Carello, und Baron (1994) |
| Sozialer Status | Vorhanden | Nicht vorhanden | Miles et al. (2011) |
| Bindung | Sichere Bindung | Unsichere Bindung | Hall, Millings, und Bouças (2012); Ramseyer und Tschacher (2008) |
| Affiliationsziel (implizit oder explizit) | Vorhanden | Nicht vorhanden | Chartrand und Lakin (2013); Duffy und Chartrand (2015), Lakin und Chartrand (2003) |
| Geschlecht | Weiblich | Männlich | van Straaten, Engels, Finkenauer, und Holland (2008); Zhao et al. (2015) |
| Psychische Störung | Keine | Störungen mit interpersonellen Einschränkungen (z.B. Schizophrenie, Autismus) | La France, Hecht, und Paluck (2003); La France und Ickes (1981); Namy, Nygaard, und Sauerteig (2002) Kupper, Ramseyer, Hoffmann, Kalbermatten, und Tschacher (2010); Nummenmaa, Lahnakoski, und Glerean (2018); Tschacher, Giersch, und Friston (2017); Varlet et al. (2014) |

| Charakteristika des Interaktionspartners | |
|--|---|
| Gruppenstatus | - in eigener Gruppe, wenn Gleichheit betont - mit anderer Gruppe, wenn Einheit zwischen Gruppen hergestellt wird - mit anderer Gruppe, wenn Einheit zwischen Gruppen hergestellt wird |
| Attraktivität des Interaktionspartners | - in eigener Gruppe, wenn Einheit zwischen Gruppen hergestellt wird - mit anderer Gruppe, wenn Gleichheit in eigener Gruppe betont |
| Ähnlichkeit | Nicht vorhanden |
| Soziales Stigma | Nicht vorhanden |
| Dyadische Charakteristika | |
| Rapport | Nicht vorhanden |
| Einstellung | übereinstimmend |
| Partnerschaft | Nicht vorhanden |
| Kontext und Interaktionsinhalt | |
| Priming mit prosozialen Inhalten | Nicht vorhanden |

Miles, Lumsden, Richardson, und Macrae (2011)
Chartrand und Lakin (2013); Duffy und Chartrand (2015),

Johnston (2002); Miles, Griffiths, Richardson, und Macrae (2009)

Chartrand und Lakin (2013); Duffy und Chartrand (2015)
Johnson (2002)

Chartrand und Lakin (2013); Duffy und Chartrand (2015)
Duran und Fusaroli (2017)
Chartrand und Lakin (2013)

Leighton, Bird, Orsini, und Heyes (2010)

| | | | |
|----------------------|---|--|---|
| Kontext | - Positiv - Neutral | - Negativ - Streit | Altmann (2011, 2013); Kimura und Daibo (2006); Paxton und Dale (2013a, 2017) |
| Aufgabe | - Kooperativ - Kompetitiv - Spaß-Aufgaben - Mit Affiliationsziel | - Nicht kooperativ - Mit professionellem Ziel | Allsop, Vaitkus, Marie, und Miles (2016); Yun, Watanabe, und Shimojo (2012) Tschacher, Rees, und Ramseyer (2014) |
| Bewusstheit | Nicht vorhanden | Vorhanden | Kulesza, Dolinski, und Wicher (2016) |
| Spiegeln | | | Van Baaren, Janssen, Chartrand, und Dijksterhuis (2009) |
| Kognitive Auslastung | Hoch | Gering | Chartrand & Lakin (2013) |

Es ist anzumerken, dass die Studien die Wirkung verschiedener Variablen auf Synchronisationsphänomene verschiedener Verhaltensmodalitäten (z.B. Bewegungen vs. Mimik) untersuchen. Die Verhaltensmodalität könnte eine wichtige Moderatorvariable für die dargestellten Einflüsse sein, wurde jedoch noch unzureichend und wenig explizit untersucht.

Im klinischen/psychotherapeutischen Setting

Innerhalb des klinischen/psychotherapeutischen Kontextes kann das Ausmaß an Aktiven Zuhörens die Synchronisationshäufigkeit beeinflussen. In medizinischen Rollenspielen wurde in Szenarien, in denen Aktives Zuhören gezeigt wurde, häufiger synchronisiert als in Szenarien ohne Aktives Zuhören (Hart et al., 2016). Weiterhin spiegeln empathische Personen ihr Gegenüber häufiger als Personen mit niedriger Empathie (Chartrand & Bargh, 1999). Einen weiteren wichtigen Einflussfaktor stellt die psychische Störung von Patienten dar. Dabei werden vor allem Störungen mit verminderten interpersonellen Kompetenzen (z.B. Schizophrenie, Autismus, Borderline Persönlichkeitsstörung und Soziale Angststörung) mit einer veränderten Synchronisationsfähigkeit in Verbindung gebracht (Duffy & Chartrand, 2015; Varlet et al., 2014). Dabei scheint Schizophrenie zum Beispiel mit einem geringen Ausmaß an interpersoneller Synchronie assoziiert zu sein (Kupper, Ramseyer, Hoffmann, & Tschacher, 2015). Da in der Arbeit eine Stichprobe von Patienten mit Sozialer Angststörung verwendet wurde, soll auf Ergebnisse im Hinblick auf diese Störung separat eingegangen werden.

Interpersonelle Synchronie bei sozialen Ängsten

Für sub-klinische Stichproben sozial ängstlicher Personen wurde der Einfluss von Synchronie bisher nur wenig untersucht; für klinische Stichproben beinahe gar nicht. Studien zum Thema lassen sich zwei verschiedenen Hypothesen zuordnen. Die erste Hypothese ist, dass sozial ängstliche Personen generell eine geringere Fähigkeit zur Verhaltensspiegelung und Imitation vorweisen, was dazu führt, dass Interaktionen mit sozial ängstlichen Personen als weniger positiv wahrgenommen werden. Dies bestätigt die angenommenen Ängste der sozial ängstlichen Personen und führt zu einer anhaltenden Vermeidung sozialer Situationen. Diese Hypothese wird dadurch gestützt, dass in sozial ängstlichen Personen generell ein geringeres Level an interpersoneller Synchronie in einer Interaktion mit einem Avatar (Vrijnsen, Lange, Becker, & Rinck, 2010), als auch partiell in einer realen Situation (Gesprächspartner wird ein Foto beschrieben) gefunden wurde (Abbott, Kocovski, & Obhi, 2018). Die Interaktion wird von den Gesprächspartnern ebenfalls nicht so positiv eingeschätzt (Heerey & Kring, 2007).

Weiterhin scheinen die zumeist positiven Effekte von Synchronisationsphänomenen nicht auf sozial ängstliche Personen zu wirken. Es konnte gezeigt werden, dass in Interaktion mit einem Avatar keine Erhöhung der Positivität, eingeschätzt durch die sozial ängstlichen Personen, gefunden werden konnte (Vrijen, Lange, Dotsch, Wigboldus, & Rinck, 2010). Gestützt wird diese Hypothese durch das kognitive Modell der Sozialen Angststörung. Durch den erhöhten Fokus auf das eigene Selbst sind Effekte sozialer Interaktion nicht so wirkungsvoll (Clark & Wells, 1995).

Die zweite Hypothese ist, dass Personen mit sozialen Ängsten generell ein höheres Level an Verhaltenspiegelung in Bezug auf positive Expressionen haben. Dijk, Fischer, Morina, van Eeuwijk, und van Kleef (2018) stellten keinen Zusammenhang zwischen dem Level von genereller Mimicry und sozialer Ängstlichkeit fest. Sie fanden hingegen, dass Patienten mit Sozialer Angststörung negative Expressionen eher wahrnehmen, jedoch nur auf positive Expressionen (Lächeln) in dem Maße reagieren, dass sie vor allem Höflichkeitslächeln zeigten. Konform zu diesen Ergebnissen, zeigten Heerey und Kring (2007) eine geringe Reziprozität des Lächelns bei sozial ängstlichen Personen: Auf echtes Lächeln und Höflichkeitslächeln der Gesprächspartner reagierten sie fast ausschließlich mit Höflichkeitslächeln. Dies könnte darauf hindeuten, dass sozial ängstliche Personen vor allem positive Expressionen spiegeln, um von anderen akzeptiert zu werden. Für das Störungsbild der Sozialen Angststörung wäre dies plausibel, da durch die Spiegelung positiver Expressionen, die vorliegende Befürchtung negativ von anderen bewertet zu werden, vermieden werden könnte.

Eine Studie an einer klinischen Stichprobe von Patienten mit SAD zeigte, dass die nicht intentionale Synchronie nicht beeinträchtigt ist. Beeinträchtigungen konnten lediglich im Hinblick auf ein intentionales Herbeiführen von Bewegungssynchronie festgestellt werden, d.h. ein Imitieren des gezeigten Verhaltens ist für Patienten mit SAD nicht problematisch, jedoch das eigene Herbeiführen und Leiten der Synchronie (Varlet et al., 2014). In Bezug auf eine klinische Stichprobe von Patienten mit Angststörungen (spezifische und soziale Phobie, Agoraphobie, Panikstörung, Zwangsstörung, generalisierte Angststörung) konnte bisher gezeigt werden, dass zu Beginn des Therapieprozesses ein im Vergleich zu Depressiven erhöhtes Maß an Bewegungssynchronie gezeigt wird. Im Therapieverlauf nahm die Bewegungssynchronie ab (Paulick et al., 2018). Die Autoren erklären die erhöhte Synchronie durch eine Aufmerksamkeitsverschiebung zum sozialen Kontext durch die Fixierung auf externe Hinweisreize und die ständige Risikoabschätzung, ob die Befürchtungen sich

verwirklichen. Im Verlauf der Therapie lernen Patienten mit Angststörungen eine Aufmerksamkeitsverschiebung, die dazu führt, dass sie sich selbst eher wahrnehmen können und nicht ständig in Referenz zur Außenwelt handeln.

3.2.3 Konsequenzen von Synchronie

Viele Studien weisen einen Einfluss von interpersoneller Synchronie nach. Allem voran steht die Studie des berühmten *Chamäleon-Effekts*. Die Autoren Chartrand und Bargh (1999) zeigten, dass die bloße Imitation einer Körperbewegung dazu führt, den Imitierenden sympathischer einzuschätzen. Zahlreiche Meta-Analysen fassen die positiven Einflüsse experimentell induzierter Synchronie seither zusammen (Mogan, Fischer, & Bulbulia, 2017; Rennung & Göritz, 2016; Vicaria & Dickens, 2016). Die Konsequenzen interpersoneller Synchronie sind in Tabelle 4 beispielhaft zusammengefasst.

Tabelle 4 *Konsequenzen von interpersoneller Synchronie*

| Konsequenz | Richtung | Beispielstudien |
|--|----------|--|
| Interpersonelle Variablen | | |
| Bewertung des Interaktionspartners (Sympathie, Vertrauen, Kooperationsbereitschaft) | + | Chartrand und Bargh (1999); Hove und Risen (2009) |
| | – | Miles et al. (2011) |
| Gefühl anderen nah zu sein | + | Vicaria und Dickens (2016) |
| Bedürfnis dazuzugehören | + | Vicaria und Dickens (2016) |
| Gefühl eine Einheit darzustellen | + | Lakens (2010); Lakens und Stel (2011); Vicaria und Dickens (2016) |
| Prosoziales Verhalten und Gedanken (Hilfsbereitschaft, Compliance, kooperatives Verhalten) | + | Hove und Risen (2009); Lakens und Stel (2011); Marsh, Richardson, und Schmidt (2009); Mogan et al. (2017); Rennung und Göritz (2016); Vicaria und Dickens (2016) |
| Gegenseitiges Verständnis | + | Shockley, Richardson, und Dale (2009) |
| Beziehungsaufbau | + | Koole und Tschacher (2016) |

3 Theoretische Einbettung in den bisherigen Forschungsstand

| | | |
|---|---|--|
| Überzeugungskraft | + | Chartrand und Lakin (2013); Maddux, Mullen, und Galinsky (2008) |
| Genauigkeit der Emotionswahrnehmung | + | Chartrand und Lakin (2013) |
| Evaluation der Interaktion | + | Vicaria und Dickens (2016) |
| Intrapersonelle Variablen | | |
| Kreatives Denken | + | Ashton-James und Chartrand (2009); Webb, Rossignac-Milon, und Higgins (2017) |
| Konfliktlösung | + | Maddux et al. (2008); Webb et al. (2017) |
| Konfliktbearbeitung | – | Bänninger-Huber und Huber (2017); Benecke und Krause (2005b) |
| Merken Gedächtnisinhalte in Bezug auf den Interaktionspartner | + | Macrae, Duffy, Miles, und Lawrence (2008); Vicaria und Dickens (2016) |
| Affekt | + | Mogan et al. (2017) |
| Soziale Angst | – | Gueguen (2011) |
| Selbstbewusstsein | + | Lumsden, Miles, und Macrae (2014) |
| Handeln frei von Stereotypen | – | Leander, Chartrand, und Wood (2011) |
| Anspannung/Kortisollevel Stress | – | Kouzakova, van Baaren, und van Knippenberg (2010) |
| Selbstregulationsfähigkeit | + | Chartrand und Lakin (2013); Koole und Tschacher (2016) |

Anmerkung: + steht für einen positiven Effekt, – für einen negativen Effekt.

Ob diese Effekte auch für den Imitierenden selbst auftreten, ist bisher noch unklar. Exemplarische Studien zeigen positive (Stel & Vonk, 2010), negative (Kot & Kulesza, 2016) oder keine Effekte für den Imitierenden (Hale & Hamilton, 2016). Eine wichtige Moderatorvariable stellt der Kontext bzw. das Ausmaß an erwarteter Synchronie dar. Ist die gezeigte Synchronie im Kontext angemessen und das Maß erwartet, treten eher positive Konsequenzen auf (Hale & Hamilton, 2016). Ein weiterer Moderator ist die Verhaltensmodalität. Mimicry führt zu höheren prosozialen Konsequenzen als Synchronie

(Rennung & Göritz, 2016; Vicaria & Dickens, 2016). Die Länge der Interaktion und der Anteil von weiblichen Versuchspersonen waren keine signifikanten Moderatoren. Einen weiteren interessanten Moderator identifizierten Mogan et al. (2017) in der Gruppengröße. Wobei eine höhere Gruppengröße positiv mit prosozialem Handeln und positivem Affekt zusammenhing, ging eine geringere Gruppengröße mit höheren positiven Effekten der interpersoneller Synchronie auf das *Social Bonding* und sozialen Kognitionen einher (Mogan et al., 2017).

Im klinischen und psychotherapeutischen Setting werden vor allem die Assoziationen zwischen nonverbaler Synchronie und Therapieoutcome (Bänninger-Huber & Huber, 2017; Paulick et al., 2017; Ramseyer & Tschacher, 2011; Rasting & Beutel, 2005), Patientenzufriedenheit (Benecke & Krause, 2005b; Greene, Adelman, Friedmann, & Charon, 1994), wahrgenommener Kompetenz (Mogan et al., 2017), therapeutischer Allianz (Koss & Rosenthal, 1997; Ramseyer & Tschacher, 2008, 2011; Reich et al., 2014) sowie Empathie (Behrends, Müller, & Dziobek, 2012; Fraenkel, 1983; Kleinbub, 2017; Kleinbub et al., 2012; Koehne, Hatri, Cacioppo, & Dziobek, 2016; Lovkvist, 2013) und Bindung (Hall et al., 2012; Håvås, Svartberg, & Ulvenes, 2015) diskutiert. Auf die Zusammenhänge hinsichtlich des Therapieerfolgs und Bindung wird im Folgenden gesondert eingegangen, da sie Untersuchungsgegenstand der in der Rahmenschrift zusammengefassten Publikationen sind.

Nonverbale Synchronie und Therapieoutcome

Die Befunde in Bezug auf die Assoziation zwischen Therapieoutcome und nonverbaler Synchronie sind überwiegend gemischt. Es gibt Studien, die einen positiven oder negativen Einfluss von Synchronie auf das Therapieergebnis sowie einen umgekehrt u-förmigen Zusammenhang von Synchronie und Therapieerfolg beschreiben. Die Studien sind in Tabelle 5 zusammengefasst dargestellt.

Tabelle 5 Studien mit ihren Ergebnissen zum Zusammenhang zwischen interpersoneller Synchronie und Therapieoutcome

| Studie | Stichprobe | Setting | Untersuchungs- gegenstand | Ergebnis | Wirkung Synchronie auf Therapieerfolg |
|--------------------------------------|--|--|------------------------------|---|--|
| Ramseyer und Tschacher (2011) | Störungsheterogene Stichprobe (N = 104) | Naturalistische KVT | Bewegungssynchronie | - Verbindung hohe Synchronie mit Symptomreduktion | + |
| Ramseyer und Tschacher (2016) | Patientin mit Depression (N = 1) | Naturalistische KVT | Bewegungssynchronie | - Handbewegungssynchronie hängt mit Sitzungserfolg zusammen | + |
| Nagaoka und Komori (2008) | Gesunde Personen (N = 4) | Psychotherapeutische Beratung, Rollenspiele | Bewegungssynchronie | - in positiven Beratungen mehr Synchronie | + |
| Galbusera, Finn, und Fuchs (2016) | Patienten mit Schizophrenie (N = 16) | Manualisierte Körper- orientierte Psychotherapie | Bewegungssynchronie | - Zunahme Synchronie über Therapieverlauf, Abnahme Symptomstärke - geringer marginal signifikanter Effekt für Zusammenhang Symptome und Synchronie bei Patienten mit steigender Synchronie | (+) |

| | | | | | |
|--|---|---|------------------------|---|---------------------|
| Bänninger-Huber und Huber (2017) | Störungsheterogene Stichprobe (N = 6) | Psychoanalytische Langzeittherapien | Gleichzeitiges Lächeln | - negativen Einfluss auf Therapieerfolg - positiven Effekt auf positive Gefühle Therapeuten gegenüber | - |
| Rasting und Beutel (2005) | Störungsheterogene Stichprobe (N = 20) | Kurzzeit- Stationsaufenthalte (psychodynamisch, multiprofessionell) | Gleichzeitiges Lächeln | - erfolgreiche Therapien mit wenig reziprokem Lächeln gekennzeichnet, in erfolglosen viel | - |
| Reich, Berman, Dale, und Levitt (2014) | Studierende (N = 157) | Psychotherapeutische Beratung | Stimmliche Synchronie | - viel gleichzeitiges Lächelns mit positivem Affekt assoziiert | - |
| Paulick et al. (2017) | Störungsheterogene Stichprobe (N = 143) | Naturalistische KVT | Bewegungssynchronie | Synchronie negativer Prädiktor für Depression - mittleres Synchronie-Level mit Verbesserung Symptome assoziiert - hohe Synchronie: kein Abbruch, keine Veränderung | Umgekehrt u- förmig |

| | | | | | | |
|-------------------------------|--|---------------------------|------------------------|--|--------------------|---|
| Paulick et al. (2017) | | | | - geringe Synchronie: Abbruch, keine Veränderung | | |
| Benecke und Krause (2005a) | Patientinnen mit Panikstörung (N = 20) | Psychodynamische Therapie | Gleichzeitiges Lächeln | - Extreme (stetiges gemeinsames Lächeln, kein negatives Lächeln) mit negativem Therapieerfolg verbunden | Umgekehrt u-förmig | |
| Ramseyer und Tschacher (2014) | Störungsheterogene Stichprobe (N = 70) | Naturalistische KVT | Bewegungssynchronie | - Synchronie der Kopfbewegungen mit globalem Outcome verbunden, Synchronie der Körperbewegungen mit Sitzungserfolg | | + |

Anmerkung: + steht für eine positive Wirkung, - für eine negative Wirkung, (+) für einen eingeschränkten/marginal signifikanten Zusammenhang mit besonderen Einschränkungen.

Die verschiedenen Befunde können jedoch auch dadurch verursacht sein, dass die Ergebnisse verschiedener Modalitäten (Stimme, Bewegung, Mimik), verschiedener Störungsbilder sowie Therapierichtungen miteinander verglichen werden. Weiterhin gibt es Hinweise darauf, dass selbst innerhalb der Bewegungssynchronie differentielle Effekte auftreten können. So zeigten Ramseyer und Tschacher (2014), dass das globale Therapieoutcome mit der Synchronie der Kopfbewegungen und dass Session-Outcome mit der Synchronie der Körperbewegungen assoziiert ist. Es ist von besonderer Relevanz, den Zusammenhang zwischen Therapieoutcome und interpersoneller Synchronie in mehreren Settings, störungshomogen sowie unter Berücksichtigung mehrerer Modalitäten zu untersuchen.

Nonverbale Synchronie und Bindung

Betrachtet man die bisherige Forschung in Bezug auf Bindung und interpersoneller Synchronie, so konnte innerhalb von Mutter-Kind Dyaden empirisch mehrfach eine positive Verbindung von interpersoneller Synchronie und sicheren Bindungsmustern, Emotionsregulationskompetenzen und der generell positiven Entwicklung nachgewiesen werden (Beebe et al., 2012; Feldman, 2012, 2015; Feldman, Magori-Cohen, Galili, Singer, & Louzoun, 2011; Isabella, Belsky, & von Eye, 1989; Lindsey & Caldera, 2015). Bisher gibt es im Erwachsenenalter und im klinischen Kontext wenige und teilweise nur indirekte Hinweise auf eine positive Beeinflussung der Regulationskompetenzen durch Synchronie. So beschreiben Ramseyer und Tschacher (2011), dass Bewegungssynchronie einen Einfluss auf den Therapieoutcome hat. Da die meisten Störungen auch Beeinträchtigungen der Emotionsregulation einschließen, nehmen Koole und Tschacher (2016) einen indirekten Einfluss von Synchronie auf Emotionsregulationskompetenzen an. Der Idee folgend, dass Bindung mit Regulationskompetenzen assoziiert ist (Dales & Jerry, 2008), zeigen weitere Studien eine Verbindung zwischen einem sicheren Bindungsstil und Bewegungssynchronie im klinischen Kontext (Ramseyer, 2008), und eine Assoziation von niedriger Synchronie und Bindungsunsicherheit (Ramseyer & Tschacher, 2011). Håvås et al. (2015) berichten, dass bei Patienten mit Cluster C Persönlichkeitsstörungen ein höheres *Affektattunement* mit einer geringeren Bindungsvermeidung assoziiert ist. Auch in Bezug auf Mimik konnte ein positiver Zusammenhang zwischen einem sicheren Bindungsmuster und einem höheren Niveau an Synchronie gezeigt werden (Hall et al., 2012). Während des Mirror Games, einer instruierten Übung, bei der die Bewegungen des Gegenübers nachgeahmt werden sollen, ist ein sicherer Bindungsstil mit einer höher eingeschätzten Komplexität und niedrigerer Synchronie

eindimensionaler Bewegungen assoziiert (Feniger-Schaal et al., 2016). Indirekte Evidenz für die Verbindung zwischen Bindung und Synchronie liefert überdies die Studie von Palmieri et al. (2018). Nach einem Bindungsprime synchronisierten die Therapeuten deutlich häufiger auf physiologischer Ebene. Die Autoren vermuten, dass der Prime dazu führt, dass der Therapeut die Interaktion mehr leitet und dadurch das Gefühl eines sicheren Hafens vermittelt.

3.2.4 Biologische Grundlagen interpersoneller Synchronie und Mechanismus

Betrachtet man die neuronalen Korrelate von nonverbaler Synchronie, so sind drei Systeme beteiligt: das Wahrnehmungs-Handlungs-Abgleich-System, das Selbst-andere System und das Belohnungszentrum (Hale & Hamilton, 2016). Das Wahrnehmungs-Handlungs-Abgleich-System wird voranging durch die Spiegelneuronen des Menschen im inferioren Parietalkortex und inferioren Frontalkortex gebildet (Hale & Hamilton, 2016; Iacoboni, 2009; Rizzolatti & Craighero, 2004). Wird eine Bewegung wahrgenommen, so führt die Aktivierung von Spiegelneuronen dazu, dass die gleichen neuronalen Repräsentationen für die Bewegung auch im Gehirn des Beobachters aktiviert werden (Rizzolatti et al., 1988). Da es gleiche Repräsentationen für die Handlung und die Wahrnehmung gibt, wird die Wahrscheinlichkeit eine ähnliche Handlung auszuführen, erhöht. Weiterhin führt assoziatives Lernen nach Hebb dazu, dass sensorische Informationen und die Verarbeitung von Handlungen und Handlungsplanung miteinander verbunden werden. Demnach scheint die gleichzeitige neuronale Aktivität die Basis für Empathie darzustellen (Iacoboni, 2009) und eine Grundlage zu sein, sich gegenseitig zu verstehen (Nummenmaa et al., 2018). Die Aktivierung des Spiegelneuron-Systems bei Synchronie konnte in Bezug auf experimentale Studien mit Fingertapping-Experimenten o.Ä. (Dumas, Laroche, & Lehmann, 2014; Novembre, Knoblich, Dunne, & Keller, 2017) sowie in sozialen Interaktionen (Hogeveen, Chartrand, & Obhi, 2014) empirisch belegt werden. Das Selbst-andere System umfasst Hirnregionen wie beispielsweise den rechten inferioren Parietalkortex und vor allem den temporalen parietalen Übergang. Bei erhöhter Überschneidung des Selbst mit anderen sind außerdem frontale Regionen (medialer Orbitofrontalkortex, ventromedialer Präfrontalkortex) aktiviert. Kinreich, Djalovski, Kraus, Louzoun, und Feldman (2017) verglichen neuronale Synchronie von Paaren und Fremden und zeigten, dass vor allem temporal-parietale Areale bei Paaren aufgrund der bestehenden Bindung eine höhere Aktivierung zeigten. Auf Verhaltensebene war diese neuronale Koppelung ebenfalls verbunden mit positivem Affekt und Blickkontakt. Diese Regionen sind ebenfalls in Bezug auf Spiegelneurone wichtig (Hale & Hamilton, 2016). Ein weiteres aktiviertes System ist das Belohnungssystem. Imitation führt zu einer Aktivierung vom

ventromedialen Präfrontalkortex, der Insula, des Striatums und des Nucleus caudatus, die ebenfalls aktiviert sind, wenn ein Individuum belohnt wird (Hale & Hamilton, 2016; Kühn et al., 2010). In Bezug auf Autismus, ein Störungsbild, was vor allem durch starke soziale Defizite und Probleme Kontakt zu anderen Personen herzustellen, gekennzeichnet ist, lässt sich keine Synchronie dieser neuronaler Netzwerke beobachten (Nummenmaa et al., 2018).

Als vermittelnder Mechanismus werden drei verschiedene Modelle diskutiert. Das erste Modell geht davon aus, dass durch die gleichzeitige Aktivierung von Spiegelneuron-System und Regionen, die verarbeiten, ob es sich um einen Selbst oder jemand anderen handelt, eine interpersonelle Nähe hergestellt wird und die Grenzen zwischen Selbst und anderem verschwimmen. Diese Nähe wird dann als positiv und belohnend wahrgenommen und führt zu prosozialen Verhaltensweisen (Hale & Hamilton, 2016; Mogan et al., 2017). Die erste Studie, die das Spiegelneuron-System als bei interpersoneller Synchronie beteiligt, nachwies, diskutiert eben diesen Mechanismus als Grundlage für die Wirkung interpersoneller Synchronie (Hogeveen et al., 2014). Das zweite Modell geht davon aus, dass die Kontingenz, die zwischen den eigenen Handlungen und der Welt gezeigt wird, einen positiven Einfluss hat. Dem Modell folgend, hat Synchronie einen positiven Einfluss auch wenn die gezeigten Handlungen nicht exakt gleich sind, wichtig ist lediglich die Kontingenz zwischen den Bewegungen. Konform des Modells konnten Studien zeigen, dass die Konsequenzen von Synchronie nicht von der Bewegungsähnlichkeit abhängig sind (Catmur & Hayes, 2013) und stark davon abhängen, welches Level an Synchronie in der erlebten Situation erwartet wird (Leander, Chartrand, & Bargh, 2012). Das dritte Modell geht davon aus, dass Menschen gern vorhersagen, was passiert. Demnach wird Synchronie der Bewegungen als belohnend empfunden, da wir das folgende Input aufgrund der Ähnlichkeit zu unseren Bewegungen besser vorhersagen können (Hale & Hamilton, 2016).

Auch auf endokrinologischer Ebene ist Synchronie mit geringeren Kortisollevel assoziiert (Kouzakova et al., 2010) und führt zu einer Ausschüttung endogener opioiden Peptiden, die ein Sicherheitsgefühl vermitteln (Machin & Dunbar, 2011).

3.2.5 In-Sync Modell der Psychotherapie

Koole und Tschacher (2016) postulieren in dem *Interpersonal Synchrony model of psychotherapy* (In-Sync Modell) die Wirkungsweise nonverbaler Synchronie innerhalb der Psychotherapie (Abb. 3).

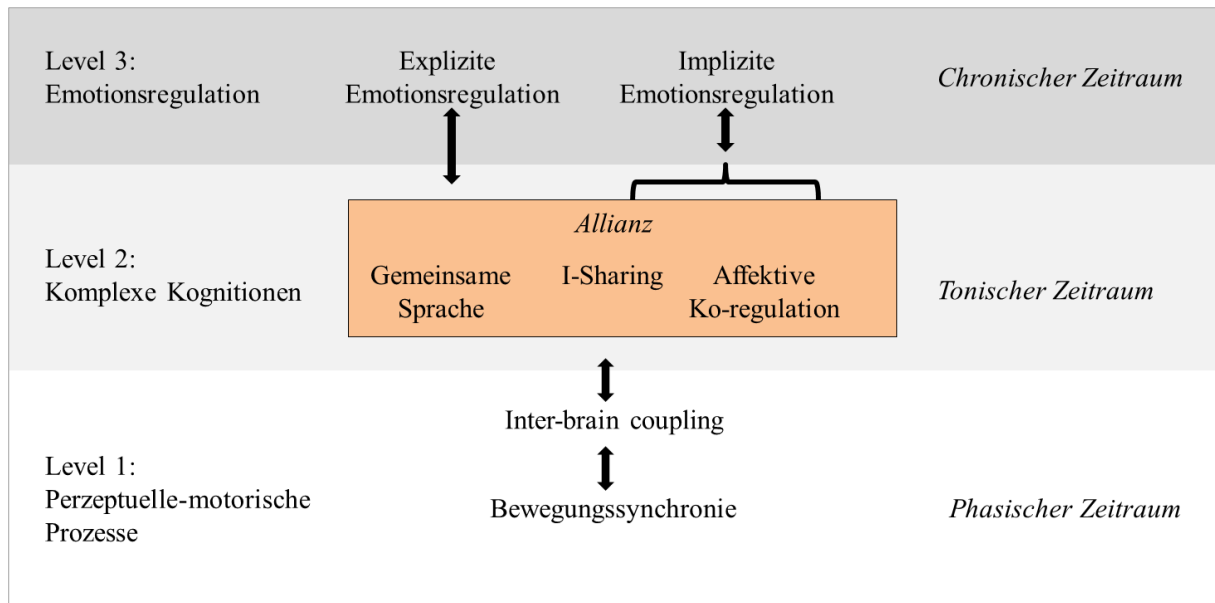


Abbildung 3 *In-Sync-Modell nach Koole und Tschacher (2016), S. 8*

Das Modell wurde entwickelt, um den Zusammenhang zwischen Bewegungssynchronie und therapeutischer Allianz darzustellen und zu erklären. Auf der ersten Ebene des Modells steht die neuronale und behaviorale Koordination des Patienten und Therapeuten. Findet diese nonverbale Synchronie statt, so führt die gleichzeitige Aktivierung der Motorareale im somatosensorischen Kortex der ausführenden Person und die Aktivierung der Spiegelneurone im präfrontalen Kortex der beobachtenden Person zu einer neuronalen Kopplung. Der Zeitraum dieser Kopplung wird als phasisch beschrieben, d.h. die Prozesse laufen auf Millisekunden-Ebene bis zu einer Zeitspanne von 10 Sekunden ab (Koole & Tschacher, 2016). Diese Kopplung auf perzeptuell-motorischer Ebene beeinflusst nachfolgend die Kognitionen auf nächster Ebene. Im Therapieprozess umfasst diese Ebene ebenfalls das Bilden einer gemeinsamen Sprache, das Teilen subjektiver Erlebnisse (I-sharing) und die gemeinsame affektive Koregulation, d.h., Affekte und physiologische Maße verändern sich nicht unabhängig, sondern in Referenz zum Interaktionspartner. Die Prozesse bewegen sich auf einer tonischen Zeitskala, sie nehmen 10 Sekunden bis hin zu einer Stunde in Anspruch. Die Bestandteile formen im In-Sync Modell die therapeutische Allianz. Diese wiederum wirkt auf höchster Ebene des Modells auf die explizite (bewusste) sowie implizite (unbewusste) Emotionsregulation des Patienten. Diese operiert auf einer chronischen Zeitebene, d.h. Veränderungen können Wochen bis hin zu Jahren persistent sein (Koole & Tschacher, 2016).

3.2.6 Erfassung Synchronie

Um den nonverbale Synchronie abzubilden und zu erfassen, haben sich in den letzten Jahren eine Reihe verschiedener Methoden etabliert. Alle wichtigen Begriffe der Erfassung nonverbaler Synchronie sind in Abbildung 4 zusammengefasst und nach ihren Prozessschritten geordnet. In den folgenden Absätzen sind die einzelnen Schritte näher beschrieben.

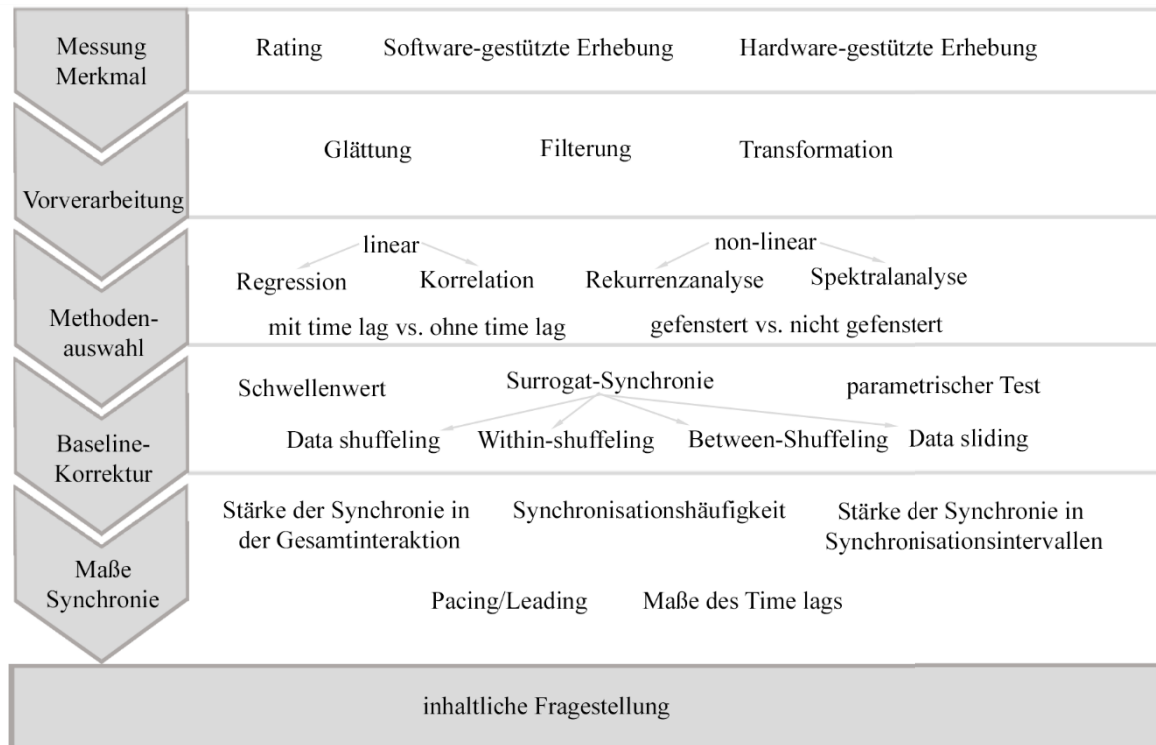


Abbildung 4 Überblick über wichtige Prozessschritte der Erfassung von nonverbaler Synchronie

Messung von Merkmalen

Die Messung nonverbaler Charakteristika wird manuell durch menschliche Ratings (*behavior coding*, z.B. Condon & Ogston) oder technikgestützt (z.B. Ramseyer und Tschacher, 2011) durchgeführt. Dabei wird zwischen kategorialer Erhebung (z.B. Lächeln vs. kein Lächeln) und kontinuierlichen Ratings (z.B. Geschwindigkeit der Bewegung) unterschieden. Vor dem manuellen Rating-Prozess werden Einheiten (*units*) unterteilt und Bezeichnungen (*labels*) für die einzuschätzenden Merkmale oder Merkmalsausprägungen vergeben. Die Zeitskala für die Ratings wird ebenfalls festgelegt (Chetouani et al., 2017; Delaherche et al., 2012). Bei manuellen Ratings ist die Anzahl der Rating-Zeitpunkte meist geringer als bei automatisierten Methoden. Da Interaktionen dynamische Prozesse darstellen, wurde in vergangenen Jahren die Anzahl der Messzeitpunkte erhöht, um nicht nur globale Einschätzungen, sondern auch Mikroprozessanalysen zu ermöglichen. Werden Merkmale

technikgestützt erhoben, so können verschiedene Ansätze unterschieden werden. 1) Hardware-basierte Erfassung: Durch bestimmte Hardware (z.B. Sensoren am Körper der Personen) ist es möglich, Bewegungen, Blickrichtung, physiologische Maße und andere nonverbale Signale direkt aufzuzeichnen und mit einer hohen zeitlichen Dichte zu speichern. Zu diesen Verfahren zählen beispielsweise *motion-tracking* oder *motion-capture* Ansätze. Ein weiteres großes Anwendungsfeld liefern Messinstrumente zur Erhebung physiologischer Daten: Durch bestimmte Sensoren können beispielweise Hautleitfähigkeit, Herzschlagfrequenz oder das im Schweiß messbare Kortisol-Level detektiert werden (Chetouani et al., 2017; Cornejo, Cuadros, Morales, & Paredes, 2017; Delaherche et al., 2012). Ebenfalls Hardware-basiert werden sämtliche neuropsychologische, hirnanatomische und hirnfunktionelle Aspekte erfasst (e.g., durch Elektroenzephalographie, Magnet-Resonanz-Tomographie, siehe Cornejo et al. (2017)). 2) Software-basierte Erfassung von Verhaltensweisen: Zu diesen Methoden zählen vor allem die *image-processing methods*. Liegen Videosequenzen von einer Interaktion vor, so ist es Software-basiert möglich, eine Vielzahl von relevanten nonverbalen Merkmalen zu erfassen. Dafür müssen als Hardware auch Videotechnik oder Mikrophone vorhanden sein. Der Unterschied besteht vor allem darin, dass bei Software-basierten Erfassungsmethoden die Hardware nicht direkt der Aufzeichnung der nonverbalen Marker dient. Diese werden später aus den gespeicherten Daten extrahiert. Beispielsweise können die Körperbewegungen durch die *Motion Energy Analysis*, eine Technik, die Pixelveränderungen nachfolgender Video-Bilder als Bewegung zählt, erfasst werden. Da diese Methode Grundlage für diese Arbeit ist, wird sie separat und detailliert im Abschnitt 4.1 erläutert. Die Bewegungsgeschwindigkeit kann durch Algorithmen, die den *optical flow* messen, bestimmt werden. Außerdem geben sogenannte *Motion History Images* ebenfalls Hinweise auf die Bewegungen der Interaktionspartner (Chetouani et al., 2017; Delaherche & Chetouani, 2010; Delaherche et al., 2012). Weitere beispielhafte Anwendungen wäre die Software OpenFace sowie Praat. Gesichtsausdrücke können durch die Software OpenFace kodiert werden. Stimmanalysen der Audiospuren im Hinblick auf die Grundfrequenz der Stimme können durch die Software Praat durchgeführt werden. Vorteile und Nachteile der beschriebenen Erfassungsansätze sind in Tabelle 6 zusammengefasst.

3 Theoretische Einbettung in den bisherigen Forschungsstand

Tabelle 6 Vor- und Nachteile verschiedener Erfassungsmethoden für nonverbale Merkmale

| | Vorteile | Nachteile |
|-----------------------------|---|--|
| Manuelle Ratings | <ul style="list-style-type: none"> - Sehr nah an menschlich wahrgenommenen Konstrukt | <ul style="list-style-type: none"> - Segmentierung von Verhalten - Beschreibungen für Einheiten oft nicht passend, Übergänge von einer Einheit zur nächsten - Reliabilität eingeschränkt, besonders bei Differenzierung mittleres Niveau - Keinen einheitlichen Rahmen für Ratings verschiedener Studien - Sehr hoher Kodieraufwand - Reduktion des real auftretenden Verhaltens - Abhängig vom Rater |
| Hardware basierte Erfassung | <ul style="list-style-type: none"> - Direkte, unverfälschte, hoch-reliable Messung des Merkmals - Objektivität hoch | <ul style="list-style-type: none"> - Hardware teuer in Anschaffung - Hardware am Körper beeinflusst normale Interaktion - Oft global (Marker über die Sequenz hinweg gemittelt), nicht lokal - Rating nicht multi-modal - Kontext, Situation oft nicht einbezogen |
| Software basierte Erfassung | <ul style="list-style-type: none"> - aus Videosequenz können Vielzahl an Markern (Mimik, Bewegung, Stimme) extrahiert werden - unabhängig vom Rater, hohe Reliabilität - Objektivität hoch | <ul style="list-style-type: none"> - Oft global (Marker über die Sequenz hinweg gemittelt), nicht lokal - Rating nicht multi-modal - Kontext, Situation oft nicht einbezogen |

Anmerkung: erweitert nach: Chetouani, 2017; Delaherche, 2010; Altdorfer, 1997

Vorverarbeitung der Messequenz & statistische Vorüberlegungen

Werden die Merkmale mit einer hohen zeitlichen Dichte erfasst, so resultieren Datensequenzen mit vielen Datenpunkten und gleichen Intervallen zwischen den Messungen. Solche Sequenzen bezeichnet man auch als Zeitreihen (Brockwell & Davis, 2013). Werden die Zeitreihen technikbasiert aufgezeichnet, so kann Hintergrundrauschen manchmal die Zeitreihen verunreinigen. Deshalb kann es hilfreich sein, die Zeitreihen vor der Synchronie-Berechnung noch einmal zu bereinigen. Weiterhin werden manchmal kleine Signalschwankungen durch die hohe zeitliche Dichte aufgezeichnet, die nicht von Bedeutung für eine nachfolgende Synchronisationsbestimmung sind. Um Hintergrundrauschen zu minimieren und kleine Signalschwankungen zu entfernen, wird die Zeitreihe geglättet. Zu diesem Zweck werden in der Interaktionsliteratur häufig *Smoothing Splines*, *Butterworth Filter* oder *Moving Average*-Transformationen durchgeführt (Altmann, 2013; Paxton & Dale, 2013b; Ramseyer & Tschacher, 2011). Bisher wurde der Einsatz solcher Verfahren allein an der Tatsache legitimiert, dass die Zeitreihen danach keine visuellen Schwankungen mehr aufweisen. Welchen Einfluss bestimmte Glättungsmechanismen und Glättungsstärken auf die Bestimmung der Synchronie haben, wurde noch nicht untersucht und ist unter anderem Bestandteil der Fragestellung der zweiten Publikation dieser Arbeit.

Werden Zeitreihen auf Grundlage von Videoaufzeichnungen erstellt, so können sie fehlerhafte Werte enthalten, die durch die schlechte Qualität des Videos resultieren. Videoaufzeichnungen mit durchgängig schlechter Qualität können vorher ausgeschlossen werden. Es gibt jedoch auch Videos, die lediglich für kurze Zeitpunkte qualitativ beeinträchtigt sind. Die Qualität wird durch nicht korrekte Pixelveränderungen oder Lichtblitze im Video verringert. Spielt man ein Video in langsamer Geschwindigkeit ab, so zeigen sich derartige Veränderungen durch komplett weiße Videoframes oder Sprünge des Bildes. Videos können automatisch auf solche Sprünge untersucht werden (Grammer, Honda, Juetten, & Schmitt, 1999; Ramseyer & Tschacher, 2011), indem die Zeitreihe auf eine große Anzahl an Pixelveränderungen, die nur für einen Frame auftreten, untersucht wird. Die so identifizierten Zeitpunkte sollten nachfolgend gesichtet und geprüft werden. Liegt ein Videofehler vor, kann der betreffende Wert durch den Mittelwert der beiden benachbarten Werte ersetzt werden (Grammer et al., 1999).

Darüber hinaus weisen Zeitreihen noch andere statistisch relevante Merkmale auf. Durch die wiederholten und zahlreichen Messungen sind Zeitreihen oft autokorreliert, d.h., ein Messpunkt zum Zeitpunkt t ist nicht unabhängig von vorangegangenen Messwerten $t-1$ oder $t-$

2 usw. Wird dieser Fakt bei der Bestimmung der Synchronie außer Acht gelassen, so kann es zur Detektion von sogenannten Scheinzusammenhängen (*spurious correlations*) kommen.

Eine weitere wichtige Annahme bei der Analyse von Zeitreihen ist die Stationarität. Bei stationären Prozessen sind Erwartungswert und die Varianz der Zeitreihe endlich und zeitunabhängig. Das bedeutet u.a., dass es keinen linearen Trend in der Zeitreihe gibt. Diese Annahme wird durch manche Zeitreihen verletzt. Dies trifft manchmal auch auf die den linearen Analysemethoden zu Grunde liegende Annahme der Normalverteilung der Daten zu. Zeitreihen, die aus der Motion Energy Analysis stammen, weisen zu einem sehr großen Anteil Nullen auf, da in einer Interaktion sehr oft keine Bewegung stattfindet. Demnach sind die Verteilungen schief und die Varianz ist nicht zeitunabhängig. Möglichkeiten zur Herstellung einer stationären Zeitreihe sind die Differenzbildung der ersten Differenz oder die Anwendung einer Varianz-stabilisierenden Transformation. Altmann (2013) zeigte, dass dabei die Anscombe-Transformation gut geeignet ist, die Verteilung einer Normalverteilung anzugleichen und die Varianz zu stabilisieren. Weitere Möglichkeiten bieten die Box-Cox-Transformation, die die normale Logarithmus-Transformation einschließt oder die Wurzel-Transformation (Altmann, 2013; Anscombe, 1948; Box & Cox, 1964). Boker, Rotondo, Xu, und King (2002) schlagen überdies vor, die Zeitreihe in Segmenten zu untersuchen. Die Annahmen der Stationarität und Normalverteilung wären somit für die einzelnen Segmente haltbar. Dafür dürfen die Segmente nicht zu groß gewählt werden. Zu kleine Segmente wiederum führen zwar nicht zu einer Verletzung der Annahmen, aber zu einer verfälschten Kovarianzschätzung im Rahmen der Bestimmung der Synchronie-Werte (Schönbrodt & Perugini, 2013). Der Schwellenwert für eine ideale Segmentgröße ist bisher noch nicht statistisch untersucht wurden.

Bestimmung der nonverbalen Synchronie

Nonverbale Synchronie kann nicht nur auf Grundlage von Zeitreihen nonverbaler Merkmale bestimmt werden. Gegenüber der *behavior coding method*, bei der einzelne nonverbale Verhaltensweisen kodiert werden, um anschließend Synchronie zwischen den Zeitreihen zu bestimmen, wurde von Bernieri, Davis, Rosenthal, und Knee (1994); Bernieri, Reznick, und Rosenthal (1988) die *judgment method* vorgeschlagen. Diese Methode orientiert sich an der Gestaltpsychologie und nimmt an, dass Rater Synchronisationsphänomene als solche in Sequenzen erkennen und bewerten können. Eine Einschätzung findet somit nicht mehr auf Mikroprozessebene statt, sondern für einzelne Zeitpunkte wird direkt wie von Bernieri et al. (1988) vorgeschlagen, die 1) Gleichzeitigkeit des Verhaltens 2) Gleichheit des

Tempos der Bewegungen und die 3) Geschmeidigkeit und Koordination eingeschätzt (vgl. Kimura & Daibo, 2006).

Dieser manuellen Methode mit eingeschränkter Reliabilität stehen zahlreiche Algorithmen-basierte Bestimmungsmethoden für die nonverbale Synchronie gegenüber, die die gewonnenen Zeitreihen analysieren. Die Ansätze können grob in lineare Methoden (Korrelation, Regression) und nicht-lineare Methoden (Rekurrenzanalysen, Spektralanalysen) eingeteilt werden (Delaherche & Chetouani, 2010; Delaherche et al., 2012).

Zu den linearen Methoden zählen vorrangig gefensterte oder nicht gefensterte Korrelationen und Regressionen, die entweder zulassen, dass auch zeitversetzte Verhaltensweisen als synchron bezeichnet werden, oder nicht. Beispiele dafür sind die Cross Lagged-Correlation (CLC), Cross Lagged-Regression (CLR), Windowed-Cross-Correlation (WCC), Windowed-Cross-Lagged-Correlation (WCLC) und die Windowed-Cross-Lagged-Regression (WCLR). In diesen Bezeichnungen zeigt das *windowed* an, ob es sich um eine gefensterte oder nicht gefensterte Methode handelt. Das *lagged* nimmt Bezug darauf, ob ein Zeitversatz (*time lag*) zwischen den gezeigten Verhaltensweisen erlaubt ist, um noch von nonverbaler Synchronie zu sprechen. Alle Methoden basieren darauf, die Kovarianz zwischen zwei Zeitreihen zu analysieren. Unterschiede liegen darin, ob sie die Assoziationen für die komplette Zeitreihe oder in Segmenten (gefenstert), unter besonderer Kontrolle der Autokorrelation (Korrelation vs. Regression) oder mit Einschluss eines Zeitversatzes bestimmen. Die Windowed Cross-Correlation schließt beispielsweise keinen Zeitversatz ein und bestimmt lediglich das Ausmaß an genau zeitgleich realisierten Verhaltensweisen, wohingegen die Windowed Cross-Lagged Regression diesen Zeitversatz einschließt und gleichzeitig durch die Aufnahme des Vorzustandes des nonverbalen Merkmals der jeweiligen Person sogar in besonderem Maße für die Autokorrelation der Zeitreihe kontrolliert. Auf eine detailliertere Darstellung der linearen Methoden soll an dieser Stelle verzichtet werden, da sie in Publikation 1 umfassend beschrieben sind.

Die Cross-Recurrence Analyse (Rekurrenzanalyse) ist eine Methode, die keinen linearen Zusammenhang zwischen den beiden Zeitreihen annimmt. Funktionsprinzip dieser Methode ist nicht die Zeitreihen auf ihre Kovarianz zu untersuchen, sondern die *Co-visitation* zu untersuchen, d.h., wann beide Zeitreihen ähnliche Muster oder unähnliche Muster in ihren Phasenverläufen aufweisen. Verbildlicht wird dies durch Cross-Recurrence Plots. Auf der X-Achse dieser Plots befindet sich die Zeitachse (P1) der einen Zeitreihe des Probanden *i*, auf der Y-Achse die Zeitachse (P2) der anderen Zeitreihe des Probanden *j*. Ein Punkt wird dann

gesetzt, wenn ein Muster in Zeitreihe A, zu einem beliebigen Zeitpunkt in Zeitreihe B auftaucht. Dadurch werden gesamte Phasenverläufe der einen Zeitreihe auf ihre Wiederholung in der anderen Zeitreihe überprüft und umgekehrt. Aus den Plots können spezifische diagonalweise Recurrence Rate (RR) Profile ausgewertet werden. Beispielsweise charakterisiert die diagonale Linie, wo $i_{S1} = j_{S2}$ (i.e., die Gerade mit einem Anstieg = 1, die durch den Koordinatenursprung geht), Phasen, in denen sich beide Systeme genau gleichzeitig gleich verhalten. Diese Plots werden mittels der Cross-recurrence quantification analysis (CRQA) ausgewertet (Coco & Dale, 2014; Marwan & Kurths, 2002; Marwan, Romano, Thiel, & Kurths, 2007). Für einen spezifischen time lag (oder eine Spanne bis zu einem maximalen time lag) werden die Recurrence Raten ausgegeben, die als Synchronisationsmaß der beiden Zeitreihen verwendet werden (vgl. Richardson, Marsh, Isenhower, Goodman, & Schmidt, 2007; Shockley, 2005; Shockley et al., 2009; Shockley, Santana, & Fowler, 2003).

Eine andere non-lineare Methode, die besonders für die Auswertung zyklischer Daten (z.B. rhythmische Bewegungen, Herzraten-Daten) geeignet ist, ist die Spektralanalyse. Bei dieser Analysetechnik wird die Zeitreihe in ihre Frequenzkomponenten zerlegt. Dabei wird die Frequenz gleicher periodischer Prozesse bestimmt. Mit einer Cross-Spectral Analysis kann dann für jede Frequenzkomponente der Zusammenhang der beiden Zeitreihen bestimmt werden. Die durchschnittliche Kohärenz der Zeitreihen und die Stärke des Zusammenhangs bei der dominierenden Frequenz können berechnet werden, um nonverbale Synchronie zu quantifizieren (Chetouani et al., 2017; Schmidt, Morr, Fitzpatrick, & Richardson, 2012).

Validität der nonverbalen Synchronie

Die Validität bezeichnet die Gültigkeit eines Messinstrumentes und ist hoch, wenn das Messinstrument das abbildet, was es zu messen vorgibt. Es können mehrere Arten von Validität unterschieden werden. Grundsätzlich spricht man von Inhaltsvalidität, wenn das Messinstrument durch die enthaltenden Items oder die generelle Funktionsweise dazu imstande ist, das Konstrukt, was gemessen werden soll, zu messen. Die Kriteriumsvalidität und Konstruktvalidität sind zwei weitere Formen der Validität, die jedoch nicht so einfach nachzuweisen sind. Zur Kriteriumsvalidität zählen die prädiktive Validität und die konkurrente Validität. Beide Formen untersuchen, wie ein gemessenes Konstrukt mit einem vorher festgelegten Kriterium zusammenhängt. Dabei umfasst die konkurrente Validität den Zusammenhang zwischen dem gleichzeitig gemessenen Konstrukt und Kriterium. Bei der prädiktiven Validität wird ein Kriterium vorhergesagt, dass zu einer anderen Zeit gemessen wird (beispielsweise die Vorhersage später erreichter Schulnoten auf Grundlage von

Testwerten eines Fragebogens). Liegt kein universelles Kriterium vor, mit dem der Testwert verglichen werden kann, so ist es wichtig, die Konstruktvalidität zu untersuchen (Cronbach & Meehl, 1955). Dies kann erfolgen, indem viele Instrumente, die das gleiche messen sollen, miteinander verglichen werden (konvergente Validität) oder indem das zu validierende Instrument mit anderen Instrumenten, die nicht das Konstrukt messen, kontrastiert wird. Um ein hoch Konstrukt-valides Instrument zu haben, sollte es eine hohe konvergente Validität haben und auch hoch diskriminant valide sein, d.h. mit Messinstrumenten, die nicht das Konstrukt messen, möglichst wenig zusammenhängen (Campbell & Fiske, 1959).

Wie bereits dargestellt wurde, gibt es eine Vielzahl von Begriffen, die für interpersonelle Synchronie nonverbaler Verhaltensweisen verwendet werden. Außerdem gibt es zahlreiche Methoden, die für die Erfassung benutzt werden. Auffällig ist jedoch, dass systematische Untersuchungen der Validität bisher nur selten oder nur in Bezug auf einen bestimmten Aspekt durchgeführt wurden (Delaherche et al., 2012). Die fehlende Validierung der Methoden ist vor allem dann problematisch, wenn Untersuchungsergebnisse verschiedener Methoden miteinander verglichen werden.

Möchte man die Kriteriumsvalidität der Messinstrumente nonverbaler Synchronie untersuchen, so gibt es einige Kriterien gegen die Synchronie verglichen werden kann. Zum einen scheint es in der bisherigen Literatur sinnvoll, die Ergebnisse von Synchronisationsberechnungen gegen den Zufall abzusichern und eine Baseline-Kontrolle vorzunehmen. Weiterhin können Algorithmen an simulierten Zeitreihen, menschlich eingeschätzten Zeitreihen oder bestehenden Methoden, die das Konstrukt messen, validiert werden (konkurrente Validität). Die prädiktive Validität kann durch die Vorhersage bestimmter Outcomevariablen im Therapieprozess oder Entwicklungsfortschritten nachgewiesen werden. Gerade in Bezug auf nonverbale Synchronie fällt jedoch auf, dass es kein wahres und konsentiertes Konstrukt gibt, was als optimales Kriterium zur Validitätsprüfung verwendet werden kann. Deshalb empfiehlt es sich ebenfalls, Algorithmen mit ähnlichem Funktionsprinzip miteinander zu vergleichen, um eventuelle Unterschiede und Gemeinsamkeiten des durch die Algorithmen gemessenen Konstruktes auszumachen.

Statistische Absicherung gegenüber einer Baseline. In psychologischer Forschung werden Testscores oft auf ihre Bedeutsamkeit geprüft. Dabei wird der Wert gegen die Nullhypothese getestet. Es wird als Nullhypothese angenommen, dass sich der Testwert nicht von null unterscheidet. Diesem Hypothesentest liegt immer eine Verteilung einer Prüfgröße zugrunde (meist *t*-verteilt). In Bezug auf die Zeitreihen-basierte Bestimmung nonverbaler

Synchronie, wird meist eine Absicherung der Ergebnisse gegenüber Zufallssynchronie empfohlen. Nehmen wir zwei Zeitreihen, die Bewegungen von Personen widerspiegeln, so kann es sein, dass durch rein zufällig gleichzeitige Bewegungen falsche Synchronie geschaffen wird, die wir als echte nonverbale Synchronie werten. Dies passiert, da Menschen sich, sei es durch die Atmung oder Gesten, stets bewegen. Durch einen Standard-Hypothesen-Test wäre es demnach nicht möglich zu testen, ob die echte (genuine) Synchronie größer ist, als zufällig entstandene Zufallssynchronie (Moulder, Boker, Ramseyer, & Tschacher, 2018). Die Zufallssynchronie wurde von Bernieri et al. (1988) unter dem Begriff der Pseudosynchronie eingeführt. Er schlägt vor, die genuine Synchronie um die Pseudosynchronie, die durch internes oder externes Mischen der Zeitreihen erhalten wird, zu bereinigen. Internales Mischen (auch *within-shuffling*) bedeutet, dass eine Zeitreihe (oder eine Videosequenz) in verschiedene Stücken zerlegt wird und diese zufällig zu einer neuen (Surrogat-)Zeitreihe angeordnet werden. Pseudosynchronie beschreibt die Synchronie zwischen der Zeitreihe A und der neu geordneten Zeitreihe B (Bernieri et al., 1988). Durch die Vermischung wird die Zeitstruktur und Abhängigkeitsstruktur des Zeitreihenpaares zerstört, sodass die Subtraktion von genuiner und Pseudosynchronie dazu führt, nur noch die nicht zufällig entstandenen Anteile der Synchronie zu messen. Die Vermischung wird durch Bootstrapping realisiert. Häufig durchgeführtes Bootstrapping führt zu einem Surrogat-Datensatz künstlich erzeugter Interaktionen, der dann für die Bestimmung einer empirischen Prüfgröße verwendet wird (Ramseyer & Tschacher, 2010). Die Analyse von Surrogat-Datensätzen stellt eine valide Möglichkeit dar, Probleme beim Standard-Hypothesen-Testen zu umgehen (Theiler, Eubank, Longtin, Galdrikian, & Farmer, 1992). Eine weitere Möglichkeit der Vermischung ist die externe Mischung (auch *between-shuffling*), bei der Zeitreihen (oder Videos) von Interaktionen miteinander verbunden werden, die nicht miteinander kombiniert waren. Beispielsweise würde die Zeitreihe A aus dem Zeitreihenpaar AB mit der Zeitreihe D aus dem Zeitreihenpaar CD kombiniert werden. Die zwischen dem artifiziellen Zeitreihenpaar (Surrogat-Zeitreihenpaar) ermittelte Pseudosynchronie gilt dann als Baseline, für die die genuine Synchronie kontrolliert wird. Das Funktionsprinzip wurde hier auf Grundlage eines Zeitreihenpaares beschrieben. Die Überprüfung, ob genuine Synchronie größer als Pseudosynchronie ist, erfolgt jedoch nicht auf Basis der individuellen Werte, sondern auf Grundlage der Verteilung der durch die Bootstraps ermittelten Parameter. Durch die wiederholte Durchführung entsteht eine empirische Prüfverteilung, die benutzt werden kann, um die Nullhypothese keiner bedeutsamen Synchronie zu testen. Es lässt sich schlussfolgern,

dass die Referenz für die Hypothesentestung für die bisher beschriebenen Surrogat-Ansätze, die empirisch ermittelte Zufälligkeit ist. Ein anderer Begriff für Pseudosynchronie ist Surrogat-Synchronie (SuSy), da die Prüfung auf Überzufälligkeit durch künstlich hergestellte Surrogat-Datensätze ermöglicht wird.

Durch die Entscheidung, welcher Surrogat-Ansatz verwendet wird, wird ebenfalls festgelegt, gegen welche Nullhypothese die genuine Synchronie getestet wird. Beispielsweise wird durch die Verwendung von within-shuffling die zeitliche Struktur der Interaktion zerstört, sodass geprüft werden kann, ob die genuine Synchronie höher ist als zufällig gezeigte Bewegungssynchronie. Geht man davon aus, dass die Interaktionen einem bestimmten Muster folgen und möchte man gegen eine restriktivere Baseline testen, so empfiehlt sich das between-shuffling. Dabei bleibt die zeitliche Struktur der einzelnen Interaktion gewahrt.

Die so gesehen liberalste Methode der Surrogat-Erstellung ist das *data shuffling*. Dabei werden einzelne Punkte der Zeitreihe ohne Zurücklegen permutiert, bis eine komplett neue Zeitreihe entstanden ist. Wie Moulder et al. (2018) zeigten, ist es für viele Datentypen möglich, diese Nullhypothese zu verwerfen, da eine zeitliche Struktur zwischen den Zeitreihen eines Zeitreihenpaares besteht (z.B. bei Motion Capture Daten, Sinus Kurven und Motion Energy Werten). Für künstlich erstelltes Rauschen konnte die Nullhypothese jedoch nicht verworfen werden, was für die Validität der Überprüfungsstrategie spricht, da Rauschen keinerlei zeitliche Struktur aufweist. Innerhalb der psychologischen Forschung wurde diese Methode beispielsweise von Louwerse, Dale, Bard, und Jeuniaux (2012) verwendet. Oft resultiert data shuffling aber in Zeitreihen, die physiologisch gar nicht realisierbar sind (z.B. sehr hohe Sprünge von Bewegungszeitreihen innerhalb sehr geringer Zeit).

Eine weitere Methode zur Erstellung eines Surrogat-Datensatzes ist das *data sliding*. Dabei wird die Zeitreihe nach ca. 60% geteilt und das Endstück der Zeitreihe wird an den Anfang verschoben. Dadurch wird die zeitliche Struktur der Kompletzeitreihe zerstört, der Mittelwert und die Varianz bleiben jedoch erhalten. Die dabei getestete Nullhypothese scheint die restriktivste Nullhypothese zu sein (Moulder et al., 2018).

Für die Wahl der passenden Strategie zur Erstellung der Surrogat-Synchronie empfehlen Moulder et al. (2018) mit dem data shuffling zu beginnen und dann abhängig von der zu testenden Hypothese zu wählen, ob within-shuffling (z.B. wenn der Untersuchungsgegenstand bestimmte Segmente der Interaktion sind), ob between-shuffling (z.B. wenn es einen klaren Experimentalablauf gibt, für den ebenfalls kontrolliert werden soll) oder ob data sliding verwendet werden soll (z.B. wenn das Maß an globaler Synchronie einer

Interaktion untersucht werden soll). Weiterhin wird ersichtlich, dass die Güte jeder Baseline-Methode abhängig vom Datentyp und den in der WCLC verwendeten Einstellungen ist. Beispielsweise ist die genuine Synchronie in Referenz zu einem data sliding-Surrogat-Datensatz für eine WCLC auf Grundlage von Motion Energy Zeitreihen mit einer Bandbreite von 30 Sekunden höher, für eine WCLC mit 2 Sekunden Bandbreite jedoch statistisch nicht unterschiedlich (Moulder et al., 2018).

Die Absicherung gegenüber der empirisch ermittelten Zufälligkeit ist nur eine Form, genuine Synchronie gegen Pseudosynchronie abzusichern. Geht man davon aus, dass die Prüfparameter einer bereits bekannten Prüfverteilung (z.B. t -Verteilung) folgen, so kann an die Stelle der empirischen Ermittlung einer Verteilung ebenfalls eine bereits etablierte Verteilung treten. Altmann (2013) sowie Gottman und Ringland (1981) beschreiben einen solchen parametrischen Test als Absicherung gegen Zufälligkeit der Ergebnisse. Dabei wird der in der Zeitreihenanalyse erhaltene Korrelationskoeffizient (bzw. R^2) anhand einer existierenden Prüfverteilung auf seine Verschiedenheit von null getestet (Altmann, 2013). Folgt man der Idee der Alpha-Fehler-Kumulierung, kann das verwendete Signifikanzniveau dabei ebenfalls angepasst werden. Eine weitere statistische Absicherung kann durch ein *prewhitening* der Zeitreihe in Kombination mit der Anpassung des Signifikanzniveaus für bedeutsame Kreuzkorrelationen geschaffen werden (Dean & Dunsmuir, 2016).

Innerhalb bestimmter Methoden wird direkt für zufällig auftretende Korrelationen kontrolliert, indem der Vorzustand der Person als autoregressives Modell mit in die Analyse eingeht (Gottman & Ringland, 1981). An einem konkreten Beispiel bedeutet das, dass bei der Vorhersage der Bewegungen des Interaktionspartners zum Zeitpunkt t nicht nur die Bewegungen der Person zu t (Jetzt-Zustand), sondern ebenfalls zu $t-1$ (Vorzustand) als Prädiktoren verwendet werden. Methoden, die dieses Prinzip nutzen, sind beispielsweise die CLR und WCLR. Im Gegensatz zu den empirischen Methoden werden bei der Anwendung dieser Methoden nur solche autoregressive Zusammenhänge kontrolliert, die den gleichen time lag wie die kreuz-korrelativen Zusammenhänge haben.

Kriteriumsvalidität. Es gibt wenige Studien, die Ergebnisse von Synchronisationsalgorithmen gegen andere Kriterien absichern. Durchgeführte Studien verglichen die Ergebnisse von Algorithmen mit a) menschlichen Ratings b) Simulationsstudien und c) generalisierteren Methoden, die zur Erfassung des gleichen Merkmals verwendet wurden. Eine einzige Studie erbrachte einen indirekten Hinweis auf die Kriteriumsvalidität der Algorithmen an menschlichen Ratings. Paxton und Dale (2013b) ließen eine Interaktion von

menschlichen Ratern und der Motion Energy Analysis hinsichtlich der gezeigten Bewegungen kodieren. Nach Anwendung der CLC konnte ein ähnliches Level der Bewegungssynchronie gezeigt werden ($r = .68$). Obwohl diese Studie oft als Referenz genannt wird, sollte sie doch eher als eine Validierung der Motion Energy Analysis betrachtet werden. Da Studien mit menschlich gerateten Synchronisationsphänomenen fehlen, war dies unter anderem Untersuchungsgegenstand der zweiten Publikation.

Auch für die Prüfung von sozialwissenschaftlich eingesetzten zeitreihenanalytischen Algorithmen an simuliertem Material fehlen bisher Studien. Es gibt lediglich eine Studie, die an simulierten Zeitreihen die WCLC und WCLR mit Peak-Picking Algorithmus testete (Altmann, 2013). Diese Studie zeigte, dass die WCLR in Verbindung mit dem Peak-Picking Algorithmus der WCLC in Bezug auf zyklische Daten überlegen ist (Altmann, 2013). Für simulierte Interaktionsverläufe fehlt eine Validierung bisher und ist deshalb ebenso Forschungsgegenstand der zweiten Publikation.

Nicht-lineare Methoden wie beispielsweise die Cross-Recurrence Analysis können als eine Generalisierung linearer Modelle verstanden werden (Coco & Dale, 2014). Die Übereinstimmung linearer Methoden mit nicht-linearen Methoden wurde bisher nur unzureichend untersucht. Wiltshire (2015) zeigte keine Assoziation zwischen linearen und nicht-linearen Methoden (Cross-Recurrence Analysis). Jedoch werden unter dem Stichwort lineare Methoden eine Vielzahl an Methoden mit unterschiedlicher Funktionsweise zusammengefasst. Als den Algorithmen gemein wird angenommen, dass alle das Konstrukt der nonverbalen Synchronie abbilden. Demnach werden auch Ergebnisse von Methoden miteinander verglichen, die einen time lag einschließen oder nicht. Innerhalb der ersten Publikation wurden deshalb verschiedene lineare Methoden in teilweise verschiedenen Versionen (CLC, CLR, WCC, WCLC_{S1}, WCLC_{S2}, WCLC_F, WCLR) mit den Ergebnissen der CRQA kontrastiert.

Konvergente Validität. In der psychologischen Interaktionsforschung werden vermehrt lineare Methoden angewendet. Bisher wurde jedoch noch nicht untersucht, ob die linearen Methoden, wendet man sie auf den gleichen Datensatz an, zu den gleichen Synchronie-Indizes führen. Bisherige Studienergebnisse weisen eher in entgegengesetzte Richtung: Nach Altmann (2011) (Verwendung WCLR) und Paxton und Dale (2013a) (Verwendung CLC) ist Synchronie in Konfliktsituationen geringer als in normalen Small-Talk-Situationen. Dies konnte unter Verwendung der WCLC nicht eindeutig gezeigt werden (Tschacher et al., 2014). Natürlich gibt es auch noch andere Unterschiede zwischen den angeführten Studien wie zum

Beispiel die Untersuchung unterschiedlicher Modalitäten (Bewegungen vs. Gesichtsausdrücke), unterschiedliche Stichproben und unterschiedliche Studiensettings. Jedoch wurde die Verwendung verschiedener Methoden noch nicht als Erklärung für die gemischten Ergebnisse untersucht. Mit dieser zentralen Forschungsfrage beschäftigt sich die erste Publikation.

Maße nonverbaler Synchronie

Für die Betrachtung der Maße von Synchronie, ist es zuerst nötig, die den Maßen zugrundeliegenden Annahmen zu formulieren. In der bisherigen Forschung existieren zwei sich gegenüberstehende Überzeugungen zu Synchronie. Die Synchronie-Berechnung von Ramseyer und Tschacher (2011, 2014) sowie Paulick et al. (2017) basiert darauf, dass angenommen wird, Synchronie sei ein entstehender Prozess, der eine gewisse Zeitspanne benötigt, um sich überhaupt zu etablieren (z.B. mindestens 30 Sekunden bis 1 Minute). Dem gegenüber steht die unter anderem von Altmann (2011, 2013) methodisch operationalisierte Überlegung, dass Synchronie in festgelegten Synchronisationsintervallen stattfindet, die sich mit Nicht-Synchronisationsintervallen innerhalb einer Interaktion abwechseln. Konform zu dieser Operationalisierung sprechen auch Boker und Rotondo (2002); Dahan, Noy, Hart, Mayo, und Alon (2016); Tronick, Als, und Brazelton (1977) von Phasen der Synchronie und Phasen des *Exits* oder des *Breaking* der Synchronie.

Wendet man die beschriebenen Algorithmen an, so resultiert stets eine Korrelationsmatrix, die auf bestimmte Art weiterverarbeitet wird. Ramseyer und Tschacher (2011) sowie Paulick et al. (2017) benutzen diese Matrix, um über die Summe der Absolutwerte der Fisher-transformierten Korrelationskoeffizienten das globale Ausmaß von Synchronie in der Interaktion festzustellen. Altmann (2011, 2013) quadriert die entstandenen Korrelationen, um 1) positive Werte zu erhalten und 2) höheren Korrelationen ein höheres Gewicht zu geben als geringen Korrelationen, die durch das Quadrieren noch geringer werden. Anschließend wird ein Peak-Picking Algorithmus benutzt, um diese Matrix nach Synchronisationsintervallen und Nicht-Synchronisationsintervallen zu filtern. Die Häufigkeit der Synchronie in Referenz zur Gesamtdauer der Interaktion wird als Synchronie-Maß empfohlen. Neben diesen beiden Maßen gibt es eine Vielzahl anderer Maße, beispielsweise kann ebenfalls der Maximalwert der Korrelationsmatrix als Synchronie-Score benutzt werden. Diese Maße werden momentan eingesetzt, ohne sie vorher systematisch verglichen bzw. miteinander in Verbindung gesetzt zu haben. Dieser Forschungsfrage nimmt sich Publikation 1 an.

Neben den Maßen für den Gesamtanteil der Synchronie einer Interaktion gibt es ebenfalls Maße, die es ermöglichen, Synchronie differenzierter zu betrachten. Die komplexeren linearen Methoden (WCLC; WCLR) ermöglichen eine Differenzierung in die Synchronie-Anteile, die durch den Patienten bzw. den Therapeuten geleitet wurden. Es können demnach detailliertere Angaben dazu gemacht werden, wer die Interaktion führt und wer folgt. Ramseyer und Tschacher (2011) zeigten bereits, dass diese Anteile mit verschiedenen Outcomes zusammenhängen. Eine weitere bedeutsame Größe ist der Zeitversatz mit dem Synchronie realisiert wird (= time lag). Das Gesamtmaß der Synchronie ist zwar besser untersucht, aber einige Studien weisen darauf hin, dass eine Betrachtung der Zeit bis zur Synchronie-Etablierung ebenfalls aufschlussreich sein kann (Feldman, Greenbaum, & Yirmiya, 1999; Kleinbub, 2017). Ein ebenfalls aus dem time lag generiertes Maß ist das von Tschacher, Ramseyer, und Koole (2017) beschriebene Maß des Social Present – der sozialen Präsenzzeit.

4 Datengrundlage & Inhalte der einzelnen Publikationen

Die in dieser Arbeit zusammengefassten Publikationen entstanden innerhalb des Projektes „Timing nonverbaler Patient-Therapeut-Interaktionen und Therapieerfolg bei sozialen Phobien (TIMPATHIN)“. Bei dem Projekt handelt es sich um eine durch die DFG geförderte multizentrische Videostudie, die einerseits eine Sekundärdatenanalyse des Projektes SOPHO-NET darstellt und andererseits die Stichprobe durch Therapien der Institutsambulanz der Universität Trier, Abteilung für Klinische Psychologie und Psychotherapie erweiterte. Innerhalb des Ursprungsprojektes SOPHO-NET wurden 495 Patienten mit Sozialer Angststörung über einen Zeitraum von drei Jahren (2007-2009) mit entweder manualisierter kognitiver Verhaltenstherapie (KVT) oder psychodynamisch-orientierter (PDT) Kurzzeittherapie behandelt. Beide Behandlungskonzepte umfassten einen Sitzungsumfang von 30 Stunden (inklusive 5 Probatorik-Sitzungen), die zum Großteil zu je 50 Minuten auf einer wöchentlichen Basis durchgeführt wurden.

Innerhalb der kognitiven Verhaltenstherapie wurde sich am Manual von Stangier, Clark, und Ehlers (2006) orientiert. Das Vorgehen innerhalb der kognitiven Verhaltenstherapie ist stark an dem in Abschnitt 3.1.1 beschriebenen kognitiven Störungsmodell der Sozialen Angststörung orientiert. In der ersten Phase der Therapie wird das individuelle Modell der Sozialen Angststörung mit dem Patienten abgeleitet. Dafür werden detaillierte Situationsanalysen durchgeführt. In der zweiten Phase wird eine kognitive Vorbereitung auf die Exposition *in vivo* vorgenommen. Dabei werden Situationen gesammelt, die Angstgefühle

auslösen und nach dem Schwierigkeitsgrad sowie der Herstellbarkeit Rollenspiele vorbereitet. In der dritten Phase erfolgen die eigentliche Exposition und die Verhaltensexperimente. In der nächsten Phase werden negative Kognitionen im Zusammenhang mit angstauslösenden Situationen auch auf verbaler Ebene überprüft. Bei dysfunktionalen Kognitionen findet anschließend mittels kognitiver Therapiestrategien eine kognitive Umstrukturierung statt. In der letzten Phase der Therapie wird sich dem Therapieabschluss und der Rückfallprophylaxe gewidmet.

Die psychodynamisch-orientierte Therapie basierte auf der supportiv-expressiven Therapie von Luborsky (1995). Dabei wurde der Fokus für die verschiedenen Therapiephasen genau vorgegeben. Innerhalb der probatorischen Sitzungen (Sitzung 1-5) wurden Beziehungs-Episoden-Interviews durchgeführt, um das der Störung zugrundeliegende Zentrale Beziehungskonfliktthema (ZBKT) herauszuarbeiten. Die Anfangsphase der Therapie (Sitzungen 6-13) soll dazu genutzt werden, durch supportive Interventionen eine tragfähige therapeutische Allianz mit dem Patienten zu etablieren. Die Mittlere Phase der Therapie (Sitzung 14-21) wird zur Verfeinerung des ZBKT genutzt. Dieser wird durch das Beziehen auf verschiedene interpersonelle Beziehungen des Patienten durchgearbeitet. Während dieser Phase werden außerdem selbstgeleitete Symptomexpositionen durchgeführt. In der folgenden Phase der Therapie (Sitzung 22-27) wird das Ende der Therapie benannt und vorbereitet (z.B. Benennung Meilensteile, Rückfälle werden auf das ZBKT bezogen). In der letzten Phase der Therapie (Sitzung 28-30) werden die Verbesserungen hinsichtlich der Symptomatik thematisiert und verstärkt. Der Patient wird zur selbstgeleiteten Exposition ermutigt. Die Eigenleistung und Eigenarbeit des Patienten werden betont.

Innerhalb der Institutsambulanz in Trier wurden Patienten im naturalistischen Setting mit entweder Kurzzeit- oder Langzeittherapien kognitiver Verhaltenstherapien behandelt. Die Therapiestunden umfassten auch hier 50 Minuten und wurden wöchentlich durchgeführt. Das Therapiekonzept orientierte sich an einem integrativen Ansatz der kognitiven Verhaltenstherapie (Lutz, Schiefele, Wucherpfennig, Rubel, & Stulz, 2016), der Psychoedukation, Entspannungstrainings, Exposition sowie auch interpersonelle und emotionsfokussierte Elemente umfasste. Die Therapeuten waren auch in diesem Zentrum mit den störungsspezifischen Manualen der Behandlung vertraut, konnten die Ansätze jedoch beliebig abwandeln, sodass sie möglichst gut auf den Patienten passten.

Einschlusskriterien waren das Vorliegen einer Sozialen Angststörung als Primär- oder Sekundärdiagnose sowie das Vorhandensein einer für die Bewegungsanalyse verwertbaren

Therapiesitzung auf Video. Patienten mit akuter Psychose oder Substanzabhängigkeiten sowie anderen hirnganischen Störungen als auch Patienten mit nicht brauchbaren Videoaufzeichnungen wurden ausgeschlossen (Details siehe Publikationen).

Um die Brauchbarkeit der SOPHO-NET Videos einzuschätzen, wurde zuerst eine manuelle Sichtung der vorhandenen Videos durchgeführt. Dafür wurden 5951 Videos, die 335 Patienten zuzuordnen sind (KVT = 223, PDT = 112), gesichtet. Eine Vielzahl von Videos musste nach der ersten Sichtung ausgeschlossen werden, wie in Abbildung 5 dargestellt ist.

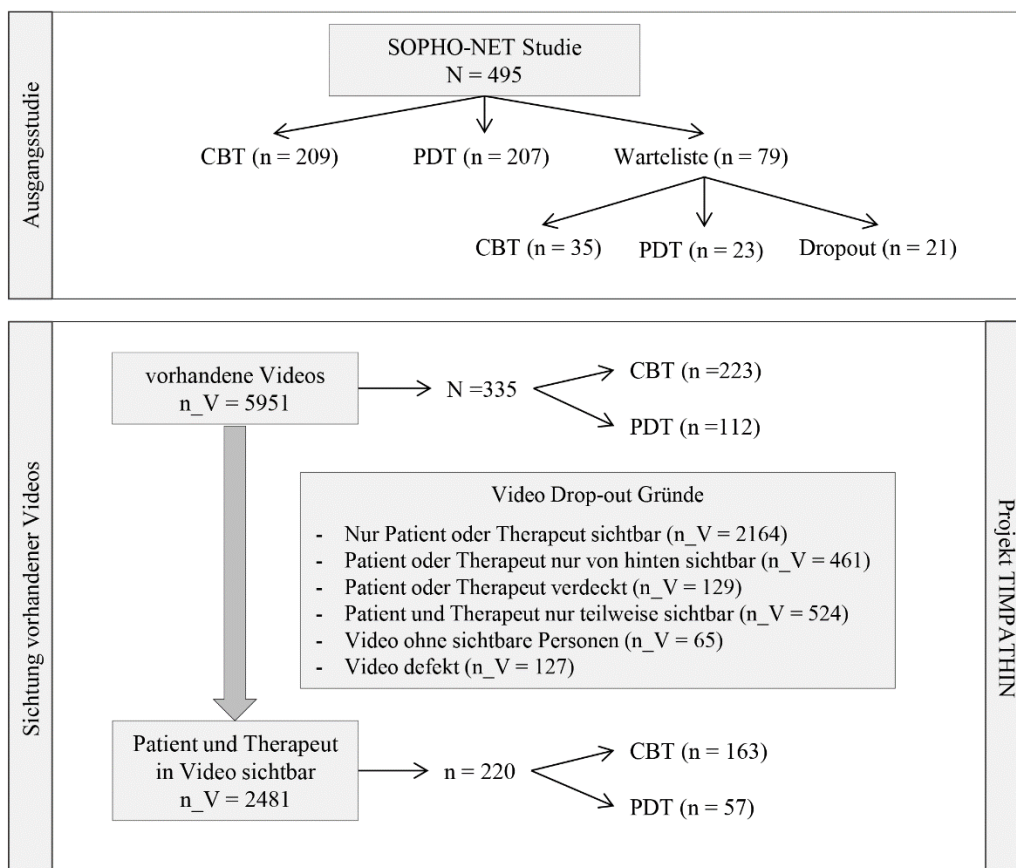


Abbildung 5 Flowchart über Studienfluss der SOPHO-NET Videos (erste manuelle Sichtung),
 n = Stichprobenanzahl Patienten, n_V = Anzahl Videos

Für die multizentrische Videostudie TIMPATHIN konnten somit $n = 220$ Patienten aus der SOPHO-NET Studie eingeschlossen werden. Da nicht von allen Sitzungen Videoaufnahmen vorlagen, wurde sich auf vier Messzeitpunkte beschränkt, die proximal zu den in SOPHO-NET verwendeten Erhebungszeitpunkten lagen. Diese Zeitpunkte waren im Einzelnen: die 3. probatorische Sitzung (Sitzung 3), die 3. antragspflichtige Sitzung (Sitzung 8), die 15. antragspflichtige Sitzung (Sitzung 20) und die 24. antragspflichtige Sitzung (Sitzung 29). Das Design der Studie ist in Abbildung 6 veranschaulicht.

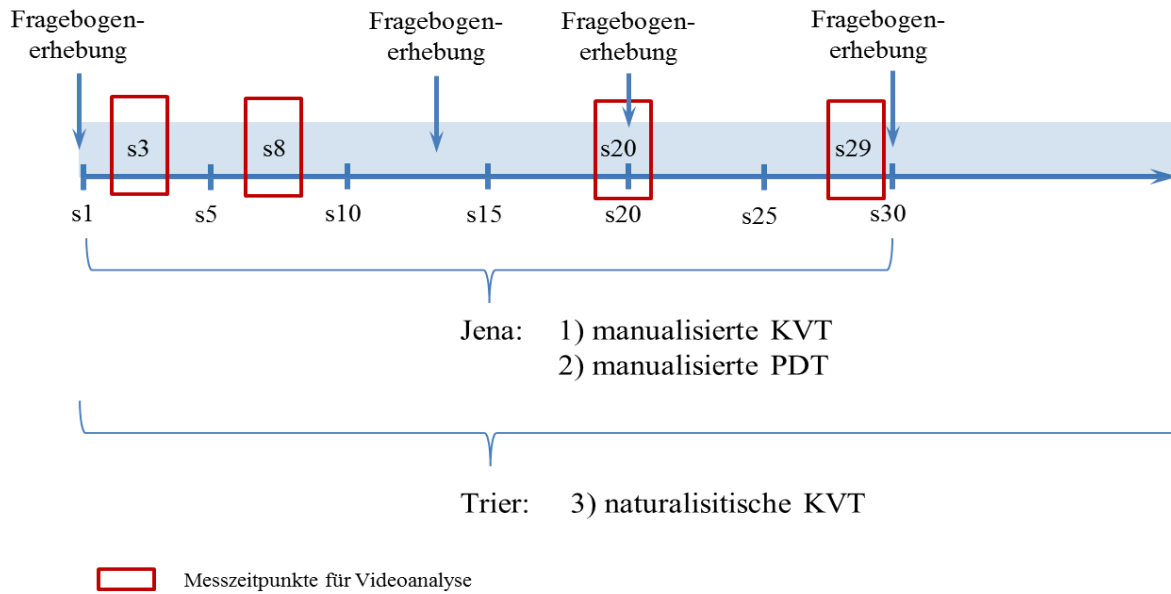


Abbildung 6 Design und Ablauf der Studie TIMPATHIN, $sX = X$. Sitzung

Um die Qualität der Videosequenzen weiterhin hinsichtlich der Eignung für die Motion Energy Analysis, der Kameraposition und hinsichtlich schwerer Bild- und Kodierfehler zu untersuchen, wurden weitere drei Sichtungen durchgeführt (siehe Tabelle 7). Nach Ausschluss der nicht geeigneten Sitzungen, wurden 156 Patienten (Behandlung mit KVT: $n = 102$, Behandlung mit PDT: $n = 54$) aus der SOPHO-NET Studie eingeschlossen.

Tabelle 7 Anzahl vorhandener Videos aus SOPHO-NET, die für TIMPATHIN eingeschlossen wurden sowie Ausschlusskriterien der durchgeführten Sitzungen

| | S3 | S8 | S20 | S29 |
|--|-----|-----|-----|-----|
| Sichtung 1 – Video vorhanden und beide gut sichtbar | 176 | 177 | 162 | 133 |
| Sichtung 2 – Eignung für MEA | | | | |
| Ausschluss weil: | | | | |
| - T steht auf | 11 | 16 | 3 | 15 |
| - Beantwortung Fragebogen nach Minute 10 | 8 | 3 | 1 | 5 |
| - Rollenspiel | 2 | 4 | 1 | 0 |
| - 3. Person im Raum | 1 | 1 | 2 | 0 |
| - T oder P außerhalb Video | 52 | 38 | 46 | 33 |
| - Familienaufstellung | 1 | 0 | 0 | 0 |
| - Lichtverhältnisse schlecht | 0 | 0 | 2 | 0 |
| Nach Sichtung 2 verbleibende Videos | 101 | 115 | 107 | 81 |

| Sichtung 3 – Position | | | | |
|-------------------------------------|-----------|------------|-----------|-----------|
| Ausschluss weil: | | | | |
| - P oder T eher von hinten zu sehen | 17 | 13 | 17 | 7 |
| Nach Sichtung 3 verbleibende Videos | 84 | 102 | 90 | 74 |
| Sichtung 4 – Kodierfehler | | | | |
| Ausschluss weil: | | | | |
| - Zu viele Kodierfehler | 0 | 2 | 0 | 1 |
| Finale Stichprobe TIMPATHIN | 84 | 100 | 90 | 73 |

Anmerkungen: SX = X. Sitzung, MEA = Motion Energy Analysis, P = Patient, T = Therapeut

Für die Institutsambulanz Trier wurden Videos von 134 Patienten mit Sozialer Angststörung gesichtet. Davon mussten $n = 23$ wegen unzureichender Videoqualität, der Kameraposition oder des Verlassens des Sitzplatzes ausgeschlossen werden, sodass $n = 111$ Patienten für die Studie TIMPATHIN eingeschlossen werden konnten (Sitzung 3: $n = 97$, Sitzung 8: $n = 79$, Sitzung 20: $n = 70$ und Sitzung 29: $n = 54$). Da in bisherigen Studien ein mittlerer Zusammenhang zwischen Synchronie und dem Therapieergebnis gefunden wurde ($r = .35$, Ramseyer und Tschacher, 2011), wurde durch G-Power die notwendige Stichprobengröße bei einer Power von $(1 - \beta = .80)$ und einem α -Fehler-Niveau von $\alpha = .01$ auf 78 Patienten geschätzt. Die erreichte Stichprobengröße von 256 Patienten ist damit als ausreichend zu betrachten.

Um die Videos in einem einheitlichen Format der Bewegungsanalyse mittels Motion Energy Analysis zuführen zu können, wurde alle Videos in ein einheitliches Format konvertiert (Details siehe Publikationen).

4.1 Erfassung von Körperbewegungen mittels Motion Energy Analysis

Die Motion Energy Analysis (MEA) wurde bereits von Watanabe (1983) zur automatisierten Erfassung von Bewegungen verwendet. Das Stimulusmaterial für die Analyse stellen Videosequenzen dar. Videosequenzen sind dadurch gekennzeichnet, dass sie aus einzelnen Bildern (Frames) bestehen, die mit einer so hohen Rate gezeigt werden, dass für den Beobachter der Eindruck bewegter Bilder entsteht. Innerhalb der MEA werden benachbarte Videobilder auf die sich ändernden Pixel untersucht. Dafür ist es zuerst nötig, die einzelnen Videobilder in Graustufen-Bilder zu konvertieren. Anschließend werden die Pixel zweier benachbarter Videobilder auf eine Intensitätsänderung untersucht. Liegt eine Intensitätsänderung vor, die einen gewissen Schwellenwert (Intensitäts-Schwelle μ) übersteigt,

so wird der Pixel als ein veränderter Pixel gezählt. Die Summe der Pixel, die sich von Videobild t bis zu Videobild $t+1$ verändern, operationalisieren die Bewegung, die zwischen diesen beiden Bildern gezeigt wird. Wendet man dieses Verfahren auf die komplette Videosequenz an, so resultiert eine Zeitreihe, die die Bewegungen innerhalb eines Videos abbildet.

Um das Verfahren besser für die Interaktionsforschung nutzen zu können, ist es nötig, zwischen verschiedenen Bewegungsanteilen innerhalb des Videos zu unterscheiden, d.h. Bewegungen vom Patienten sollten von Bewegungen des Therapeuten unterscheidbar sein, auch wenn sie mit einer Kamera aufgenommen wurden. Dafür können sogenannte *Regions of Interest* (ROIs) bestimmt werden. Markiert man den Patienten mit einer ROI und den Therapeuten mit einer anderen, sich nicht mit der ersten überschneidenden ROI, resultieren als Output zwei Zeitreihen: eine für die Bewegungen der Patienten-ROI und eine für die Bewegungen der Therapeuten-ROI. Die einzelnen Schritte vom Video zur Zeitreihe sind in Abbildung 7 veranschaulicht. Damit die Zeitreihen tatsächlich sinnhafte Bewegungen enthalten, müssen einige technische und inhaltliche Voraussetzungen erfüllt sein, die in Tabelle 8 zusammengefasst sind.

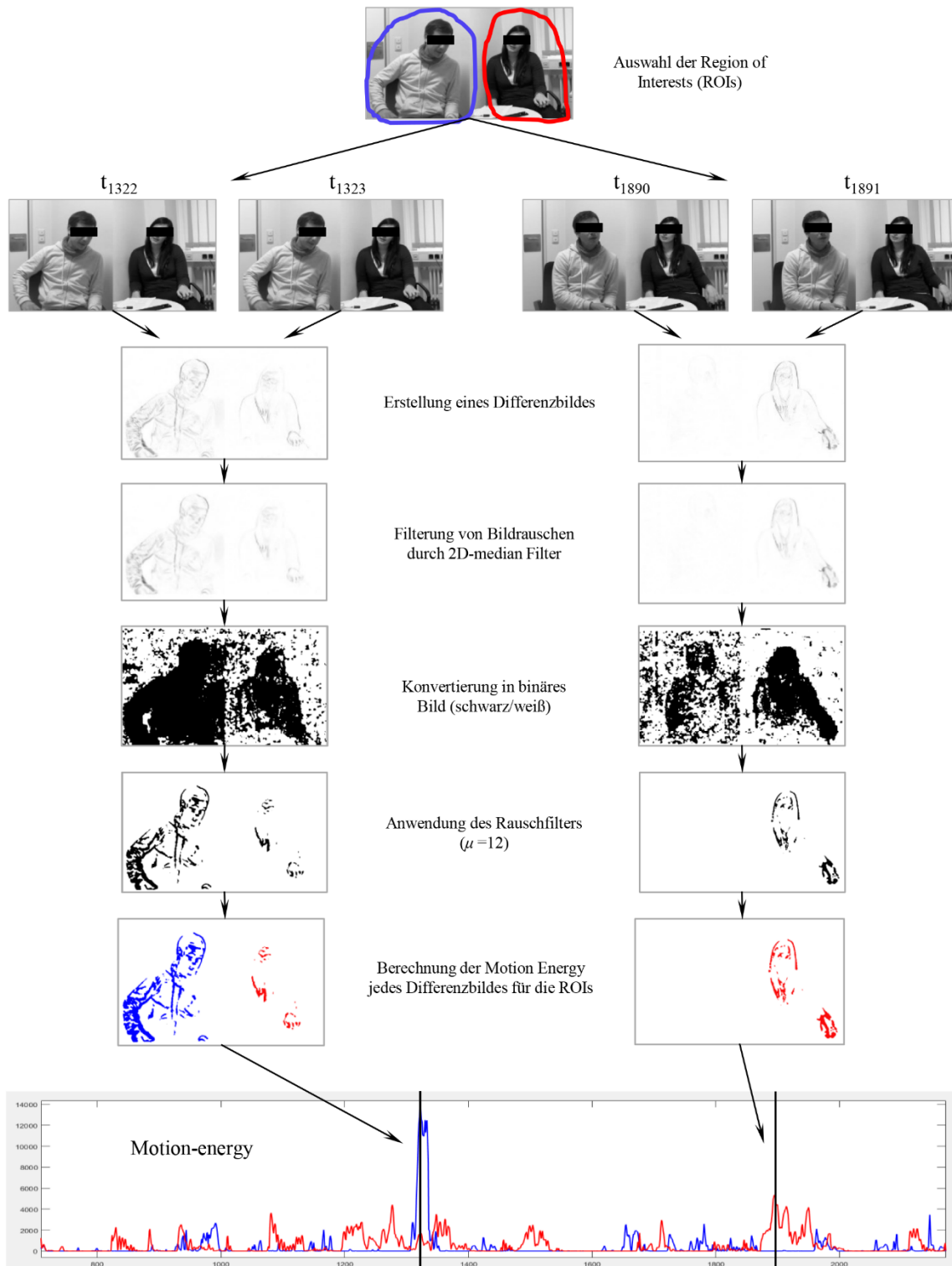


Abbildung 7 Ablaufschema für die Erfassung von Bewegungen aus einem Video mittels Motion Energy Analysis (MEA)

Tabelle 8 *Technische und inhaltliche Voraussetzungen der Motion Energy Analysis*

| Technische Voraussetzungen | Inhaltliche Voraussetzungen |
|--|---|
| Patient und Therapeut im Video nicht überlappend zu sehen | Konstantes Sitzen, ohne Aufstehen |
| Stabile und gute Lichtverhältnisse | Kein Besprechen von Fragebögen bzw. Ausfüllen von Formblättern |
| Keine Bewegung der Kamera, kein Auto-Zoom, keine Verzerrungen | Andere organisatorische Dinge wie Ausmachen eines erneuten Termins ausschließen |
| Stabiler Hintergrund (keine Fenster mit Vorbeifahrenden Autos etc.) | Keine Rollenspiele oder Entspannungstechniken, nur verbale Interaktionen |
| Dopplung der Bewegungen vermeiden (keine Spiegelungen, große Schattenbildung) | Festgelegte Bereiche für Interaktionspartner (kein Plätze-Tauschen) |
| Person frontal bis seitlich im Bild | |
| Größe der Interaktionspartner im Video vergleichbar | |
| Kleidung der Personen beachten (nicht zu ähnlich zum Hintergrund; dunkle, einfarbige Kleidung führt zu weniger Kontrast innerhalb des Kleidungsstücks vgl. Abb. 7; gemusterte Kleidung führt zu vielen Pixelveränderungen) | |

Anmerkung: Tabelle nach Altmann (2013), Ramseyer und Tschacher (2008) mit inhaltlichen Ergänzungen

Auch der Kamerawinkel spielt eine wichtige Rolle in der Aufzeichnung von Videosequenzen. In einer Vorstudie zu den hier beschriebenen Originalarbeiten wurde der Einfluss der Kameraposition untersucht. Als Stimulusmaterial wurden Videos verwendet, die eine Winkekatze aus verschiedenen Positionen zeigten. Die Verwendung einer Winkekatze ermöglichte es, mit ein und derselben Kamera aus verschiedenen Positionen ein und dieselbe Bewegung aufzuzeichnen. Es konnte gezeigt werden, dass Positionen vergleichbar sind, die die Interaktionspartner lateral oder seitlich zeigten (siehe Abb. 8).

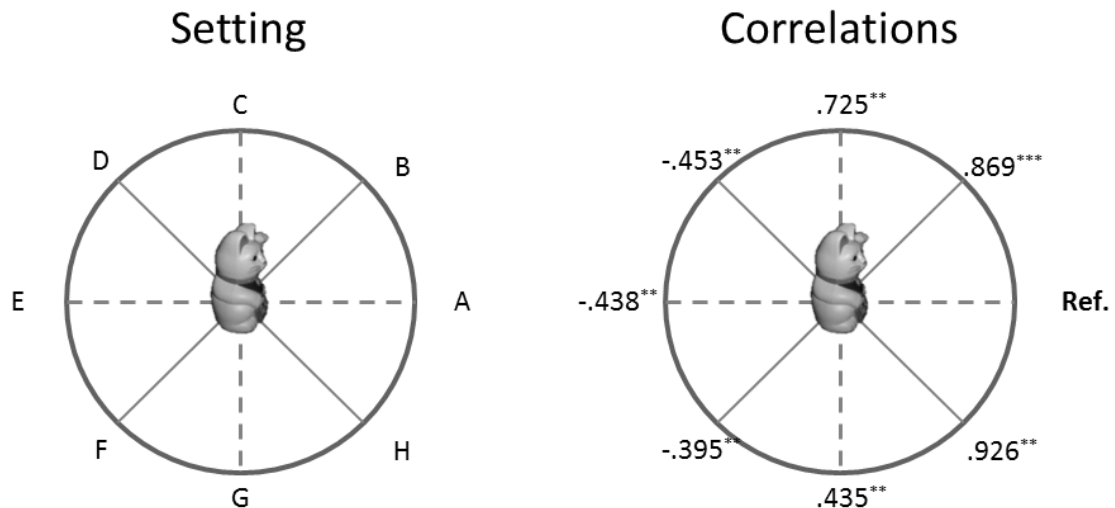


Abbildung 8 Korrelationen von Bewegungen aus verschiedenen horizontalen Kamerapositionen (Abb. aus Altmann et al. (2019))

Für die Anwendung der MEA existieren aktuell mehrere Möglichkeiten: Mit einem Apple Betriebssystem kann die Software MEA von Ramseyer und Tschacher verwendet werden (Ramseyer & Tschacher, 2011). Diese liegt momentan in der MEA 4.03a & MEA 4.03b Version vor (Download: http://psync.ch/?page_id=74). Innerhalb der Software können Optionen für die Erfassung der Motion Energy eingestellt werden (z.B. der Intensitätsschwellenwert). Zwei andere auf Funktionen in Matlab basierende Ansätze sind der von Altmann (2013) und der von Paxton und Dale (2013b). Innerhalb der Matlab-Ansätze sind zwei Regionen für die Interaktionspartner vorgegeben (50% des Videobildes). Damit können für ein Standardsetting mit gut positionierten Interaktionspartnern Zeitreihen der Bewegung bestimmt werden. Die Zeitreihen, die aus den drei verschiedenen Verfahren gewonnen werden, stimmen signifikant überein ($r = .894, \dots, .996$). Da das Projekt TIMPATHIN eine Sekundärdatenanalyse darstellt, gab es verschiedene Experimentalsettings, sodass den beiden Interaktionspartnern kein standardisierter Raum zugewiesen werden konnte. Vielmehr musste innerhalb jedes Videos festgelegt werden, wo Therapeut und Patient saßen. Dafür wurden die Matlab Skripte von Altmann (2013) weiterentwickelt. Es wurde eine grafische Benutzeroberfläche eingeführt, die es erlaubt, die ROIs frei Hand einzuzeichnen und automatisch abzuspeichern. Weiterhin wurde die ROI Größe jeder Person gespeichert, sodass die resultierende Zeitreihe an der Größe der verwendeten ROI standardisiert werden konnte. Diese Standardisierung machte es möglich, dass die Zeitreihen selbst bei verschiedener Größe der Personen trotzdem vergleichbar im Ausmaß der gezeigten Bewegungen waren und die

ROI-Größe zu keinen systematischen Verzerrungen führte (Skript auf GitHub: <https://github.com/10101-00001/MEA>).

Weiterhin wurde die Intensitätsschwelle μ , wie von Altmann (2013) empfohlen, empirisch bestimmt. Dafür wurden die in Abbildung 9 dargestellten Pixel im Hintergrund des Therapiesettings verwendet. Für diese Pixel wurden die Differenzwerte zwischen den Videoframes durch die MEA bestimmt. Da in diesen Bildbereichen keinerlei Bewegung stattfindet, gehen die in den Zeitreihen sichtbaren Intensitätsveränderungen auf Schwankungen des Lichtes, Bildrauschen o.Ä. zurück. Deshalb wurde bestehend auf $N = 20$ Videos das 99% Quantil der Motion Energy Werte bestimmt. Um Bewegungen gut von zufällig auftretenden Artefakten zu unterscheiden, wurde der resultierende Intensitätsschwellenwert von 12 verwendet.



Abbildung 9 *Kontroll-Pixel im Hintergrund (rot)*

Große Artefakte wie beispielsweise Lichtblitze oder Pixelstörungen könne durch diesen Schwellenwert jedoch nicht gefiltert werden. Dafür war es nötig, zwei zusätzliche Filterkriterien einzufügen: 1) Ähnlich wie bei der Bestimmung des nötigen Intensitätsschwellenwertes, wurden zwei ROIs mit der Größe von 10x10 Pixeln im Hintergrund gesetzt. Die Hintergrund ROIs waren entweder links und rechts im oberen oder links und rechts im unteren Bildbereich. Welche ROIs verwendet wurden, wurde durch die Einzeichnung der ROIs von Patient und Therapeut festgelegt: Befanden sich die ROIs der

interessierenden Personen eher im oberen Bildbereich, sodass es wahrscheinlich zu Überlappungen mit den oberen Hintergrund-ROIs kommen würde, so wurden die unteren verwendet; ansonsten die oberen. Durch die MEA wurden nicht nur die Zeitreihen für die eingezeichneten ROIs der Interaktionspartner erstellt, sondern ebenfalls Zeitreihen für die Hintergrund-ROIs. Anschließend wurden die Zeitreihen der Interaktionspartner auf *missing* gesetzt, wenn in einer der Hintergrund-ROIs eine Aktivierung von mehr als 5% gezeigt wurde. Dahinter steckt die Annahme, dass sich in den Hintergrund-ROIs keine Bewegung ereignen sollte. Wurde dennoch eine Pixelveränderung detektiert, so resultierte diese von Videoartefakten, wie in einer manuellen Sichtung der betreffenden Stellen des Videos festgestellt werden konnte. Dennoch gab es Videostörungen, die nur lokal auftraten und nicht das komplette Videobild betrafen. Dies waren zum Beispiel Pixelveränderungen, die kurzzeitig auftraten (für ein Videobild) und beispielsweise nur in der ROI des Patienten sichtbar waren. Diese Artefakte führten in der Zeitreihe zu Sprüngen, die dadurch gekennzeichnet waren, dass die Zeitreihe von einem Wert auf den anderen rasant anstieg, um beim nächsten Messwert wieder stark zu fallen. 2) Um eine statistische Differenzierung und Detektion solcher Phänomene zu schaffen, wurden alle Veränderungen, die die Zeitreihen zwischen zwei Messwerten zeigten, bestimmt und das 95% Quantil aller Differenzwerte ermittelt. Dies zeigte für Körperzeitreihen einen Wert von 15% und für Kopfzeitreihen einen Wert von 25%. Die Videostellen mit Sprüngen, die diese Schwellenwerte überstiegen, wurden gesichtet. Es wurde festgestellt, dass all die Differenzwerte, die über dem Schwellenwert lagen, durch Artefakte verursacht wurden. Deshalb wurden sie auf *missing* gesetzt. Im letzten Schritt wurden die fehlenden Werte durch lineare Interpolation der benachbarten Werte ersetzt, vorausgesetzt, dass nicht mehr als acht zusammenhängend fehlende Werte auftraten. Um kleine Unebenheiten in der Zeitreihe zu glätten, wurde zuletzt ein Moving Median mit einer geringen Bandbreite (5) eingesetzt.

4.2 Untersuchungen innerhalb der Originalarbeiten

In den nächsten Absätzen sind die in dieser Arbeit zusammengefassten Publikationen vorgestellt. Weiterhin werden Zusatzanalysen zu den Publikationen berichtet, die über die Analysen in den Publikationen hinausgehen aber zum übergreifenden Verständnis beitragen oder Quellen der Verzerrung der Ergebnisse ausschließen. Vor der publikationsübergreifenden Diskussion sind ebenfalls die Arbeitsanteile an den einzelnen Publikationen berichtet.

4.2.1 Untersuchung der konvergenten und kriteriumsorientierten Validität zeitreihenbasierter linearer Methoden (Publikation 1)

Schoenherr, D., Paulick, J., Worrack, S., Strauss, B., Rubel, J. A., Schwartz, B., . . . Altmann, U. (2018). Quantification of nonverbal synchrony using linear time series analysis methods: Lack of convergent validity and evidence for facets of synchrony. *Behavior Research Methods*, 1-23. doi:10.3758/s13428-018-1139-z

Wie bereits ausgeführt, gibt es zahlreiche linear-basierte Zeitreihenanalyse-Verfahren, die zur Bestimmung eines Synchronie-Maßes verwendet werden. Die Augenscheinvalidität der Verfahren ist vermehrt gegeben und Studien belegen oft die prädiktive Validität. Jedoch wurden die verschiedenen Maße nicht für einen einheitlichen Datensatz bestimmt. Zudem werden Parametereinstellungen für die einzelnen Verfahren oft nur unzureichend berichtet und erschweren die Nachvollziehbarkeit der Funktionsweise sowie die Transparenz. Innerhalb der ersten Originalarbeit wurden zuerst sieben verschiedene linear-basierte zeitreihenanalytische Verfahren vorgestellt und deren Funktionsweise und wichtige Parameter erläutert, um die Transparenz der Methoden zu erhöhen und so die Objektivität verwendeter Methoden zu sichern. Im zweiten Teil wurden diese sieben Methoden auf einen identischen Datensatz von $N = 84$ Zeitreihen, generiert aus Psychotherapievideos, angewendet. Dabei wurden verschiedene Synchronisationsmaße bestimmt (z.B. durchschnittliche Stärke der Synchronie, Häufigkeit der Synchronie, maximale Stärke der Synchronie). Diese Maße wurden zum einen unter Benutzung komplett identischer Parametereinstellungen für alle Algorithmen berechnet, um maximale Vergleichbarkeit innerhalb der Studie zu sichern. Zum anderen wurden sie mit den (wenn vorhanden) in der Literatur berichteten Einstellungen angewendet, um maximale Vergleichbarkeit zu bisherigen Ergebnissen zu schaffen. Die konvergente Validität wurde mittels Korrelationen und explorative Faktorenanalyse geprüft. Die prädiktive Validität wurde hinsichtlich der Veränderung interpersoneller Probleme in der Therapie und die konkurrente Validität in Bezug zu einer generalisierten non-linearen Methode (der Cross-Recurrence Analysis) untersucht.

Es stellte sich heraus, dass einige Synchronisationsmaße nur schwach miteinander korrelieren. Es konnte weder für einheitliche noch für literatur-basierte Parameter ein passendes Ein-Faktoren-Modell nachgewiesen werden. Für die in der Literatur berichteten Parameter konnte jedoch mit acht Synchronisationsmaßen, die sechs verschiedenen Algorithmen zuzuordnen sind, ein Drei-Faktoren-Modell mit übergeordnetem Second-Order Synchronie-Faktor nachgewiesen werden. Die Studie liefert Evidenz dafür, dass verschiedene

Algorithmen, seien sie auch alle linear-basiert, nicht das gleiche Konstrukt, sondern vielmehr verschiedene Facetten von Synchronie messen. Als Facetten konnten die Stärke der Synchronie der gesamten Interaktion, die Stärke der Synchronie innerhalb von Synchronisationsintervallen und die Häufigkeit der Synchronie unterschieden werden. In Bezug auf die prädiktive Validität, zeigten sich vor allem die Windowed Cross-Lagged Correlation sowie die Windowed Cross-Lagged Regression mit Peak-Picking Algorithmus von Altmann (2013) und die Windowed Cross-Correlation als valide. Beinahe alle Synchronisationsmaße, mit Ausnahme der Cross-Lagged Correlation und Cross-Lagged Regression zeigten signifikante Zusammenhänge zur generalisierten non-linearen Methode (literatur-basierte Parameter). Die Ergebnisse der Studie zeigen erstmalig, dass der Methodenfaktor bei der Bestimmung von Synchronie eine große Rolle spielt. Demnach sollte zukünftig auf die genaue Beschreibung des verwendeten Verfahrens sowie eine explizite Definition des Synchronie-Konstruktes geachtet werden. Bisherige widersprüchliche Untersuchungsergebnisse verschiedener Studien können auf die verwendete Methode zurückgehen.

4.2.2 Identifikation der besten Parameterkonfiguration für die zeitreihenbasierten Methoden (Publikation 2)

Schoenherr, D., Paulick, J., Strauss, B., Deisenhofer, A.-K., Schwartz, B., Rubel, J., Lutz, W., Stangier, U., Altmann, U. (2019). Identification of movement synchrony: Validation of windowed cross-lagged correlation and -regression with peak-picking algorithm. *PLoS One*, 14(2). doi:10.1371/journal.pone.0211494

Innerhalb dieser Originalarbeit wurden die als valide identifizierten Methoden Windowed Cross-Lagged Correlation mit Peak-Picking Algorithmus (WCLC_F) sowie die Windowed Cross-Lagged Regression mit Peak-Picking Algorithmus (WCLR) näher untersucht. Diese beiden Methoden und die Verwendung der Synchronie-Häufigkeit wurden gewählt, da sie 1) den Validitätsanforderungen genügen, nachgewiesen durch die erste Publikation, 2) die Häufigkeit der Synchronie in lokalen Synchronisationsintervallen operationalisieren und somit ein Vergleich der Lokalisierung von Synchronie zwischen menschlichen Ratern bzw. simulierten Zeitreihen und den Ergebnissen der Algorithmen leicht möglich ist und 3) Ratings der Lokalisierung von Bewegungssynchronie auf Grundlage der Videosequenzen und Zeitreihen einfacher sind als Ratings der Synchronie-Stärke. Bei der Synchronie-Stärke ist es schwierig einzuschätzen, ob beispielsweise Bewegungssynchronie, die durch gleichzeitige Kopfbewegungen entsteht, genauso stark ist wie Synchronie, die entsteht, weil auf die

Armbewegung des Interaktionspartners eine Rumpfbewegung gezeigt wird. Bisher sind jedoch für keine Methode Validierungs-basierte Empfehlungen für die verwendeten Parameter gegeben. Diese Forschungslücke wird für die $WCLC_F$ und $WCLR$ mit dieser Arbeit geschlossen. Beide Methoden wurden hinsichtlich artifiziell hergestellter Synchronisationsintervalle und menschlich-ingeschätzten Synchronisationsintervallen getestet. Außerdem wurde sichergestellt, dass die Methoden in Intervallen ohne Synchronisation ebenfalls valide sind, d.h. einen möglichst kleinen Anteil an falsch positiven Synchronie-Zeitpunkten bestimmen. Dabei wurden 600 verschiedene Parameterkonfigurationen angewendet. Die Parameterkonfigurationen unterschieden sich hinsichtlich der verwendeten Glättung und Transformation der Zeitreihen, der verwendeten Methode ($WCLC_F$, $WCLR$) und angewendeten Bandbreite sowie einem eingesetzten R^2 -Schwellenwert zur Filterung von zufällig auftretenden Synchronisationsintervallen. Die verschiedenen Parameterkonfigurationen wurden auf drei verschiedene Bedingungen angewendet: 1) natürlich eingebettete Synchronisationsintervalle/ Nicht-Synchronisationsintervalle 2) natürlich isolierte Synchronisationsintervalle/Nicht-Synchronisationsintervalle und 3) künstlich hergestellte Synchronisationsintervalle/Nicht-Synchronisationsintervalle. Die natürlich eingebetteten Synchronisationsintervalle/Nicht-Synchronisationsintervalle resultierten aus einem Ratingprozess dreier Rater, die 40 Sequenzen dichotom einschätzten (Synchronie vorhanden vs. Synchronie nicht vorhanden). Aus den übereinstimmend eingeschätzten Sequenzen wurden dann zufällig 10 Sequenzen ausgewählt, die Synchronie enthielten (Synchronisationsintervalle) und 10, die keine Synchronie enthielten (Nicht-Synchronisationsintervalle). Diese Synchronie-Sequenzen wurden in Bezug auf die Länge des Phänomens gematcht zu den Nicht-Synchronie-Sequenzen ausgewählt. Um die Sequenzen zeitreihenanalytischen Methoden zuführen zu können, wurde dann die Motion Energy Analysis durchgeführt. Auf Grundlage der entstandenen Zeitreihen, wurden die beiden weiteren Bedingungen gebildet. Dabei umfasste die natürlich isolierte Bedingung ausschließlich das eingeschätzte Synchronisationsphänomen. Alle anderen Bewegungen wurden in der Zeitreihe auf null gesetzt. Die künstliche Bedingung umfasste Zeitreihen, die die zusammenhängenden Peaks der Synchronisationsphänomene benutzen, diese jedoch duplizieren, um perfekte zeitversetzte Synchronie zu generieren. Um die Übereinstimmung der menschlichen Ratings und der Computerratings als auch die Übereinstimmung zwischen simuliertem Synchronisationsintervall und Computerratings zu quantifizieren, wurde Cohen's kappa berechnet. Für die Intervalle ohne Synchronisationsphänomene wurde der Prozentsatz

an überidentifizierten Zeitpunkten an der Gesamtdauer der Sequenz bestimmt. Es wurde die Parameterkonfiguration gewählt, die die beste Identifikationsrate aufwies, d.h. ein hohes kappa in Synchronisationssequenzen und ein geringen Prozentsatz überidentifizierter Frames in den Nicht-Synchronisationssequenzen zeigte. Dabei wurden die verschiedenen Bedingungen nach wachsender Komplexität der Zeitreihen getestet (künstliche Bedingung < natürlich isolierte Bedingung < natürlich eingebettete Bedingung) und nur die Konfigurationen geprüft, die sich auf einer geringeren Komplexitätsebene als zufriedenstellend gezeigt hatten.

In der künstlichen Bedingung konnten 76 Konfigurationen mit guter Identifikationsrate bestimmt werden. Eine geringe Glättung der Zeitreihen, geringe Bandbreiten, die Verwendung der WCLR sowie das Setzen eines R^2 -Schwellenwertes von 0.25 für die Synchronisationsintervalle zeigten sich als die Identifikationsrate positiv beeinflussende Faktoren. In der natürlich isolierten Bedingung konnten lediglich zwei Konfigurationen mit zufriedenstellender Identifikationsrate bestimmt werden (WCLC, Logarithmus-transformierte Zeitreihen, Bandbreite 75/125 Frames pro Sekunde (fps) bei 25fps, keine Glättung, R^2 -Schwelle 0.25); in der natürlich eingebetteten keine. Die Ergebnisse zeigen zum einen, dass die Wahl der Parameter mit der Güte der Bestimmung der Synchronie verbunden ist. Aus den zwei weniger komplexen Bedingungen konnten Empfehlungen für die Parameter abgeleitet werden, sodass in den folgenden inhaltlichen Publikationen die Methode WCLC, Logarithmus-transformierte Zeitreihen, Bandbreite 125fps bei 25fps, keine Glättung, R^2 -Schwelle 0.25 verwendet wurde. Zum anderen bleibt im Hinblick auf die menschlich eingeschätzten Intervalle zu konstatieren, dass die Konstrukte von Mensch und Algorithmus unterschiedlich sind. Dies kann dadurch begründet sein, dass Menschen mehrere Aspekte der Interaktion wahrnehmen (Blickkontakt, Semantik, etc.), die von zeitreihenanalytischen Methoden nicht zur Identifikation von Synchronisationsintervallen verwendet werden.

Zusatzuntersuchungen zur Validität der verwendeten R^2 Schwelle von 0.25

Da die R^2 -Schwelle in der Validierungsstudie das erste Mal verwendet wurde, um zufällig aufgetretene Synchronie zu filtern, wurde eine weitere kleine Untersuchung an zwei Beispielprobanden vorgenommen. Innerhalb des Projektes „Irritationsfreies und emotionssensitives Trainingssystem“ (IRESTRA) wurde am Institut für Psychosoziale Medizin und Psychotherapie untersucht, wie sich nonverbale Synchronie innerhalb von Trainingssituationen etabliert und was positive oder negative Prädiktoren sind. Dazu wurde eine Interaktion mit einem spezifischen Experimentalablauf verwendet: Zuerst wurde mit gesunden Probanden oder Patienten mit Fazialis-Parese ein standardisierter Small-Talk geführt

(z.B. wie die Probanden das Institut gefunden haben), danach folgte eine Meditationsphase. Im dritten Teil fand das eigentliche Training statt. Da es sich bei den eingeschlossenen Patienten um Patienten mit einer Gesichtslähmung handelte, wurden spezielle Yoga-Übungen zur Entspannung des Schulter- und Nackenbereichs sowie des Gesichtes angeleitet und durchgeführt. Für die Untersuchung der Güte des R^2 -Schwellenwertes ergeben sich hier Vorteile: Durch das standardisierte Setting ist festgelegt, wann Synchronie auftreten sollte (Small-Talk: natürliche Synchronie; Übungen: instruierte Synchronie) und wann keine Bewegungen und demnach keine Synchronie zu sehen ist (Meditation). In Abbildung 10 und 11 sind die Ergebnisse zweier zufällig ausgewählter Probanden gezeigt. Es wurden zwei Probanden ausgewählt, um zu zeigen, dass die Befunde unabhängig davon sind, wann die Meditation stattfindet und wie lange sie dauert. Die schwarzen Striche bzw. Rechtecke kennzeichnen die durch den Algorithmus identifizierten Synchronisationsintervalle (dichotom: vorhanden vs. nicht vorhanden). In der oberen Abbildung sieht man alle identifizierten Synchronisationsintervalle, in der unteren die Intervalle, die nach der Anwendung des R^2 -Schwellenwertes für die Synchronisationsintervalle bestehen geblieben sind. Die Abbildungen verdeutlichen den Nutzen des R^2 -Schwellenwertes: Durch seine Anwendung werden zufällig identifizierte Synchronisationsintervalle gefiltert und nur Intervalle mit sichtbaren Bewegungen werden für die Synchronisationsbestimmung benutzt. Somit wird eine Bewegungssynchronie, die eventuell auch mit geschlossenen Augen auftritt wie beispielsweise die Synchronie der Atemzyklen durch diese Anwendung des Schwellenwertes bewusst ausgeschlossen.

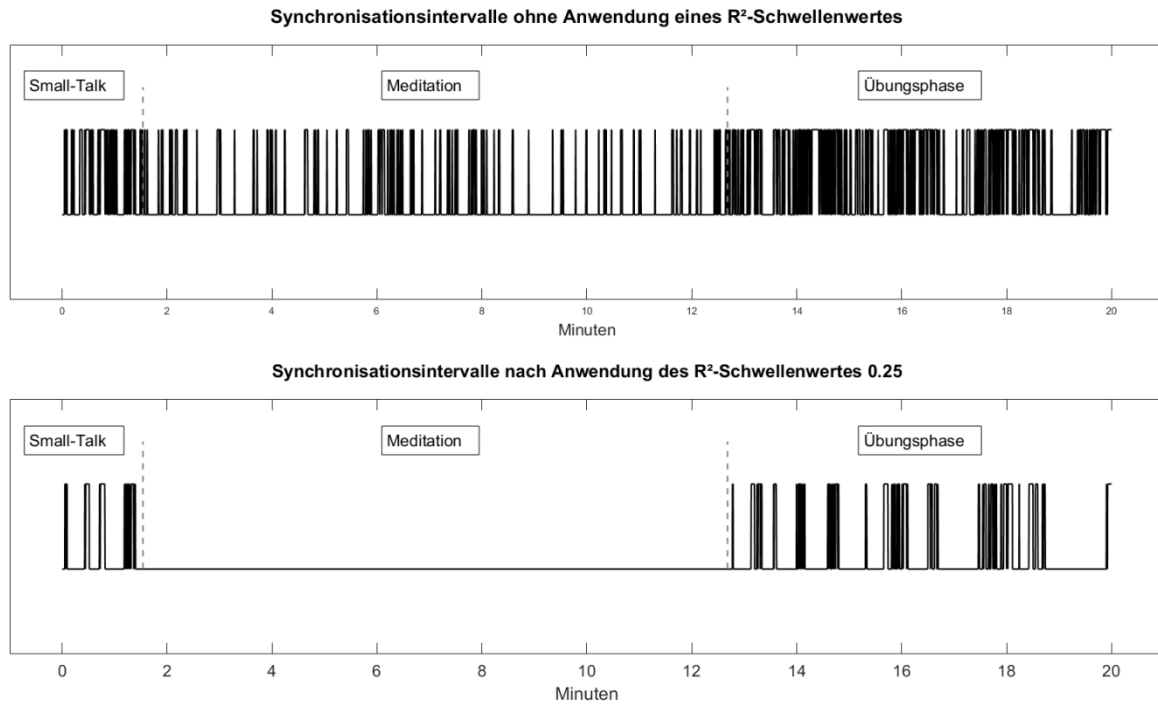


Abbildung 10 Vergleich aller Synchronisationsintervalle und der Intervalle bei Anwendung eines R²-Schwellenwertes (Beispielproband 1)

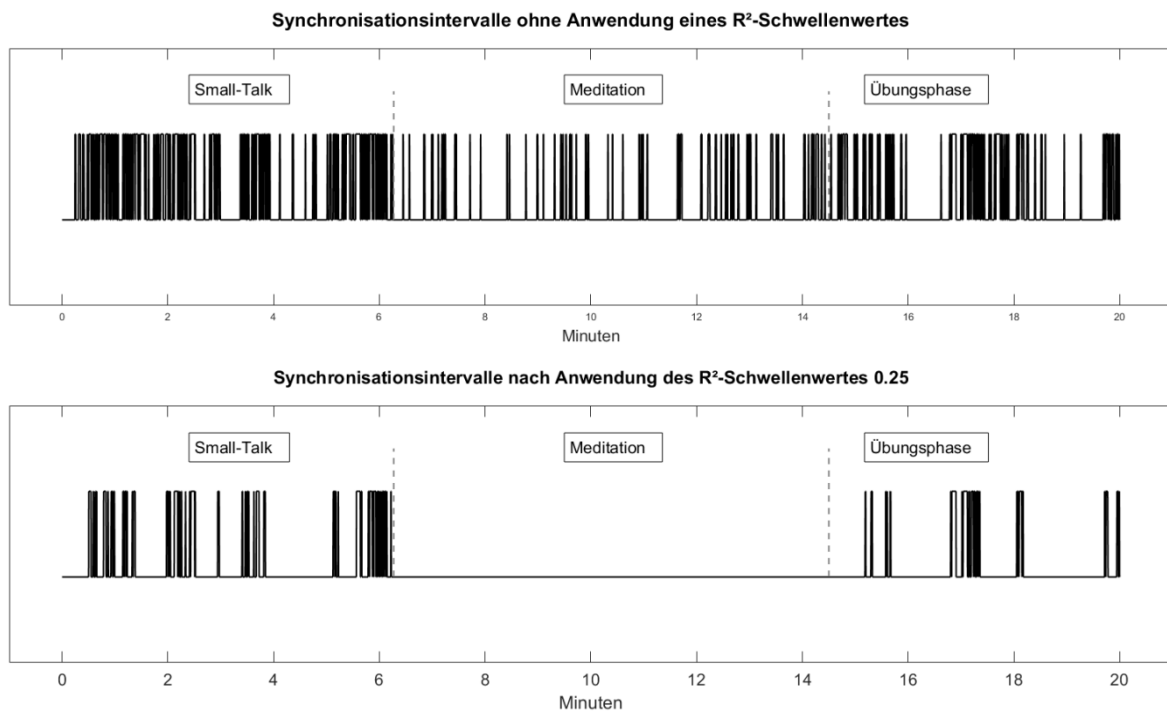


Abbildung 11 Vergleich aller Synchronisationsintervalle und der Intervalle bei Anwendung eines R²-Schwellenwertes (Beispielproband 2)

4.2.3 Zusammenhang zwischen Therapieabbruch und nonverbaler Synchronie (Publikation 3)

Schoenherr, D., Paulick, J., Strauss, B., Deisenhofer, A.-K., Schwartz, B., Rubel, J., Lutz, W., Stangier, U., Altmann, U. (zur Publikation angenommen, 06.12.2018). Nonverbal synchrony predicts premature termination of psychotherapy for social anxiety disorder. *Psychotherapy*.

Bewegungssynchronie erscheint als ein vielversprechender Prädiktor für die bisher nur unzureichend und widersprüchlich vorhersagbaren Risikogruppen für einen Therapieabbruch, da es den Prozess der Therapie näher charakterisiert. Deshalb wurde in der dritten Publikation der Zusammenhang zwischen Bewegungssynchronie und dem Therapieabbruch untersucht. Es wurde sich aufgrund der Störungsabhängigkeit von Synchronie auf eine Stichprobe von Patienten mit Sozialer Angststörung fokussiert (TIMPATHIN Stichprobe N = 267). Die Patienten durchliefen entweder eine manualisierte Kurzzeittherapie der kognitiven Verhaltenstherapie oder psychodynamisch-orientierten Therapie. Weiterhin gab es einen Therapiearm naturalistischer Therapien, die je nach Indikation auch in Langzeittherapien überführt werden konnten. Untersuchungsgegenstand dieser Arbeit war die Bewegungssynchronie innerhalb der 3. probatorischen Sitzung, also dem 3. Aufeinandertreffen von Patient und Therapeut. Diese wurde mittels MEA und der validierten Methode der WCLC mit Peak Picking Algorithmus erfasst. Methodisch wurde die Fragestellung, ob die Bewegungssynchronie den Therapieabbruch vorhersagen kann, auf mehrere verschiedene Arten bearbeitet. Zum einen wurden logistische Regressionen (Kriterium: Abbruch) verwendet, um die Ergebnisse mit bisherigen Studien vergleichbar zu machen. Um einen frühen Abbruch von späteren Abbrüchen zu unterscheiden, wurden ebenfalls multinomiale Regressionen mit kategorialer Dropout-Variable durchgeführt. Basierend auf der Diskussion zur Operationalisierung von Abbruch (vgl. Swift, Callahan, & Levine, 2009) wurde der *Clinically significant change* (CSC) der interpersonellen Probleme innerhalb der Therapie ebenfalls vorhergesagt. Um die Zeitabhängigkeit des Dropouts hochaufgelöst mit zu modellieren, wurde eine Cox Regression berechnet. In allen Modellen wurde für den Therapeuten im Rahmen von *random effects* kontrolliert. Alle Analysen wurden nicht nur mit der Gesamt-Häufigkeit der Synchronie, sondern ebenfalls mit den Anteilen, die der Patient initiierte und denen, die der Therapeut initiierte, durchgeführt.

Die Ergebnisse zeigen, dass Bewegungssynchronie ein valider Prädiktor für die Vorhersage von Therapieabbrüchen ist. Dabei scheint es für die Abbruchrate unabhängig von der Sitzungsanzahl vor Abbruch, besonders auf die durch den Patienten initiierte Synchronie

anzukommen. Um aber frühe Abbrüche zu vermeiden, haben beide Komponenten (Therapeut leitet Synchronie sowie Patient leitet Synchronie) eine Bedeutung. Betrachtet man das Kriterium klinisch signifikanter Veränderung (CSC), so fällt auf, dass für eine Verbesserung der interpersonellen Probleme gerade das Initiieren der Synchronie durch den Therapeuten prädiktiv ist. Besonders die Gruppe der Patienten, die ohne Verbesserung die Therapie abbrechen, wurde in der Probatorik eine geringe Synchronie zwischen Therapeut und Patient beobachtet. Eine Quantifizierung des Effekts durch die Cox Regression zeigte, dass sich die Abbruchrate pro steigendem Prozent an Bewegungssynchronie um 5% verringert (entspricht ca. 9 Sekunden zusätzlicher Bewegungssynchronie). Die Ergebnisse verdeutlichen die Bedeutung der nonverbalen Synchronie für die klinische Forschung und Praxis. Die Ergebnisse deuten darauf hin, dass durch eine vermehrte Bewegungssynchronie zu Beginn der Therapie Abbrüche verhindert werden könnten. Dabei kommen den verschiedenen Anteilen der Synchronie (Patient leitet vs. Therapeut leitet) wahrscheinlich unterschiedliche Bedeutungen zu. Wenn der Patient leitet, steht dies in Zusammenhang damit, ob er generell in der Therapie bleibt, hat jedoch keinen Bezug zur Veränderung seiner interpersonellen Probleme (gemessen durch das Inventar Interpersoneller Probleme [IIP]). Leitet der Therapeut, so ist dies auch wichtig, um einen frühen Abbruch zu vermeiden, jedoch hat es vorrangig einen Einfluss auf die interpersonellen Probleme des Patienten. Zukünftige Forschung sollte die Untersuchungen semantisch ankoppeln, um das Verständnis der differentiellen Wirkung von Synchronie zu ermöglichen. Innerhalb der Originalarbeit wurden weiterhin Subgruppenanalysen durchgeführt, die zeigten, dass ein hohes Maß an Synchronie vor allem bei gemischtgeschlechtlichen Dyaden und Patienten, die sich in psychodynamisch-orientierter Therapie befanden, wichtig war. Auch diese Ergebnisse unterstreichen die Notwendigkeit für das Koppeln der Synchronie-Befunde mit den semantischen Inhalten der Sitzungen.

Zusatzanalysen – andere Moderatoren

Innerhalb der Untersuchung wurden ebenfalls die Primäre Diagnose, der Bildungsstand, der aktuelle Beziehungsstatus, die Anzahl der komorbiden Diagnosen sowie die Belastung durch depressive Symptome (Beck Depression Inventory) innerhalb von Subgruppenanalysen auf moderierende Einflüsse untersucht. Aus Platzgründen wurde auf eine Darstellung innerhalb der Originalarbeit verzichtet (siehe Tabelle 9).

Tabelle 9 Hazard ratios mit 95% Konfidenzintervallen für verschiedene Subgruppenanalysen

| Partnerschaft | | | | |
|----------------------|--------------|--------------|-------------------|--------------|
| Kein Partner | Partner | | | |
| 0.94 [†] | 0.97 | | | |
| [0.89, 1.00] | [0.91, 1.03] | | | |
| Bildungsstand | | | | |
| Kein Abitur | Abitur | | | |
| 0.92* | 0.98 | | | |
| [0.86, 0.99] | [0.92, 1.04] | | | |
| BDI-Klassifikation | | | | |
| unter 8 | 9-13 | 14-19 | 20-28 | über 25 |
| .93 | 0.94 | 1.03 | 0.90 [†] | 1.26 |
| [0.84, 1.04] | [0.86, 1.02] | [0.93, 1.14] | [0.81,1.00] | [0.46, 3.40] |
| Anzahl der Diagnosen | | | | |
| 1 Diagnose | 2 Diagnosen | 3 Diagnosen | 4 Diagnosen | |
| 1.00 | 0.89* | 0.94 | 0.95 | |
| [0.91, 1.08] | [0.80, 0.99] | [0.85; 1.04] | [0.83, 1.09] | |
| Primärdiagnose | | | | |
| SAD | Keine SAD | | | |
| 0.94 | 0.97 | | | |
| [0.85, 1.04] | [0.92, 1.02] | | | |

Anmerkung: SAD = Soziale Angststörung, BDI = Beck Depressions-Inventar, [95% Konfidenzintervall], [†] marginal signifikanter Effekt, * signifikanter Effekt unter Nutzung eines 5% Fehlerniveaus.

Die Analysen zeigen einen möglichen moderierenden Einfluss des Beziehungsstatus und Bildungsstandes. Für die BDI-Klassifikation und die Anzahl der Störungen ließen sich keine konsistenten Befunde finden. Außerdem machte es keinen Unterschied in Bezug auf die gefundenen Zusammenhänge zwischen Therapieabbruch und Bewegungssynchronie, ob der Patient als Primär- oder als Sekundärdiagnose die Soziale Angststörung hatte. Die Ergebnisse in Bezug auf den Bildungsstand lassen gemeinsam mit den Befunden zum Dyadentyp vermuten, dass eine unzureichende Passung zwischen Therapeut und Patient (aufgrund demografischer Unterschiede) durch eine nonverbale Passung (in Form von Bewegungssynchronie) kompensiert werden kann. Betrachten wir Therapien von Patienten mit niedrigem Bildungsstand, so könnte es sein, dass wenn auf verbaler Ebene keine optimale Passung gefunden wird, die nonverbale Ebene umso wichtiger ist, um dadurch einen *common*

ground zu ebnet und den Patienten in der Therapie zu halten. Diese Erklärungen lassen sich jedoch aus unseren Daten nicht beweisen und sollten weiter untersucht werden.

Zusatzanalysen – andere Zeitpunkte

Innerhalb des Projektes TIMPATHIN wurden ebenfalls drei andere Messzeitpunkte untersucht. Diese drei Messzeitpunkte waren: Sitzung 8, Sitzung 20, Sitzung 29. Für die anderen drei Messzeitpunkte konnten keine Zusammenhänge zwischen Bewegungssynchronie und Therapieabbruch unter Verwendung eines Cox Regressionsmodells mit Propensity-Scores und random effects für den Therapeuten gefunden werden (Sitzung 8 Synchronie: $b = 0.00$, $p = .90$; Sitzung 20 Synchronie: $b = -0.00$, $p = .96$; Sitzung 29 Synchronie: $b = 0.01$, $p = .92$). Dabei wurden die Prädiktoren Synchronie und Anzahl komorbider Diagnosen verwendet (analog zur Publikation). Die Ergebnisse sind übersichtsartig mit einem reduzierten Modell in Abbildung 12 gezeigt. Dargestellt sind Unterschiede in der Abbruchrate zwischen Dyaden mit über- und unterdurchschnittlicher Bewegungssynchronie (HLM-Struktur nicht beachtet). Dabei sieht man, dass sich vor allem Dyaden mit hoher Bewegungssynchronie von Dyaden mit geringer Bewegungssynchronie hinsichtlich ihrer Abbruchrate unterscheiden (Sitzung 3, Bild A). Für die anderen drei Zeitpunkte ist dieser charakteristische Unterschied nicht zu sehen und ebenfalls statistisch nicht nachzuweisen (Sitzung 3 Loglikelihood Test = 5.16, $df = 1$, $p = .03$; Sitzung 8 Loglikelihood Test = 0.38, $df = 1$, $p = .50$; Sitzung 20 Loglikelihood Test = 0.32, $df = 1$, $p = .60$; Sitzung 29 Loglikelihood Test = .01, $df = 1$, $p = .90$).

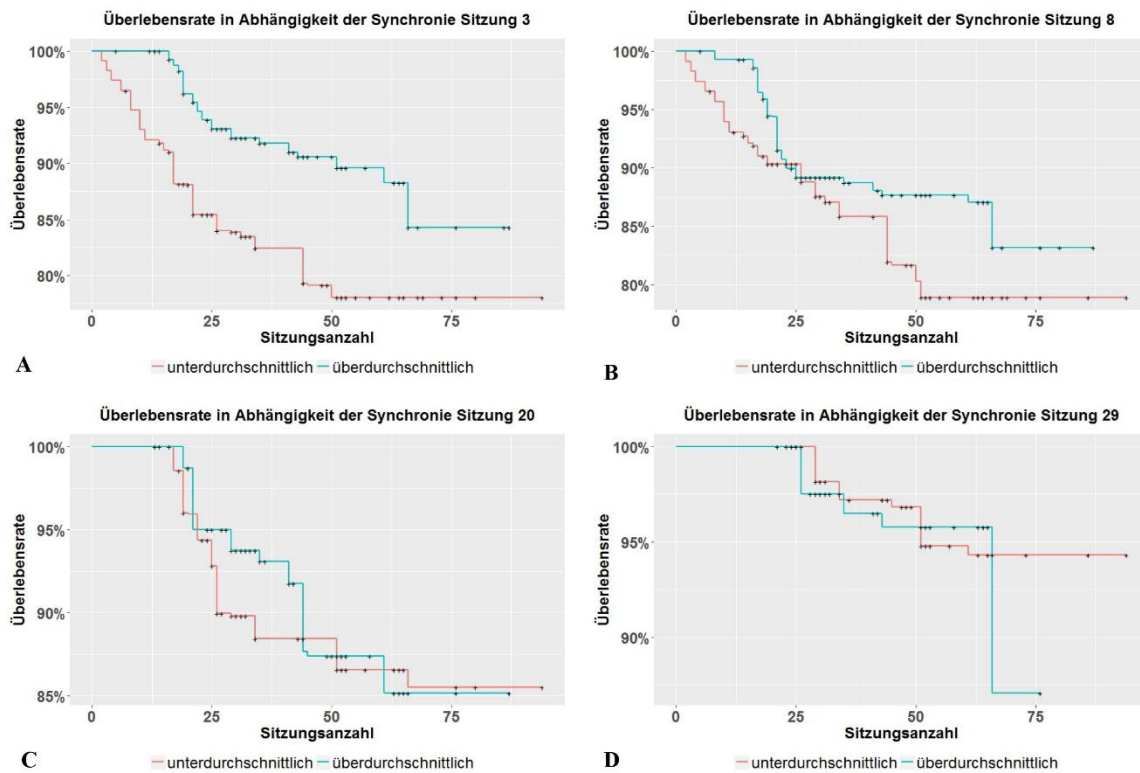


Abbildung 12 Veranschaulichung der Ergebnisse der Cox Regressionen zu den anderen Zeitpunkten

Auch diese Befunde verdeutlichen, dass Bewegungssynchronie kein immer gleich wirkender Faktor innerhalb von Psychotherapien ist. Innerhalb der dritten Sitzung der psychodynamisch-orientierten Therapie wurden Beziehungs-Episoden-Interviews durchgeführt, die in besonderer Art interpersonelle Beziehungen in Vergangenheit und Gegenwart beleuchten. Dadurch erscheint es durchaus plausibel, dass Bewegungssynchronie als interpersonelles Konstrukt in diesen Sitzungen für Patienten besonders wichtig ist, um in der Therapie zu bleiben.

Zusatzanalysen zum Vergleich des Algorithmus mit Anwendung zweier verschiedener maximaler time lags

Da für die inhaltlichen Analysen der von Ramseyer und Tschacher (2011) vorgeschlagene maximale time lag von 5 Sekunden (= 125 Frames) und nicht der in der Validierungsstudie benutzte time lag von 3 Sekunden verwendet wurde (= 75 frames), wurde mit einem Teil des Datensatzes eine Sensitivitätsanalyse durchgeführt. Dazu wurden die N = 84 Zeitreihen der manualisierten KVT und PDT, Sitzung 3 einer erneuten Zeitreihenanalyse unterzogen. Die Parameter dafür stammten aus der Validierungsstudie, lediglich der maximale time lag wurde von 3 Sekunden auf 5 Sekunden verändert. Folgende

Korrelationen konnten zwischen den Indizes der Verfahren mit verschiedenen time lags bestimmt werden: Synchronisationshäufigkeit $r = .97^*$, Synchronisationshäufigkeit geleitet durch den Patienten: $r = .93^*$, Synchronisationshäufigkeit geleitet durch den Therapeuten: $r = .94^*$, Synchronisationsstärke innerhalb der Synchronisationsintervalle: $r = .87$, Mittlerer time lag: $r = .53^*$. Die Korrelationen der Synchronie-Indizes zeigen, dass die sich die Ergebnisse bei der Verwendung von 3 Sekunden als maximalem time lag und 5 Sekunden als maximalem Time lag nicht wesentlich voneinander unterscheiden. Der größte Unterschied besteht bei der Bestimmung des mittleren time lags, was jedoch plausibel ist, da der Wert direkt von der Festlegung des maximalen time lags abhängt.

Zusatzanalyse: Robustheit der Effekte unter Hinzunahme der Variable „therapeutische Allianz“, gemessen durch den Helping Alliance Questionnaire

Um die Ergebnisse der Analyse zum Zusammenhang zwischen nonverbaler Synchronie und Therapieabbruch dahingehend abzusichern, dass die nonverbale Synchronie nicht nur die therapeutische Allianz abbildet, wurden Zusatzanalysen durchgeführt. In diesen wurde zusätzlich die therapeutische Allianz in die Modelle aufgenommen. Sie wurde durch den *Helping Alliance Questionnaire* (HAQ) um die 13. Sitzung gemessen. In Tabelle 10 sind die Synchronie-Koeffizienten zum Vergleich dargestellt. Die in Klammern dargestellten Koeffizienten beziehen sich auf die Modelle, die zusätzlich den Prädiktor HAQ enthielten.

Tabelle 10 *Vergleich der Synchronie-Koeffizienten mit und ohne Hinzunahme des HAQ als Prädiktor*

| | | Modell 1 | Modell 2 | Modell 3 |
|---|----------------|---------------------------|--------------------------------|----------------------------------|
| | | Synchronie total | Patienten-geleitete Synchronie | Therapeuten-geleitete Synchronie |
| Logistische Regression (Kriterium: Dropout dichotom) | | $b = -0.06^*$ (-0.06*) | $b = -0.13^*$ (-0.12*) | $b = -0.08$ (-0.08) |
| Multinomiale Regression (Kriterium: Dropout kategorial) | Früher Abbruch | $b = -0.22^*$ (-0.25*) | $b = -0.46^*$ (-0.74) | $b = -0.36^*$ (-0.40*) |
| | Später Abbruch | $b = 0.04$ (0.04) | $b = 0.07$ (0.07) | $b = 0.08$ (0.08) |
| Logistische Regression (Kriterium: Klinisch) | | $b = 0.08$ (0.08) | $b = 0.07$ (0.08) | $b = 0.22^*$ (0.23*) |

| bedeutsame Veränderung) | | | | |
|--|---------------------------|--|---------------------------|---------------------------|
| Multinomiale Regression (Kriterium: Klinisch Signifikante Veränderung x Abbruch) | Regulär beendet | $b = -0.07$ (-0.08) | $b = -0.07$ (-0.08) | $b = -0.20^*$ (-0.22*) |
| | Verbesserung | $b = -0.08$ (-0.18) | $b = -0.21$ (-0.36) | $b = -0.08$ (-0.25) |
| | Abbruch ohne Verbesserung | $b = -0.20^*$ (-0.20*) | $b = -0.31^*$ (-0.29*) | $b = -0.41^*$ (-0.41*) |
| | Abbruch mit Verbesserung | $b = -0.05^*$ (-0.04 ⁺) | $b = -0.11^*$ (-0.10*) | $b = -0.07$ (-0.07) |

Anmerkung: Die Modelle sind äquivalent zu den in der Publikation berechneten Modellen, für Details siehe Publikation.

Die Ergebnisse zeigen, dass Bewegungssynchronie einen Effekt hat, der über die therapeutische Allianz, erhoben durch den HAQ, hinausgeht. Sie sind auch unter Hinzunahme des zusätzlichen Prädiktors robust. Bewegungssynchronie misst demnach etwas, was über die therapeutischen Allianz hinausgeht.

4.2.4 Zusammenhang zwischen Bindungsangst und –vermeidung und nonverbaler Synchronie (Publikation 4)

Schoenherr, D., Strauss, B., Paulick, J., Deisenhofer, A.-K., Schwartz, B., Rubel, J., Lutz, W., Stangier, U., Altmann, U. (in Begutachtung, eingereicht am 07.03.2019). Movement synchrony and attachment related anxiety and avoidance in social anxiety disorder. *Journal of Counseling Psychology*.

Die Bindungstheorie kann auch als Regulationstheorie verstanden werden, da der Bindungsstil maßgeblich beeinflusst, wie wir uns selbst in Bezug auf die Umwelt oder im Umgang mit anderen Menschen regulieren können. Befunde der nonverbalen Synchronie weisen darauf hin, dass frühe Synchronie-Erfahrungen mit einer positiven Entwicklung einhergehen (Feldman, 2012, 2015). Außerdem legen Studien aus dem therapeutischen Kontext nahe, dass auch hier ein Zusammenhang zwischen Bindungsmustern und Synchronisationsfähigkeit besteht (Håvås et al., 2015; Ramseyer & Tschacher, 2011). Deshalb sollte im Rahmen dieser Originalarbeit

untersucht werden, wie die Bindungsdimensionen Bindungsangst sowie Bindungsvermeidung, gemessen durch die deutsche Version des *Experiences in Close Relationship Questionnaire*, mit nonverbaler Synchronie im therapeutischen Kontext (gemessen in Sitzung 8) zusammenhängen. Da bei Patienten mit Sozialer Angststörung ein großer Zusammenhang zwischen Bindungsangst und –vermeidung sowie Sozialer Angst und Vermeidung besteht, wurde der Zusammenhang zwischen Synchronie und Sozialer Angst/Vermeidung ebenfalls untersucht. Die Soziale Angst/Vermeidung wurde durch die deutsche Version der *Liebowitz Social Anxiety Scale* gemessen. Neben Korrelationen zwischen den Variablen wurden komplexe hierarchisch lineare Modelle angewendet. Diese schlossen die Bindungsvariablen und Angstvariablen gleichermaßen ein, um die Schlussfolgerungen der Korrelationen unter Hinzunahme wichtiger Kovariaten zu prüfen. Es wurden neben der totalen Häufigkeit der Bewegungssynchronie die Leading/Pacing-Anteile der Patienten und Therapeuten unterschieden und der Mittelwert sowie der Range der Verzögerung (d.h. des time lags) betrachtet. Für die Maße des time lags wurde ebenfalls zwischen dem time lag des Patienten und dem des Therapeuten unterschieden.

Die Ergebnisse zeigten, dass die in der achten Sitzung gezeigte Synchronie zwischen Therapeut und Patient nicht durch die Bindungsangst und –vermeidung des Patienten zu Beginn der Therapie vorhergesagt werden konnte. Die nonverbale Synchronie, die in der Therapie erlebt wird, kann jedoch vorhersagen, wie groß die Bindungsangst des Patienten zum Ende der Therapie ist. Je höher die erlebte Synchronie, desto geringer ist die Bindungsangst des Patienten am Ende der Therapie. Dabei kommt es vor allem auf die vom Therapeuten geleitete Synchronie an. Dieser Befund konnte speziell in der KVT nachgewiesen werden. In Bezug auf den time lag wurde festgestellt, dass hohe Werte der Bindungsvermeidung mit höheren time lags zu Beginn und zum Ende der Therapie in Zusammenhang stehen. Wie zu erwarten, war dabei vor allem der time lag des Patienten prädiktiv. Die Zusammenhänge zwischen Sozialer Angst und Vermeidung, die in den Korrelationen gefunden wurden, konnten durch komplexere Modelle nicht bestätigt werden. Dies deutet darauf hin, dass Bewegungssynchronie eher eine proximale Variable zu interpersonellen Variablen (z.B. Bindung, interpersonelle Probleme) ist. In Bezug auf Bindungsmuster scheint es für einen Therapeuten sinnvoll in der ersten Phase der Therapie eine leitende Funktion einzunehmen und dem Patienten das Gefühl eines sicheren Hafens zu geben.

4.2.5 Vorhersage von Veränderungen der interpersonellen Probleme und Depression durch nonverbale Synchronie (Publikation 5)

Altmann, U., Schoenherr, D., Paulick, J., Deisenhofer, A.-K., Schwartz, B., Rubel, J., Stangier, U., Lutz, W., Strauss, B. (in Begutachtung, eingereicht am 14.02.2019). Associations between movement synchrony and outcome in patients with social anxiety disorder: Evidence for treatment specific effects. *Psychotherapy Research*.

In dieser Originalarbeit wurde untersucht, ob der durch die Forschung oft für Bewegungssynchronie postulierte Zusammenhang zwischen Therapieoutcome und nonverbaler Synchronie auch innerhalb einer homogenen Störungsgruppe von Patienten mit Sozialer Angststörung vorhanden ist. Dafür wurden die in TIMPATHIN vorliegenden drei Therapierichtungen manualisierte KVT, naturalistische KVT und manualisierte PDT miteinander verglichen, um zu untersuchen, ob nonverbale Synchronie ein allen Therapierichtungen zugrundeliegender Wirkfaktor ist. Zur Quantifizierung des Therapieerfolgs wurden die Veränderungen der interpersonellen Probleme, gemessen durch die deutsche Version des *Inventory of Interpersonal Problems*, die Veränderung der Depressionen, gemessen durch die deutsche Version des *Beck-Depression-Inventory*, und die therapeutische Allianz, erfasst durch den *Helping Alliance Questionnaire* verwendet. Bewegungssynchronie wurde erneut über Motion Energy Analysis und eine nachfolgende Zeitreihenanalyse operationalisiert. Die Zusammenhänge wurden mittels hierarchisch linearer Modelle unter Kontrolle eines Therapeuteneffekts untersucht. Weiterhin wurde eine Leading-Variable in die Modelle aufgenommen, die indizierte, ob die Synchronie eher vom Patienten oder vom Therapeuten geleitet wurde. Der Zeitverlauf der Synchronie wurde zwischen den Therapierichtungen verglichen.

Die Ergebnisse zeigen, dass Bewegungssynchronie in beiden Bedingungen mit KVT häufiger vorkam als in der PDT. In den manualisierten Therapien ist der Verlauf der Bewegungssynchronie über die Therapiedauer fast konstant. In der naturalistischen KVT ist ein Abfall der Synchronie hin zur 29. Sitzung zu beobachten. Betrachtet man die reine Bewegungshäufigkeit, so konnte auch festgestellt werden, dass in sich Therapeuten in den KVT Bedingungen signifikant häufiger bewegten als Therapeuten in der PDT. Das Leading/Pacing Verhalten von Therapeut und Patient scheint sich innerhalb der Therapierichtungen nicht zu unterscheiden.

In beiden KVT Gruppen konnte ein Zusammenhang zwischen den Veränderungen der interpersonellen Probleme und der Bewegungssynchronie in Sitzung 3 und 8 gezeigt werden.

Häufige Bewegungssynchronie führte zu einer Verringerung der interpersonellen Probleme. Dabei hatte der Anteil an Patienten-geleiteter Synchronie einen eher negativen Einfluss. Leitet der Patient die Synchronie hauptsächlich und der Therapeut folgt stets, so führt dies zu höheren interpersonellen Problemen und höheren Depressionen beim Patienten. Das Multi-Gruppen-Modell zeigte, dass die Effekte des Leading vor allem durch die manualisierte KVT verursacht werden. Die Effekte genereller Synchronie sind in beiden KVT Bedingungen aber nicht in der PDT sichtbar. In der PDT und naturalistischer KVT konnten jedoch höhere Zusammenhänge zwischen Bewegungssynchronie in der 3. und 8. Sitzung und der therapeutischen Allianz in Sitzung 20 gefunden werden. Eine höhere Bewegungssynchronie prädizierte eine höhere therapeutische Allianz. Der Effekt scheint größer zu sein, wenn der Therapeut in Sitzung 8 leitet und der Patient folgt.

Der Zusammenhang zwischen Therapieoutcome und Bewegungssynchronie konnte auch in unserer Stichprobe gezeigt werden, manifestiert sich jedoch vor allem für die Vorhersage von interpersonellen Problemen in der KVT. In manualisierter KVT kommt es überdies zu negativen Einflüssen (d.h. einer geringeren Reduktion der interpersonellen Probleme des Patienten am Ende der Therapie, höheren Depressionswerten), wenn der Therapeut den Patienten häufig imitiert (Patient Leading gering, Therapeut Pacing hoch). Dieses Verhalten hat in der PDT und naturalistischen KVT einen negativen Einfluss auf die therapeutische Allianz.

Zusatzanalyse: Veränderung der Synchronie-Häufigkeit und der Leading-Anteile innerhalb der Therapie

Da in der Publikation vorrangig die Synchronie der verschiedenen Behandlungskonzepte verglichen wurde, wurden gepaarte *t*-Tests zum Vergleich der Synchronie-Häufigkeit im Therapieverlauf durchgeführt. Die Ergebnisse sind in den Abbildungen 13 zusammengefasst.

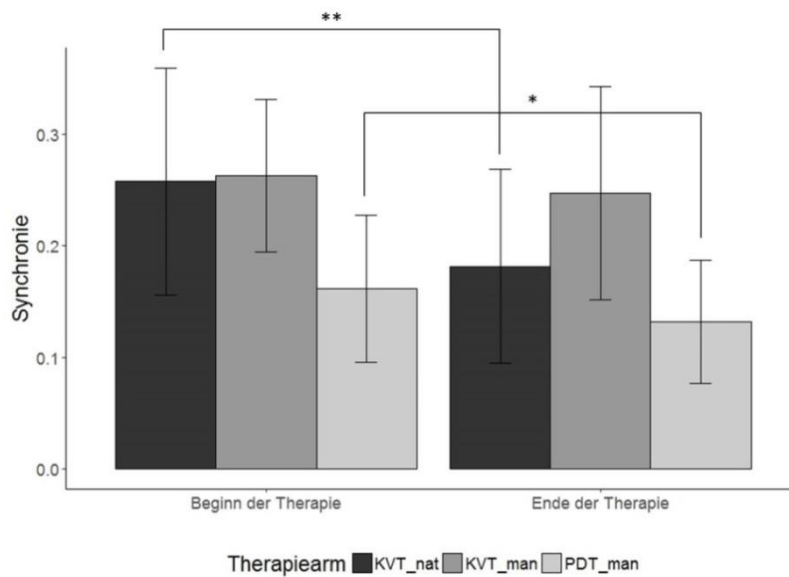


Abbildung 13 Vergleich der Synchronie zu Beginn der Therapie und zum Ende der Therapie innerhalb der verschiedenen Therapierichtungen; t-Test für verbundene Stichproben: ** entspricht $p < .01$, * entspricht $p < .05$, Abbildung und t-tests zur Illustration ergänzt.

Die Ergebnisse zeigen den auch von Paulick et al. (2018) postulierten Abfall der Synchronie-Häufigkeit in der naturalistischen KVT. Für die PDT konnte dieser Abfall ebenfalls gefunden werden. Die manualisierte KVT scheint einen besonderen Standpunkt einzunehmen, da die Synchronie-Häufigkeit zu Beginn der Therapie und zum Ende der Therapie nicht signifikant verschieden sind. Dieser Befund könnte darauf hindeuten, dass die Überangepasstheit der Patienten zum Ende der Therapie abnimmt. Auf Therapeutenseite zeigt dies ein stärkeres Bezugnehmen auf den Patienten zu Beginn der Therapie, wenn die Etablierung einer therapeutischen Allianz das Hauptziel darstellt. Im späteren Therapieverlauf scheinen andere Interventionen ein geringeres Level an Synchronie zu zeigen (siehe publikationsübergreifende Diskussion).

Die Veränderungen der Leading/Pacing-Anteile innerhalb der Therapie sind in Abbildung 14 veranschaulicht.

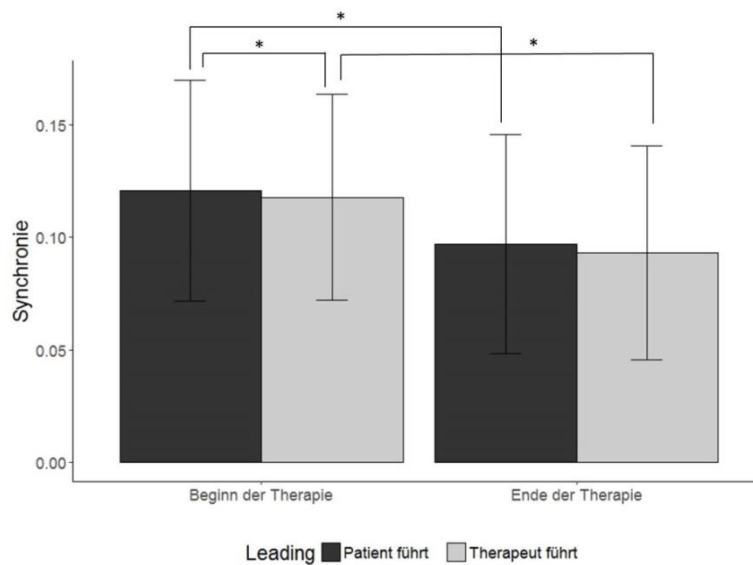


Abbildung 14 Vergleich der Leading-Anteile zu Beginn der Therapie und zum Ende der Therapie, * entspricht $p < .05$

Zu Beginn der Therapie liegt der höhere Anteil des Leadings beim Patienten. Es wurden demnach häufiger synchrone Bewegung durch den Patienten initiiert. Am Ende der Therapie unterschied sich das Ausmaß der Leading-Anteile nicht mehr. Es gibt gleichermaßen Synchronisationsintervalle, die vom Patienten geleitet werden als auch Synchronie, die vom Therapeuten initiiert wird. Es ist plausibel, dass zu Beginn der Therapie, der Patienten durch die Erklärung seiner Symptome die Sitzungen eher leitet.

4.2.6 Zusammenhang zwischen Bewegungsparametern und Selbst-Ratings zu interpersonellen Problemen (Zusatzanalyse)

Da bisherige Forschungsergebnisse in Bezug auf die nonverbalen Korrelate von Angst teilweise widersprüchlich sind (Freezing-Hypothese vs. Hypothese erhöhter Bewegungen aufgrund von Nervosität und Unruhe), wurde das allgemeine Bewegungsverhalten der Patienten näher untersucht. Dafür wurden Patienten eingeschlossen, von denen Motion Energy Zeitreihen vom Anfang der Therapie (s3) als auch vom Ende der Therapie (s29) vorhanden waren ($N = 86$). Die Bewegungshäufigkeit und die Bewegungsgeschwindigkeit wurde in Anlehnung an die von Grammer et al. (1999) postulierten Maße bestimmt. Dafür wird zuerst ein personenabhängiger Schwellenwert für eine minimale Bewegung gewählt, der Hintergrund-Rauschen optimal von Bewegungen trennt. Danach werden die einzelnen Motion Energy Werte, die den Schwellenwert überschreiten, analysiert. Die Fläche unter den Peaks operationalisiert die Bewegungsgröße. Die Dauer des Peaks charakterisiert die Bewegungsdauer (siehe Abbildung 15). Der Quotient dieser beiden Werte ist die

Bewegungsgeschwindigkeit jeder Bewegung. Weiterhin kann die Bewegungshäufigkeit als prozentuale Zeit, in der sich bewegt wurde, bezeichnet werden (in Referenz zur Gesamtzeit der Interaktion). Das heißt, sie gibt an, wie viele Motion Energy Werte den Schwellenwert überstiegen haben.

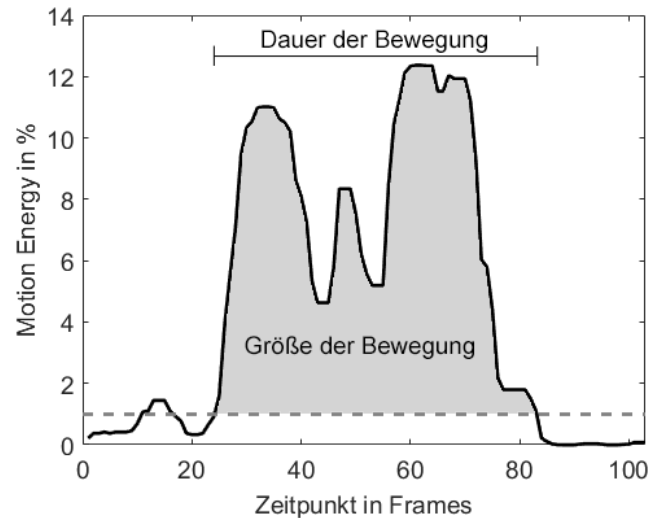


Abbildung 15 Grammer-Maße Größe der Bewegung und Dauer der Bewegung an einem Zeitreihen-Beispiel

In Bezug auf die verwendeten Daten wurde Hintergrund-Rauschen bereits durch diverse Vorverarbeitungsschritte minimiert, sodass kein personenabhängiger Schwellenwert sondern ein konstanter Schwellenwert von 1 verwendet wurde. Der Schwellenwert von 1 wurde gewählt, um kleinste Bewegungen, die beispielweise durch das Heben und Senken des Brustkorbes beim Atmen verursacht werden, auszuschließen. Das R-Skript, was zur Erfassung der Maße geschrieben und benutzt wurde, befindet sich im digitalen Anhang dieser Arbeit, zusammen mit einer Funktion, die für die Stapelverarbeitung der Daten verwendet wurde. Die Funktion kann mit einem konstanten oder personenspezifischen Schwellenwert benutzt werden. Output der Funktion ist eine Text-Datei, in der die Bewegungsmaße der Patienten gesammelt sind. Eine kurze Anleitung, welche Parameter für das Benutzen der Funktion gesetzt werden müssen, befindet sich im ebenfalls Anhang. Anschließend wurde untersucht, ob sich die Bewegungsmaße vom Anfang der Therapie zum Ende der Therapie veränderten (für Therapeuten und Patienten; t -test für verbundene Stichproben) und ob sie mit demografischen Variablen oder den IIP-Skalen assoziiert sind. Für die dichotomen Variablen wurden t -tests verwendet, für kategoriale, nicht ordinale Daten eine ANOVA und für metrische Variablen Pearson Produkt Moment Korrelationen. Die deskriptiven Daten der Stichprobe sind im Anhang dargestellt.

Es konnte weder für Patienten, noch für die Therapeuten ein Unterschied der Bewegungsmaße von prä zu post festgestellt werden (Patienten: PrB: $t = 1.18$, $df=85$ $p = .24$; Speed: $t = -.834$, $df = 85$, $p = .41$; Therapeuten: PrB: $t = 1.22$, $df = 85$ $p = .22$; Speed: $t = -1.23$, $df = 85$, $p = .22$). Die Ergebnisse der Korrelationsanalysen, t -tests und ANOVAs zeigen einen Einfluss des Therapiearms (s3 Bewegungsgeschwindigkeit) sowie Zusammenhänge zwischen der FG Skala des IIPs (zu introvertiert/vermeidend) und den Bewegungsmaßen (mit Ausnahme von der Bewegungshäufigkeit zu s3; siehe Tabelle 11). Weiterhin konnte ein Zusammenhang zwischen einer abweisend/kalten Haltung am Ende der Therapie sowie dem Gesamtscore des IIP zum Ende der Therapie und wenig Bewegungen gefunden werden.

Tabelle 11 *Ergebnisse Korrelationen, t-test und ANOVA für die Abhängigkeiten der Bewegungsparameter von demografischen Variablen und den IIP-Skalen*

| | PrB | | Speed | |
|---------------------------|--|--|--|---|
| | S3 | S29 | S3 | S29 |
| Geschlecht | $t = .70$, $df = 84$, $p = .49$ | $t = 1.15$, $df = 84$, $p = .25$ | $t = .72$, $df = 84$, $p = .48$ | $t = -.85$, $df = 84$, $p = .93$ |
| Abitur | $t = 1.21$, $df = 84$, $p = .23$ | $t = -.037$, $df = 84$, $p = .71$ | $t = .93$, $df = 84$, $p = .36$ | $t = -.14$, $df = 84$, $p = .89$ |
| Partner | $t = -.70$, $df = 84$, $p = .49$ | $t = .75$, $df = 84$, $p = .46$ | $t = -.28$, $df = 84$, $p = .78$ | $t = -.31$, $df = 84$, $p = .76$ |
| Therapiearm | $F = 2.33$, $df = 2;83$, $p = .10$ | $F = 1.74$, $df = 2;83$, $p = .18$ | $F = 6.82$, $df = 2;83$, $p < .05$ | $F = .77$, $df = 2;83$, $p = .46$ |
| Alter | .086 | .029 | .000 | .141 |
| | IIP Skalen | | | |
| PA – dominant | -.014 | -.095 | .079 | -.096 |
| BC – streitsüchtig | .009 | -.158 | .100 | -.030 |
| DE – zu abweisend/kalt | -.026 | -.223* | -.112 | -.191 |

4 Datengrundlage & Inhalte der einzelnen Publikationen

| | | | | |
|---------------------------|-------|--------|--------|--------|
| FG – sozial vermeidend | -.139 | -.361* | -.225* | -.230* |
| HI – selbstunsicher | -.032 | -.202 | -.036 | -.051 |
| JK – zu nachgiebig | -.045 | -.139 | -.036 | .049 |
| LM – zu freundlich | .060 | .013 | .047 | .078 |
| NO – zu aufdringlich | .060 | .022 | .093 | .168 |
| Gesamt | -.043 | -.227* | -.030 | -.071 |

Anmerkungen: PrB = Prozentsatz an Bewegungen = Bewegungshäufigkeit, Speed = Geschwindigkeit der Bewegungen, *t*-test für Geschlecht, Abitur, Partner, ANOVA für Therapie, Pearson-produkt Moment Korrelationen für IIP Skalen, Alter.

Als weitere Analyse wurde ein Strukturgleichungsmodell modelliert, was für die FG-Skala des IIPs die Zusammenhänge beider Messzeitpunkte modelliert. Dafür wurde zum einen die Regression von den Bewegungsmaßen (Bewegungshäufigkeit und Bewegungsgeschwindigkeit) auf die IIP Werte des Prä-Zeitpunktes und zum anderen die Regression der Bewegungsmaße auf die IIP-Werte des Post-Zeitpunktes spezifiziert. Innerhalb des Modells wurde weiterhin über einen Modell-Test geprüft, ob die Ladungen der Regressionen gleich sind. Lässt man eine Kovarianz zwischen dem Bewegungsparameter zu s29 und dem IIP-Wert des Prä-Zeitpunktes zu, so passen beide Modelle auf die Daten (Bewegungshäufigkeit: $\chi^2 = 1.03$, $df = 1$, $p = .31$, RMSEA = .02; Bewegungsgeschwindigkeit: $\chi^2 = 1.05$, $df = 1$, $p = .31$, RMSEA = .03). Für beide Modelle sind die Ladungen jedoch unterschiedlich voneinander (Bewegungshäufigkeit: $diff = .02$, $p = .03$, Bewegungsgeschwindigkeit: $diff = .03$, $p = .02$). Die Pfad-Koeffizienten sind in Abbildung 16 dargestellt (Bewegungshäufigkeit: A, Bewegungsgeschwindigkeit: B).

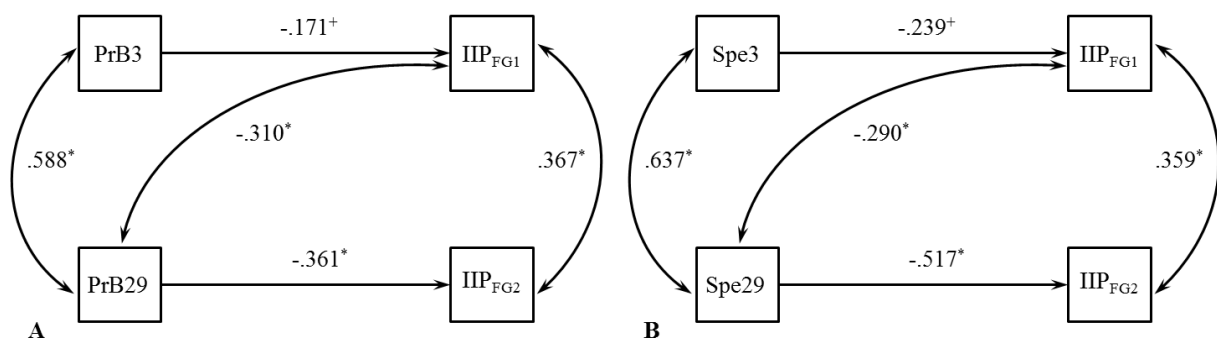


Abbildung 16 Strukturgleichungsmodelle für die Vorhersage der IIP Skala FG (introvertiert, sozial vermeidend) durch die Bewegungsparameter Häufigkeit und Geschwindigkeit vor der Therapie (IIP_{FG1}) und nach der Therapie (IIP_{FG2}).

Die Ergebnisse weisen darauf hin, dass die Bewegungshäufigkeit und -geschwindigkeit mit der sozialen Vermeidung zusammenhängt, in dem Sinne, dass sich Patienten mit größeren Schwierigkeiten im interpersonellen Kontakt weniger und langsamer bewegen. Da sich die Bewegungsparameter innerhalb der Therapie nicht verändern, weisen die Befunde eher auf soziale Vermeidung als Persönlichkeitseigenschaft hin. Personen mit höherer sozialer Vermeidung scheinen sich demnach weniger und langsamer als Personen zu bewegen, die nicht introvertiert sind. Dies könnte auch durch bestimmtes Sicherheitsverhalten zustande kommen. Um motorische Unruhe zu verstecken, könnten Patienten mit Sozialer Angststörung ihre Hände bewusst unter der Tischplatte halten, um zu vermeiden, dass die Unruhe der Hände bemerkt wird. Dies würde durch unsere 2-dimensionalen Auswertungsmethoden nicht detektiert werden und somit in geringeren Bewegungsindizes resultieren. Ramseyer (2011) zeigte, dass die Bewegungen von Patienten innerhalb der Therapie zunahmten und eine Steigerung in der Bewegungshäufigkeit mit einem besseren Funktionsniveau zusammenhing. Diese Steigerung konnte innerhalb unserer Stichprobe nicht gefunden werden. Grund dafür könnten die störungshomogene Stichprobe, die fehlende Untersuchung von demografischen Variablen als Moderatoren aufgrund der geringen Stichprobengröße sowie die etwas andere Erfassung der Bewegungsparameter sein. Generell kann konstatiert werden, dass Personen mit einem sehr geringen Bewegungsniveau und langsamen Bewegungen besondere Beachtung in der Therapie bekommen sollten, da dies auf ein geringeres Funktionsniveau generell (Ramseyer, 2011), eine Disposition zur sozialen Vermeidung oder ein ausgeprägtes Sicherheitsverhalten hinweisen könnte.

4.2.7 Arbeitsanteil D. Schönherr innerhalb der Publikationen

Die Eigenanteile der Autorin D. Schönherr sind in Bezug auf die Originalarbeiten in Tabelle 12 zusammengefasst.

Tabelle 12 *Eigenanteile innerhalb der Publikationen*

| | Publikation 1 | Publikation 2 | Publikation 3 | Publikation 4 | Publikation 5 |
|---------------------------|--------------------|--------------------|--------------------|--------------------|---------------------|
| | <i>Erstautorin</i> | <i>Erstautorin</i> | <i>Erstautorin</i> | <i>Erstautorin</i> | <i>Zweitautorin</i> |
| Inhaltliche Konzeption | ++ | ++ | + | ++ | |
| Auswahl Methoden | ++++ | ++++ | ++++ | +++ | |
| Programmierung | + | + | + | + | + |

4 Datengrundlage & Inhalte der einzelnen Publikationen

| | | | | | |
|---|------|------|------|------|------|
| Literaturrecherche | ++++ | ++++ | ++++ | ++++ | + |
| Datenerhebung (Synchronie-Maße), Aufbereitung, Imputation | +++ | +++ | +++ | +++ | +++ |
| Datenanalyse | ++++ | ++++ | ++++ | ++++ | |
| Ergebnisdiskussion | ++++ | ++++ | ++++ | ++++ | |
| Verfassen Manuskript | ++++ | ++++ | ++++ | ++++ | |
| Begutachtung & Kommentierung Manuskript | | | | | ++ |
| Überarbeitung des Manuskriptes (auf Basis der Kommentare der Koautoren) & Fertigstellung | ++++ | ++++ | ++++ | ++++ | |
| Zusatzanalysen | | ++++ | ++++ | | ++++ |

Anmerkung: Die Eigenanteile wurden unter Supervision von Dr. Uwe Altmann erbracht. ++++ entspricht vollständiger Bearbeitung, +++ überwiegender Bearbeitung, ++ mehrheitlicher Bearbeitung, + entspricht der Bearbeitung in Teilen.

5 Publikationsübergreifende Diskussion

Aus den hier zusammengefassten Originalarbeiten und Zusatzanalysen können wichtige Schlussfolgerungen hinsichtlich der Validität der Synchronie-Erfassung, der Algorithmen-basierten Definition von Synchronie und der differentiellen Bedeutung der Synchronie im therapeutischen Kontext abgeleitet werden. In den nachfolgenden Absätzen sind diese publikationsübergreifend dargestellt.

5.1 Bedeutung der methodische Arbeiten

Ein großes Problem der bisherigen Synchronisationsforschung besteht darin, dass viele Verfahren zur Erfassung nonverbaler Synchronie angewendet werden, die nicht umfassend auf ihre Validität untersucht wurden. Diese Forschungslücke wurde mittels der ersten

Publikationen (1+2) geschlossen. Die aus der ersten Publikation gewonnenen Erkenntnisse sind in Tabelle 13 noch einmal überblicksartig zusammengefasst.

Tabelle 13 *Überblick über die aus Publikation 1 gewonnenen Erkenntnisse – Validität verschiedener linearer Zeitreihenanalyse-Techniken*

| Verfahren | Konvergente Validität | Prädiktive Validität | Konkurrente Validität | Gemeinsamer Synchronie-Faktor |
|--------------------------|-----------------------|----------------------|-----------------------|-------------------------------|
| CLC | | | | |
| Homogene Parameter | -- | -- | -- | / |
| Heterogene Parameter | - + | -- | -- | + |
| CLR | | | | |
| Homogene Parameter | --- | --- | --- | / |
| Heterogene Parameter | --- | --- | --- | - |
| WCC | | | | |
| Homogene Parameter | +++ | +++ | +++ | / |
| Heterogene Parameter | +++ | +++ | +++ | - |
| WCLC_{S1} | | | | |
| Homogene Parameter | -- | -- | -- | / |
| Heterogene Parameter | ++ | - | ++ | + |
| WCLC_{S2} | | | | |
| Homogene Parameter | +++ | - | +++ | / |
| Heterogene Parameter | +++ | - | +++ | + |
| WCLC_F | | | | |
| Homogene Parameter | +++ | +++ | +++ | / |
| Heterogene Parameter | +++ | +++ | +++ | + |
| WCLR | | | | |
| Homogene Parameter | +++ | +++ | +++ | / |
| Heterogene Parameter | +++ | +++ | +++ | + |

Anmerkungen: CLC = Cross Lagged Correlation, CLR = Cross Lagged Regression, WCLC = Windowed Cross-Lagged Correlation, S1 = Stärkemaß Nr. 1 (vgl. Paulick et al., 2017), S2 = Stärkemaß Nr.2 (vgl. Ramseyer und Tschacher, 2011), F = Frequenzmaß, WCLR = Windowed Cross-Lagged Regression, + bis +++ = gute bis sehr gute Validität, - bis --- = schlechte bis sehr schlechte Validität, / = kein Modell gefunden, prädiktive Validität in Bezug auf die Vorhersage der Veränderungen der interpersonellen Probleme (IIP), konkurrente Validität in Bezug auf die non-lineare Methode Cross Recurrence Quantification Analysis.

Nicht alle Synchronie-Maße bisher verwendeter linear-basierter Zeitreihenanalysemethoden korrelieren miteinander. Getestet wurden die Methoden: Cross Lagged Correlation (CLC), Cross Lagged-Regression (CLR), Windowed Cross-Correlation (WCC), verschiedene Versionen der Windowed Cross-Lagged Correlation (in Anlehnung an Paulick et al. (2017): $WCLC_{S1}$, in Anlehnung an Ramseyer und Tschacher (2011): $WCLC_{S2}$, in Anlehnung an Altmann (2013): $WCLC_F$), und die Windowed Cross-Lagged Regression (WCLR). Die Maße der WCC, $WCLC_{S2}$, WCLC (Stärkemaß, Häufigkeitsmaß) und WCLR hingen hoch miteinander zusammen (homogene Parametereinstellungen: $r = .60 - .99$, heterogene Parametereinstellungen: $r = .55 - .99$). Für die Maße der CLC und $WCLC_{S1}$ waren die verwendeten Parametereinstellungen von zentraler Bedeutung: Bei homogenen Einstellungen konnten keine Zusammenhänge zu anderen Maßen nachgewiesen werden. Für heterogene Parametereinstellungen zeigte die CLC Zusammenhänge mit den Maßen der $WCLC_{S1}$ und $WCLC_{S2}$, für die $WCLC_{S1}$ konnten überdies auch Verbindungen zu allen anderen Maßen (außer CLR) festgestellt werden ($r = .35 - .71$).

Besonders hervorzuheben sind die guten Validitätseigenschaften der WCC, $WCLC_{S2}$, $WCLC_F$ und WCLR. Die WCC scheint eine gute Analysemethode zu sein, wenn lediglich das generelle Niveau der Bewegungssynchronie erforscht werden soll. Die gefensterten Methoden bieten darüber hinaus die Bestimmung anderer Synchronie-Maße (z.B. Pacing/Leading). Weiterhin zeigt die Studie, dass die in dieser Arbeit verwendete Methode $WCLC_F$ eine andere Facette von Synchronie misst als die bisher in der Psychotherapieforschung verwendeten Methoden $WCLC_{S2}$ (Ramseyer & Tschacher, 2011, 2014) und $WCLC_{S1}$ (Paulick et al., 2017; Paulick et al., 2018). Diese sind jedoch demselben übergeordneten Konstrukt der Bewegungssynchronie zuzuordnen und korrelieren bei heterogenen Parametereinstellungen substantiell miteinander (mit $WCLC_{S1}$: $r = .44^*$, mit $WCLC_{S2}$: $r = .55^*$).

Gleichzeitig deutete die Studie darauf hin, dass es von fundamentaler Bedeutung ist, welche Parameter innerhalb der Analyse verwendet werden, ob die Zeitreihen beispielsweise vorher transformiert oder geglättet sind etc. Deshalb wurden in Publikation 2 verschiedene Parameter variiert und an mehreren externen Kriterien (simulierte Zeitreihen, menschlich-ingeschätzte Zeitreihen) die Sinnhaftigkeit der Parameter-Konfigurationen untersucht. In der bisherigen Literatur wurden die Parameter-Einstellungen für inhaltliche Analysen lediglich mit inhaltlichen Aspekten oder der Operationalisierung von Synchronie gerechtfertigt. Die in dieser Arbeit dargestellte Studie zeigte, dass die validen Methoden $WCLC_F$ und WCLR für die Synchronie-Bestimmung geeignet sind. Dafür sollten folgende Parametereinstellungen

verwendet werden: Bandbreite 75 bzw. 125 frames (= 3 bzw. 5 Sekunden), keine Glättung, Logarithmustransformation der Zeitreihen und ein R^2 -Schwellenwert von 0.25. Bei Verwendung dieser Einstellungen erfassen beide Methoden perfektes Echoing und Synchronisationsphänomene, die losgelöst vom generellen Verhaltensstrom sind. Weiterhin liefert die Studie wichtige Hinweise zur Abgrenzung der menschlich empfundenen Synchronie und der durch Algorithmen bestimmten Synchronie: Das menschliche Konstrukt ist semantisch und multidimensional. Die Algorithmen orientieren sich lediglich an den Kurvenverläufen der Motion Energy. Demnach bestehen große Unterschiede zwischen den beiden Konstrukten. Konkretisiert man das durch die $WCLC_F$ gemessene Konstrukt auf Basis der in dieser Arbeit berichteten Untersuchungen zur Validität, so hat dies folgende Eigenschaften: 1) Synchronie ist operationalisiert als ein On-/Off-Prozess, der sich in Synchronisationsintervallen und Nicht-Synchronisationsintervallen manifestiert. 2) Das durch die $WCLC_F$ gemessene Konstrukt operationalisiert die Häufigkeit der Synchronie oder die Synchronie-Stärke in Bezug auf Synchronisationsintervalle. 3) Synchronie ist nicht semantisch, d.h. gleichzeitige Bewegungen werden als zusammengehörig angesehen, wenn sie statistisch (linear) miteinander zusammenhängen. 4) Durch die Einführung des R^2 -Schwellenwertes sind Atemsynchronie und generelle Körperschwankungen, die in Synchronie passieren, spezifisch ausgeschlossen. 5) Synchronie ist valide erfasst (prädiktive, konkurrenente, konvergente Validität) und steht in Zusammenhang mit bisher in der Psychotherapieforschung berichteten Synchronie-Maßen.

5.2 Wichtigkeit der inhaltlichen Arbeiten & Erweiterung des In-Sync Modells

Innerhalb der inhaltlichen Arbeiten wurde auf eine Vielzahl von Arten untersucht, ob es einen Zusammenhang zwischen dem Therapieerfolg und Bewegungssynchronie gibt. Dafür wurden folgende Variablen untersucht, die alle auf verschiedene Art den Therapieerfolg operationalisieren: 1) Therapieabbrüche, 2) der Veränderung interpersonellen Probleme als kategoriale Variable (klinisch bedeutsame Veränderung) sowie als kontinuierliche Variable, 3) Veränderung der Depressivität und 4) die Veränderung der Sozialen Ängste und Sozialer Vermeidung. Außerdem wurden die Bindungsdimensionen Angst und Vermeidung auf einen Zusammenhang untersucht.

Alle Untersuchungen wurden an einer Stichprobe von Patienten mit Sozialer Angststörung durchgeführt. Die Ergebnisse zeigen, dass die Synchronie-Häufigkeit von diesen Patienten über den Therapieverlauf abnimmt, obwohl die Bewegungshäufigkeit statistisch konstant bleibt. Ähnlich wie von Dijk et al. (2018) gezeigt, könnten gerade zu Therapiebeginn vermehrt Synchronisationsphänomene auftreten, um dem Interaktionspartner zu gefallen. Auf

Patienten-Seite steht dies im Zusammenhang mit dem starken Monitoring des eigenen Körpers in Bezug zur Wirkung auf die Außenwelt. Auf Therapeuten-Seite könnte ein erhöhtes Level an Synchronie dafür sprechen, dem Patienten das Gefühl zu geben, ihm zuzuhören und ihn wahrzunehmen. Der Therapeut könnte durch vermehrtes unbewusstes Imitieren ein angenehmes Setting für den Patienten schaffen, was ebenfalls den Aufbau einer therapeutischen Allianz begünstigt. Gerade zu Beginn der Therapie (Probatorik) scheint der Patient mehr Bewegungen zu initiieren, auf die der Therapeut reagiert. Je höher der Anteil an geleiteter Synchronie durch den Patienten ist, desto geringer ist das Risiko eines Therapieabbruchs. Durch fehlende Synchronie auf die vom Patienten initiierten Bewegungen, könnte eine negative Interaktion erlebt werden, die die Befürchtungen des Patienten unterstreicht und zum Therapieabbruch führt. Dies könnte bereits sehr früh in der Therapie ein Indikator für eine ungünstige Passung zwischen Therapeut und Patient oder eine nicht ausreichend etablierte therapeutische Allianz sein.

Im Rahmen der Therapie wird von einem Patienten mit Angststörung zunehmend gelernt, die Aufmerksamkeit auf die eigenen Bedürfnisse zu legen und wieder aktiv zu werden. Nach der Probatorik scheint es besonders günstig, durch den Therapeuten geleitet zu werden. Dadurch werden die therapeutische Allianz erhöht sowie interpersonellen Probleme und die Bindungsangst der Patienten reduziert. Bei gleichem Bewegungsniveau, werden am Ende der Therapie nicht mehr so viele Synchronisationsphänomene beobachtet. Der Patient gewinnt neues Selbstbewusstsein und handelt nicht stets überangepasst (Paulick et al., 2018). Für den Therapeuten ist ein Rückgang der Synchronie ebenfalls erstrebenswert, um den Patienten eigenverantwortlich aus der Therapie zu entlassen.

In Bezug auf die Therapierichtung zeigen die Befunde eine starke Abhängigkeit der Ergebnisse vom betrachteten Leading/Pacing-Anteil sowie dem jeweiligen Outcome. Ein positiver Einfluss der Bewegungssynchronie auf den Therapieoutcome konnte sowohl in der kognitiven Verhaltenstherapie als auch in der psychodynamisch-orientierten Therapie gezeigt werden. Dieser Einfluss fällt jedoch unterschiedlich aus. In der PDT ist es gerade in der Probatorik sinnvoll dem Patienten zu Beginn der Therapie viel Raum und die Möglichkeit zur Initiierung von Bewegungssynchronie zu geben, wohingegen der Therapeut auf den Patienten reagieren sollte. In der KVT scheint dies eine geringere Bedeutung zu haben. Betrachtet man jedoch die Therapiestunden nach der Probatorik, so ist es in der KVT eher förderlich, wenn der Patient den initiierten Bewegungen des Therapeuten folgt. Dies hat einen positiven Einfluss auf die interpersonellen Muster des Patienten zum Ende der Therapie. Dieser Befund könnte

durch den Kontext der Sitzungen plausibilisiert werden: In der PDT werden in der Probatorik Beziehungs-Episoden-Interviews durchgeführt, in denen der Patient von bisherigen schwierigen Beziehungen berichtet. Um einen Abbruch zu vermeiden, könnte es deshalb auf nonverbaler Ebene entscheidend sein, zu synchronisieren. In der KVT werden zu diesem Zeitpunkt detaillierte Situationsanalysen durchgeführt und ein auf den Patienten passendes Störungsmodell erarbeitet. Die Sitzungen laufen sehr strukturiert ab – eine Struktur, die vom Therapeuten gegeben wird. Möglich wäre, dass je mehr Struktur der Therapeut innerhalb dieser Sitzungen vermittelt, desto positiver der Einfluss auf die interpersonellen Probleme des Patienten ist. Weiterhin ist Therapeuten-geleitete Synchronie auch in der nächsten Phase der kognitiven Vorbereitung auf die Exposition sinnvoll (ca. Sitzung 8 bis 14). Sie führt dazu, dass die interpersonellen Probleme zum Ende der Therapie geringer sind und der Patient eine geringere Bindungsangst zeigt. Jedoch sind die hier gegebenen Erklärungsansätze nur grob aus den Sitzungsmanualen abgeleitet. Es wäre für zukünftige Forschung von großer Bedeutung Bewegungssynchronie genauer auf den Kontext (d.h. die Interventionen oder Themen innerhalb der Therapiesitzungen) zu beziehen, in dem sie gezeigt wird.

Durch die Analysen in den Originalarbeiten wurde festgestellt, dass Outcome-Maße wie Soziale Angst/Vermeidung und Depression nicht oder kaum mit Bewegungssynchronie in Zusammenhang stehen. Dies steht im Widerspruch zu den von Ramseyer und Tschacher (2011) postulierten Befunden, eines Zusammenhangs verschiedener störungsspezifischer Outcomes. Eine Erklärung hierfür könnte zum einen das andere Synchronie-Maß geben: Im Gegensatz zu Ramseyer und Tschacher (2011) wurde in dieser Arbeit nicht die Synchronie-Stärke der Gesamtinteraktion, sondern die Synchronisationshäufigkeit betrachtet. Zum anderen unterschieden sich die untersuchten Outcome-Maße (Ramseyer und Tschacher, 2011: Brief Symptom Inventory), die Stichprobe (bei Ramseyer und Tschacher, 2011 eine störungsheterogene Gruppe nur gleichgeschlechtlicher Dyaden), die Therapierichtung (bei Ramseyer und Tschacher, 2011: naturalistische KVT) sowie der Untersuchungszeitraum der Bewegungssynchronie (Ramseyer und Tschacher, 2011: erstes Therapiedrittel). Die Befunde aus den hier dargestellten Originalarbeiten weisen darauf hin, dass Bewegungssynchronie eher eine proximale Variable zu interpersonellen Konstrukten (z.B. Bindung, interpersonelle Probleme) ist. Weiterhin konnte durch eine Zusatzanalyse gezeigt werden, dass Bewegungssynchronie nicht gleichzusetzen mit der therapeutischen Allianz ist, diese jedoch vorhersagen kann. Die Effekte auf die Abbruchrate waren robust unter Hinzunahme der verbal eingeschätzten therapeutischen Beziehung.

Die Merkmale des Synchronie-Konstruktes, das in dieser Arbeit Verwendung fand, sind nachfolgend noch einmal zusammenfassend dargestellt:

- interpersonelles Konstrukt, kein Maß der reinen Symptomatik
- \neq therapeutische Allianz, misst etwas anderes/mehr
- On-/Off-Prozess, der sich in Synchronisationsintervallen und Nicht-Synchronisationsintervallen manifestiert
- nicht semantisch, d.h. gleichzeitige Bewegungen werden als zusammengehörig angesehen, wenn sie statistisch (linear) miteinander zusammenhängen
- Atemsynchronie und generelle Körperschwankungen in Synchronie sind ausgeschlossen
- Maße: Häufigkeit der Synchronie oder die Synchronie-Stärke in Bezug auf Synchronisationsintervalle
- Nachgewiesene prädiktive, konkurrente, konvergente Validität, Zusammenhang mit bisher in der Psychotherapieforschung berichteten Synchronie-Maßen

Die Befunde sprechen für eine Erweiterung des von Koole und Tschacher (2016) postulierten In-Sync-Modells der Psychotherapie. Eine mögliche Erweiterung, die die Befunde der bisherigen Arbeit integriert, ist in Abbildung 17 vorgeschlagen und soll im Folgenden näher beschrieben werden.

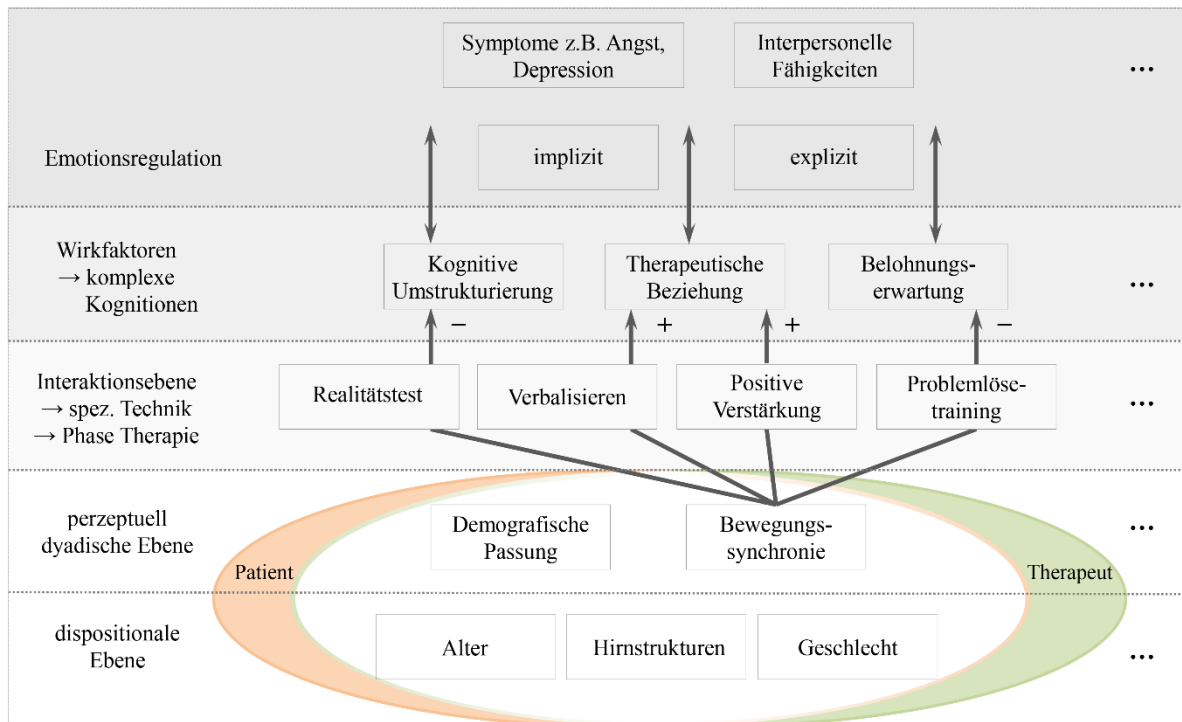


Abbildung 17 *Erweitertes In-Sync-Modell der Psychotherapie*, + entspricht einem positiven Einfluss, - entspricht einem negativen Einfluss, ... deuten an, dass das Modell horizontal durch weitere Aspekte erweitert werden kann

Die unterste Ebene des Modells stellt die *dispositionale Ebene* dar. Damit sind Merkmale des Therapeuten und des Patienten gemeint, die die zwischen beiden stattfindende Interaktion beeinflussen können. In Bezug auf die Betrachtung von Bewegungssynchronie sind dies beispielsweise das Geschlecht, die hirnanatomischen Grundstrukturen (in Bezug auf Synchronie z.B. Spiegelneurone, Areale des temporal-parietalen Übergangs, Areale des Belohnungssystems) oder die bisher gemachten Erfahrungen des Patienten und des Therapeuten. Auf Patienten-Seite ist weiterhin die psychische Störung als eine Determinante zu nennen. Zahlreiche Studien konnten bisher die förderlichen und hemmenden Einflussfaktoren auf Synchronie zusammenfassen, beispielsweise prosoziale Orientierung (Lumsden et al., 2012), das Geschlecht (La France & Ickes, 1981), die psychische Störung (Paulick et al., 2018) und der Einfluss der Aktivierung der bei der Imitation förderlichen Hirnstrukturen (Nummenmaa et al., 2018). Die drei Punkte auf jeder Ebene der Abbildung deuten an, dass die dargestellten Bestandteile der Ebenen lediglich Beispiele darstellen und sie die Ebenen nicht allumfassend charakterisieren. Dabei ist weiterhin zu erwähnen, dass dabei Kognitionen, Verhaltensweisen, Emotionen sowie körperliche Empfindungen eine Rolle spielen. Der Körper stellt dabei im Sinne des Embodiment-Ansatzes ein Mediator zwischen

der externen Welt und dem inneren Erleben sowie dem Verhalten dar des Individuums dar (Tschacher & Storch, 2010).

Die Ebene darüber ist die *perzeptuelle dyadische Ebene*. Diese Ebene umfasst alle auch von außen wahrnehmbaren Eigenschaften der dyadischen Interaktion. Dazu gehören zum Beispiel die demografische Passung und die gezeigte Bewegungssynchronie. Diese Variablen sind von der dispositionalen Ebene beeinflusst und beschreiben die Interaktion von Patient und Therapeut. Die Variablen sind von zentraler Bedeutung, da sie das Geschehen innerhalb der Psychotherapie maßgeblich beeinflussen. So zeigten Wintersteen, Mensinger, und Diamond (2005), dass ein Matching hinsichtlich des Geschlechtes bei Jugendlichen zu höheren Allianz-Ratings sowie einer geringeren Abbruchrate führte. Innerhalb der hier dargestellten Originalarbeiten wurde gezeigt, dass Bewegungssynchronie eine höhere Wichtigkeit hat, wenn gemischt-geschlechtliche Dyaden oder ein ungleicher Bildungsstand vorliegen. Deshalb sind beide Variablen auch als Beispiele auf einer Ebene angeordnet. Da die Bewegungssynchronie in gleichgeschlechtlichen Dyaden und bei gleichem Bildungsstand nur eine untergeordnete Rolle spielt, scheinen beide Elemente kompensatorisch zu sein. Für die Praxis soll dies keineswegs heißen, dass Therapeuten und Patienten möglichst gut aufeinander abgestimmt werden sollten, es offeriert vielmehr die Möglichkeit der nonverbalen Kommunikation als Kompensationsmöglichkeit für demografische Unterschiede. Weiterhin könnte die Variable Bewegungssynchronie differenzierter betrachtet werden. Es wurde gezeigt, dass verschiedene Facetten von Bewegungssynchronie unterschieden werden können, die ebenfalls differentiell mit verschiedenen Outcome-Variablen zusammenhängen könnten. Das Leading/Pacing-Verhalten macht darüber hinaus einen wichtigen Unterschied. Ramseyer und Tschacher (2011) stellten fest, dass das Leading des Patienten eher mit der Selbstwirksamkeitseinschätzung und das Leading des Therapeuten eher mit der therapeutischen Allianz assoziiert ist. Reich et al. (2014) zeigten, dass ein erhöhtes Folgen der paraverbalen Merkmale des Patienten zu einer erhöhten Depressivität am Ende der Therapie führen kann. Dies unterstreicht weiterhin eine wichtige Erweiterungsmöglichkeit des Modells, die Modalität. Betrachtet man Forschungsergebnisse verschiedener Modalitäten, beispielweise Mimik, Bewegungen und Stimme, so kann ebenfalls die Modalität ein wichtiger Moderator sein. Beispielsweise fanden Rasting und Beutel (2005) für Synchronie von Gesichtsausdrücken einen negativen Zusammenhang, wohingegen Ramseyer und Tschacher (2011) einen positiven Einfluss von Bewegungssynchronie auf den Outcome nachwiesen. In dem hier beschriebenen Modell

werden demnach lediglich Ideen für Bewegungssynchronie dargestellt. Für andere Modalitäten könnte Synchronie einen komplett anderen Einfluss haben.

Die darüber liegende Ebene ist die *Ebene der Interaktion*, die die angewandte spezifische Technik der Psychotherapie sowie den Kontext der Interaktion einschließt. Als in der Psychotherapie hauptsächlich vorherrschender Kontext werden in diesem Modell spezifische Techniken der Psychotherapie beschrieben. Wie im Theoretischen Teil jedoch bereits verdeutlicht wurde, hat eine Vielzahl an Kontexten einen Einfluss darauf, wie Synchronie wirkt. Mögliche spezifische Techniken der Psychotherapie sind beispielsweise die Realitätstestung, das Verbalisieren, positive Verstärkungen sowie das Problemlösetraining (Pfammatter et al., 2012). Diese spezifischen Techniken wirken durch bestimmte, auf der Ebene über ihnen zusammengefasste *Wirkfaktoren* der Psychotherapie. Diese Wirkfaktoren sind beispielsweise die kognitive Umstrukturierung, die Therapiebeziehung sowie die Besserungserwartung (Pfammatter et al., 2012). Anhand dieser Wirkfaktoren wird eine Änderung der *impliziten und expliziten Emotionsregulation*, der Therapieerfolg des Patienten erreicht. Nimmt man auf perzeptueller Ebene Bewegungssynchronie wahr, so hat dies je nach Kontext, d.h. je nach spezifischer Technik verschiedene Auswirkungen auf die dahinterliegenden Wirkfaktoren. In dem Modell wird demnach die perzeptuell dyadische Ebene als ein wichtiger Moderator zwischen der Interaktionsebene und der Wirkfaktorebene beschrieben. Die Pfeile und daran befindliche Zeichen geben eine mögliche Wirkungsweise vor. Wird beispielsweise Bewegungssynchronie während des Verbalisierens oder positiver Verstärkung gezeigt, so wirkt sich dies positiv auf die Therapiebeziehung und so auf den Therapieoutcome des Patienten aus. Wird Bewegungssynchronie jedoch innerhalb einer Realitätstestung oder während des Problemlösetrainings gezeigt, so kann sie keinen Einfluss oder sogar einen negativen Einfluss haben. Gestützt wird diese Modellannahme durch zahlreiche Studien, die die Kontextabhängigkeit der Wirkung von Synchronie unterstreichen. Der Einfluss von nonverbaler Synchronie konnte der Studie von Leander et al. (2011) nur beobachtet werden, wenn die Interaktion ein Affiliationsziel hatte. Lag jedoch ein professionelles Ziel vor (z.B. gemeinsames Lösen einer Aufgabe), hatte nonverbaler Synchronie keinen Einfluss. Auch scheint nicht die generelle Höhe der gezeigten nonverbalen Synchronie ausschlaggebend, sondern ob und wie nonverbale Synchronie erwartet wird (Hale & Hamilton, 2016). Wird eine professionelle Aufgabe bearbeitet und keine Synchronie erwartet, so hat diese einen negativen Einfluss (Leander et al., 2012). Dies könnte auf die Psychotherapie übertragen werden: Ist der dahinterliegende Wirkfaktor die Etablierung einer

Therapiebeziehung, so ist Synchronie zuträglich; wird dem entgegengesetzt ein eher professionelles Ziel verfolgt (z.B. Realitätsprüfung, Perspektivenwechsel → kognitive Umstrukturierung) so ist es plausibel, dass Synchronie keinen oder einen negativen Einfluss haben könnte. Hirnanatomisch konnte von Santiesteban et al. (2012) gezeigt werden, dass gerade für den Perspektivenwechsel das Verschwimmen der Self-Other-Grenze, welches durch nonverbale Synchronie induziert wird, nicht zuträglich ist und sich bei Imitations-Inhibition die Fähigkeit zum Perspektivenwechsel erhöht. Auch Bänninger-Huber und Huber (2017) berichten einen positiven Effekt von Synchronie auf die therapeutische Beziehung, wohingegen für eine Symptomverbesserung eine Konfliktspannung benötigt wird, die nicht förderlich von Synchronie beeinflusst wird. Diese Ergebnisse beruhen jedoch auf der Untersuchung von gleichzeitigem Lächeln. Da angenommen werden kann, dass die Modalität einen kritischen Einfluss auf die Wirkung von nonverbaler Synchronie hat, sollten diese Hypothesen spezifisch für Bewegungssynchronie untersucht werden.

Die hier berichteten Originalarbeiten lassen sich in das postulierte Modell folgendermaßen integrieren: Für die früheren Messzeitpunkte konnten Einflüsse der Bewegungssynchronie auf die interpersonellen Fähigkeiten, die Bindungsangst und das Abbruchverhalten gezeigt werden. Geht man über Sitzung 8 hinaus, wurden diese Zusammenhänge jedoch nicht mehr gefunden. Es kann angenommen werden, dass im Laufe einer Psychotherapie die Etablierung der therapeutischen Allianz eine weniger wichtige Rolle einnimmt. Nach der Etablierung dieser zu Beginn, können in der Therapie andere Ziele bearbeitet werden. Im Vergleich dazu, nehmen Problemzentrierte Techniken (Exposition, kognitive Umstrukturierung, Durcharbeiten des ZBKT) zu. Deshalb könnte es nach dem vorgestellten Modell plausibel sein, dass in diesen Phasen der Therapie keine direkten Zusammenhänge zwischen Maßen des Therapieerfolgs und Bewegungssynchronie gefunden wurden. Die Modellannahme der differentiellen Wirkung von Bewegungssynchronie impliziert außerdem, dass nicht alle Therapieerfolgsmaße gleichermaßen von Bewegungssynchronie beeinflusst werden. Demnach sind Maße mit stärkerem interpersonellem Fokus (interpersonelle Probleme, Bindung) eher von Bewegungssynchronie beeinflusst als Maße, die direkt mit Problem-orientierten Strategien zusammenhängen (z.B. Soziale Angst, Depressionen). Weiterhin sollten verschiedene Facetten von Bewegungssynchronie sowie verschiedene Pacing-/Leading-Anteile untersucht werden, da diese unterschiedlich mit verschiedenen Outcome-Variablen assoziiert zu sein scheinen.

5.3 Stärken und Einschränkungen

Eine wichtige Stärke der vorliegenden Arbeiten ist, dass über den kompletten Forschungsprozess sichergestellt wurde, dass die verwendeten Methoden dazu geeignet sind, Synchronie unter Erfüllung aller psychologischen Gütekriterien zu messen. Durch die automatisierte Erfassung der Körperbewegungen durch die MEA wurde ein reliables und valides Verfahren benutzt (Altmann, 2010; Paxton & Dale, 2013b). Da die bisherige Forschung wenig Validierungsarbeiten zur Berechnung der Synchronie aus Zeitreihen liefert, wurden die Methoden erst hinsichtlich zahlreicher Validitätskriterien geprüft und die Parameterkonfiguration bestimmt, die die besten Ergebnisse für die Messung der Synchronie liefert. Um die Validität zu prüfen, wurden nicht nur Korrelationen verwendet, sondern auch explorative und konfirmatorische Faktorenanalysen benutzt. Weiterhin fanden ordinal logistische Regressionen Anwendung, sodass von einer sehr guten statistischen Absicherung der Schlussfolgerungen gesprochen werden kann. Bei wiederholter Durchführung resultieren durch den Algorithmus die exakt gleichen Synchronisationsintervalle. Durch die öffentliche Zugänglichkeit jeglicher Skripte auf GitHub (<https://github.com/10101-00001/MEA> bzw. https://github.com/10101-00001/sync_ident) und die detaillierte Beschreibung der verwendeten Parameter in den Publikationen, können die Ergebnisse repliziert werden und andere Arbeitsgruppen können transparent nachvollziehen, welche Schritte zu den Synchronie-Indizes geführt haben. Weiterhin erfüllen sie so das Gütekriterium der Objektivität. Die Methode zur Synchronie-Bestimmung selbst ist damit objektiv, reliabel, transparent und valide.

Die inhaltlichen Analysen, in denen der generierte Synchronie-Index verwendet wurde, sind von hoher methodischer Qualität. Die Stichprobe von $N = 267$ Patienten einer homogenen Stichprobe übersteigt die bisherigen Stichprobengrößen der Synchronie-Forschung (siehe Tabelle 5). Weiterhin wurden für die fehlende Randomisierung der Versuchspersonen in den Therapiebedingungen stets Propensity-Scores verwendet, um die Gruppen miteinander vergleichen zu können. Patienten und Therapeuten in der beschriebenen Studie waren verblindet hinsichtlich der Untersuchungsfrage (Videos aufgenommen zwischen 2007-2009, Forschungsfrage entwickelt: 2011). Da Therapeuteneffekte in der Psychotherapieforschung hinsichtlich vieler Kriterien gut belegt sind, wurden außerdem stets hierarchische lineare Modelle verwendet, um für den Therapeuteneffekt zu kontrollieren. Weiterhin wurde auf mehrere Analysestrategien zurückgegriffen (z.B. Publikation 3: Verwendung von logistischen Regression & Überlebenszeitanalyse, von Drop-out als Kriterium sowie der klinisch

signifikanten Veränderung). Dieses Vorgehen zeigt, dass die gewonnenen Erkenntnisse unabhängig von der verwendeten Methodik sind.

Ein weiterer wichtiger Punkt ist das Konstrukt der Synchronie. Durch die vorliegende Arbeit kann die Definition des Konstruktes Synchronie konkretisiert werden. Durch die Nähe zu interpersonellen Konstrukten kann angenommen werden, dass Bewegungssynchronie nicht nur ein nonverbales Maß der Symptomatik des Patienten ist, sondern eher einen dyadischen Prozess abbildet.

Jedoch sind auf Grundlage der methodischen Analyseverfahren einige Einschränkungen nicht vermeidbar. Die MEA aggregiert Bewegungen zu einem Gesamtbewegungsindex, sodass nicht differenziert werden kann, ob die Bewegung dadurch entstanden ist, dass sich der Interaktionspartner vor oder zurückgelehnt hat. Diese Bewegungen sind semantisch jedoch völlig verschieden: Ein Vorlehnen könnte Annäherung, ein Zurücklehnen Ablehnung zeigen. Bewegungen, die von verschiedenen Extremitäten ausgehen, können nicht unterschieden werden. In Bezug auf das Aufnahmesetting ist zu konstatieren, dass die Videos nicht mit identischen Kameras aufgezeichnet wurden, was eventuelle Verzerrungen hervorgerufen haben könnte. Diese Verzerrungen sollten aufgrund folgender Gründe eher gering sein: Die Synchronie wurde stets in Bezug auf eine Dyade aufgezeichnet und mit einer identischen Kamera betrachtet. Anschließend wurden alle Videos in ein einheitliches Format gebracht und außerdem eine ROI-Standardisierung zur besseren Vergleichbarkeit innerhalb eines Videos und zwischen den Videos einführt. Jegliche Auffälligkeiten in den Synchronie-Indizes wurden stets durch manuelle Sichtungen untersucht.

Bei der Validierung der zeitreihenanalytischen Methoden wurden 600 verschiedene Parametereinstellungen überprüft. Diese Art der Studie und Anzahl der variierten Parameter ist bisher in der Forschung einmalig. Jedoch hätte der maximale time lag ebenfalls variiert werden sollen. Die durchgeführte Zusatzanalyse zum Vergleich zweier verschiedener time lags lässt vermuten, dass geringe Veränderungen dieses Parameters für die Synchronie-Maße weniger entscheidend sind und eher eine Rolle bei der Bestimmung von time lag relativen Variablen (mittlerer time lag, Social Present) spielen. Eine feingliedrigere Untersuchung der Parameter-Levels könnte darüber hinaus sinnvoll sein, um zu quantifizieren, in welchem Range die Parameter zu identischen Ergebnissen führen und ab welchen Schwellenwerten von einer substantiellen Parameteränderung ausgegangen werden kann. Gerade Interaktionen zwischen den Einstellungen (verschiedene Transformationen angewendet, dann die Bandbreite betrachtet) könnten ebenfalls zu interessanten Ergebnissen führen. Wenn für jede

Zeitreihenanalyse-Methode optimale Parameter bestimmt sind, sollten die Untersuchungen zur konvergenten Validität erneut durchgeführt werden, da sich durch die Wahl geeigneter Parameter ein eindimensionales Synchronie-Konstrukt zeigen könnte.

Für die inhaltlichen Analysen wurden vier Messzeitpunkte verwendet. Lag das Video der genauen Sitzung nicht vor oder konnte nicht verwendet werden, so wurden Videos der angrenzenden Sitzungen verwendet. Aus den Schlussfolgerungen dieser Arbeit und des erweiterten In-Sync Modells wird jedoch ersichtlich, dass damit Sitzungen miteinander verglichen wurden, die eventuell keinen vergleichbaren Inhalt hatten. In Zukunft empfiehlt es sich ein standardisiertes Setting oder qualitative Ratings des Gesprächsinhalts zu verwenden, um die Grundlagen der Wirkung von Bewegungssynchronie zu untersuchen.

Die Befunde für die Soziale Angststörung wurden ohne Referenz einer gesunden Stichprobe interpretiert. Ohne diese Referenz sind die Interpretationen des Ausmaßes der Synchronie nicht robust. Studiendesigns sollten eine gesunde Vergleichsgruppe einschließen, um die Synchronie-Veränderungen in Referenz zu dieser zu setzen.

Die Unterschiede zwischen der manualisierten PDT und KVT könnten neben den verschiedenen Therapiekonzepten auch durch besondere Charakteristika der Stichprobe erklärt werden. Die Selektion der Videos zur Anwendung der MEA resultierte beispielsweise darin, dass die Therapeuten der PDT Bedingung ein höheres Alter zeigten als in der KVT Bedingung. Die KVT Therapeuten bewegten sich generell mehr als die PDT Therapeuten.

Weiterhin wurde in dieser Arbeit lediglich eine Art der nonverbalen Synchronie, die Bewegungssynchronie betrachtet. Bisherige Forschung liefert Hinweise darauf, dass interpersonelle Synchronie und deren Einfluss nicht unabhängig von der untersuchten Verhaltensmodalität ist. Dahingehend sollte das In-Sync Modell auch hinsichtlich anderer Modalitäten (z.B. mimische Synchronie, Synchronie paraverbalen Merkmale,...) erweitert werden, um die differentiellen Effekte der Synchronie verschiedener Modalitäten in ein Modell zu integrieren. Innerhalb des hier beschriebenen Projektes TIMPATHIN wurde eine Fortsetzung des Projektes genehmigt, die an dem hier beschriebenen Datensatz den Einfluss stimmlicher Synchronie (Stimmfrequenz, Sprechrate) untersuchen wird.

Weiterhin könnte man in den Analysen für das generelle Bewegungsniveau von Patient und Therapeut kontrollieren. Bewegen sich Patient und Therapeut häufig, ist die Detektion von Synchronisationsintervallen einfacher, da Bewegungen stattfinden. Im Gegensatz dazu würde in einer Interaktion ohne jegliche Bewegung auch keine Bewegungssynchronie gefunden werden, obwohl sich der Therapeut an das Verhalten des Patienten angepasst haben könnte. Da

Paulick et al. (2018) keinen Einfluss der generellen Bewegungshäufigkeit auf die Bewegungssynchronie zeigen konnten, schlossen wir es als Prädiktor für unsere Synchronie-Analysen aus.

5.4 Implikationen für die Psychotherapie

Die durch die Arbeit gewonnen Erkenntnisse deuten auf mehrere Anwendungsmöglichkeiten im Psychotherapiekontext hin. Sie könnten für die Psychotherapeutische Ausbildung und als diagnostisches oder Feedback-Instrument innerhalb der Psychotherapie genutzt werden.

Nonverbale Kommunikation nimmt einen beträchtlichen Anteil unserer generellen Kommunikation ein. Deshalb ist es wichtig, Psychotherapeuten für den Einsatz und das Wahrnehmen nonverbaler Signale zu schulen. Lovkvist (2013) schlägt vor, dass Therapeuten auf eine Abnahme der Muskelspannung beim Patienten achten und die eigenen Bewegungen innerhalb der Therapie erhöhen sollten. Damit könnten mehr Möglichkeiten gegeben sein, dass sich der Patient auf der Bewegungsebene mit dem Therapeuten synchronisiert. Es sollte jedoch definitiv davon abgesehen werden, den Patienten nachzuahmen, da die positiven Effekte von nonverbaler Synchronie lediglich auftreten, wenn das Verhalten unbewusst, authentisch und unbemerkt gezeigt wird (Bailenson & Yee, 2005; Van Baaren et al., 2009). Studien im Rahmen der Spiegelneuron-Forschung legen nahe, dass das Spiegelneuron-System abhängig von den gemachten Erfahrungen ist (Jacoboni, 2009). Demnach könnte es möglich sein, dass durch vermehrtes Üben und Lenken der Aufmerksamkeit auf das nonverbale Verhalten auch der implizite und unbewusste Umgang mit Patienten auf nonverbaler Ebene beeinflusst werden kann. Weiterhin kann es innerhalb der Psychotherapie von enormen Nutzen sein, nonverbales wie verbales Verhalten wahrzunehmen, um Inkongruenzen der Kommunikationsaspekte zu bemerken, die wichtige Indikatoren für unverarbeitete Konflikte oder mentale Repräsentationen der Patienten liefern können.

Werden nonverbale Merkmale expliziter in die Psychotherapeutische Ausbildung einbezogen, beispielsweise durch Rollenspiele oder das Üben mit virtuellen Avataren, so scheint es möglich, innerhalb der Psychotherapiestunden das nonverbale Verhalten des Patienten als diagnostisches Instrument oder zum Feedback zu benutzen. Bemerkte man zum Beispiel ein sehr geringes Level an Bewegungssynchronie, so könnte dies in der Therapie explizit angesprochen werden. Damit könnte die Hemmschwelle des Patienten eventuelle Probleme zu berichten, herabgesetzt werden und eine transparente Arbeitsatmosphäre geschaffen werden. Dafür ist es jedoch zwingend erforderlich, zu untersuchen, wie sich

Synchronie innerhalb der Sitzungen manifestiert: Ist Synchronie beispielsweise gerade nach dem erfolgreichen Reparieren einer *alliance rupture* größer oder kündigen sich solche Events in dem nonverbalen Verhalten an?

Sind optimale Synchronie-Werte für verschiedene Settings und Störungsgruppen untersucht und konsentiert, so könnte die Synchronie-Messung ebenfalls an die Stelle der Prä-Treatment, Post-Treatment sowie den Verlaufsmessungen innerhalb der Psychotherapiesitzungen treten. Im Gegensatz zu dem Einsatz eines Fragebogens ist die Erfassung von Synchronie nicht auf Selbst- oder Fremdauskünfte angewiesen und bietet ein Messinstrument ohne Verzerrungen durch *differential item functioning* oder soziale Erwünschtheit. Damit wäre es denkbar, bereits durch die in der Probatorik gezeigte Synchronie einen Indikator für die Passung zwischen Therapeut und Patient zu bekommen, was Abbrüche wegen einer fehlenden Sympathie der Interaktionspartner verhindern könnte. Innerhalb des Prozesses der Therapie könnte eine zu geringe Synchronie ein Warnsignal für den Therapeuten sein, das anzeigt, dass es nötig ist, noch weiter in die therapeutische Beziehung zu investieren.

Ein anderer Aspekt ist, dass die gewonnenen Erkenntnisse auch dazu benutzt werden können, Bewegungssynchronie aktiv im Therapieprozess als eine spezifische Technik einzusetzen. Hale und Hamilton (2016) schlagen vor, die Effekte von Synchronie, im Rahmen von virtual Mimicry oder einem Mimicry Priming aktiv für die Therapie zu benutzen. Gerade bei Patienten mit Sozialer Angststörung scheint dies eine sinnvolle Ergänzung zur kognitiven Verhaltenstherapie zu sein (Del-Monte et al., 2016). Auch in der Behandlung von Autismus scheint es sinnvoll gleichzeitige Bewegungen auszuführen (wie beispielsweise beim Tanzen), da diese Empathie-fördernd sind (Behrends et al., 2012; Koehne et al., 2016). Im therapeutischen Kontext sollte jedoch vor allem erst untersucht werden, ob der Grad der Bewegungssynchronie aktiv manipuliert werden kann oder ob eine Schulung in nonverbalen Merkmalen dazu führt, dass auch unbewusst „besser“ nonverbal synchronisiert wird. Die Schwierigkeit der aktiven Manipulation von Synchronie liegt vor allem darin, dass der Therapeut weiterhin als authentisch wahrgenommen wird.

Um direkte Anwendungsmöglichkeiten für die psychotherapeutische Praxis abzuleiten, ist es jedoch zu früh. Um konkrete Handlungsanweisungen stellen zu können, müssen deutlich mehr Experimente durchgeführt werden. Das erweiterte In-Sync Modell der Psychotherapie, das im Abschnitt 5.2 vorgestellt wurde, zeigt mögliche Integrationsmöglichkeiten verschiedener Ergebnisse, jedoch sollten diese vorzugsweise im standardisierten Setting untersucht werden.

5.5 Implikationen für die Forschung

Aus den hier beschriebenen methodischen Arbeiten lassen sich konkrete Handlungsanweisungen für die zukünftige Forschung ableiten. Es sollte beim Einsatz der Methoden deutlich darauf geachtet werden, dass das intendierte Konstrukt auch tatsächlich gemessen wird. Eine Augenscheinvalidität ist dabei nicht genug, um die Güte eines Verfahrens zu beweisen. Es sollten systematische Validierungsstudien durchgeführt werden, die das gemessene Konstrukt besser greifbar machen und die Algorithmen in Zusammenhang zu bereits bestehenden Algorithmen stellen. Dafür wäre es äußerst vorteilhaft, eine gemeinsame Datenbank von Videos und Zeitreihen zu erstellen. Besteht für die Zeitreihen und Videos Konsens hinsichtlich der dargestellten Synchronie, so können sie verwendet werden, um neue Algorithmen zu testen, bestehende Algorithmen an dieser Datenbank weiterzuentwickeln und zu einer Methodenstandardisierung zu gelangen. Ein konsertiertes Konstrukt der Synchronie könnte auch im Bereich des *Machine-Learnings* dazu führen, durch Optimierungsfunktionen geeignete Parameter für eine optimale Synchronie-Detektion zu identifizieren. Sind weder Parameter noch Konstrukt explizit festgelegt, gibt es zu viele Freiheitsgrade für den Einsatz dieser Methoden. Weiterhin ist es unabdingbar, die verwendeten Parameter in Publikationen klar zu benennen, um anderen Arbeitsgruppen zu ermöglichen, die Ergebnisse zu replizieren. Eine Veröffentlichung der verwendeten Algorithmen ist demnach ebenfalls anzustreben.

Herausforderungen auf methodischer Ebene sind für die Zukunft vor allem das Setting, die Messmethode sowie das multimodale Messen (vgl. Cornejo et al., 2017). Es gibt zahlreiche Synchronisationsexperimente im Experimentalsetting. Diese geben jedoch wenig Aufschluss über die direkten Wirkmechanismen von Synchronie innerhalb menschlicher natürlicher Interaktionen. Ist das Setting jedoch komplett unstandardisiert, so kann der Effekt von Synchronie ebenfalls nicht genau entschlüsselt werden. Die Herausforderung zukünftiger Forschung liegt darin, ein Setting zu etablieren, welches standardisiert ist, jedoch auch dem natürlichen Kontext entspricht. Ein Beispiel hierfür könnte die Untersuchung von Bewegungssynchronie innerhalb von standardisierten Interviews oder während der Durchführung einer Situationsanalyse sein. Weiterhin ist es erstrebenswert, nicht nur einen Parameter nonverbaler Synchronie zu messen. In dieser Arbeit wurde sich auf Bewegungssynchronie konzentriert, jedoch gibt es noch weitere Arten von Synchronie wie beispielsweise die der Stimme, Mimik oder Hirnströme, die nicht gemessen wurden. Um nonverbale Synchronie als komplettes Konstrukt verstehen zu können, ist es für zukünftige Forschung wünschenswert, auch andere Modalitäten zu betrachten.

Auf inhaltlicher Ebene sind die Herausforderungen vor allem die Manipulation der Synchronie, differentielle Wirkungen sowie Moderatoren (vgl. Hale & Hamilton, 2016). Zukünftige Forschung sollte vor allem untersuchen, wie Synchronie aktiv und bewusst im Therapieprozess benutzt werden kann. Da es sich bei der Bewegungssynchronie um ein implizites Maß handelt, ist dies schwer kontrollier- und umsetzbar. Des Weiteren müssen standardisierte Settings verwendet werden, um die differentielle Wirkung von Synchronie zu verstehen. Dabei sollten wichtige demografische und klinische Variablen stets auf ihren moderierenden Einfluss untersucht werden.

Aus den hier gewonnenen Erkenntnissen lassen sich auf inhaltlicher Ebene spezifische Forschungsfragen für die Zukunft ableiten. Da Synchronie in verschiedenen Therapiephasen unterschiedlich wirkt und die Leading-/Pacing-Anteile eine wichtige Rolle spielen, sollte die Veränderung der Synchronie (Leading-/Pacing-Anteile, time lag Maße) innerhalb der Sitzungen sowie über den Therapieprozess hinweg in Relation zu den in den Sitzungen besprochenen Themen oder Interventionen kleinschrittig untersucht werden. Anhand der Verläufe könnten günstige und nicht günstige Verlaufsmuster identifiziert werden, die den Therapieprozess umfassender charakterisieren. Weiterhin ist es erstrebenswert, Synchronie mit mehreren Facetten (Stärke der Synchronie, Häufigkeit der Synchronie) und multimodal zu messen (Mimik, Stimme, Physiologie, Bewegungen, Wort-Synchronie, usw.), um der Multidimensionalität nonverbalen Verhaltens gerecht zu werden. In einem nächsten Schritt ist es erstrebenswert, die praktische Umsetzbarkeit und Einsetzbarkeit der Erkenntnisse zu überprüfen. Dabei sollte aus den Erkenntnissen ein Trainingsprogramm (z.B. Schulungen mit anschließender Behandlung von Avatar-Patienten und Supervision) für Psychotherapeuten etabliert werden und in einer randomisiert-kontrollierten Studie untersucht werden, ob trainierte Psychotherapeuten bessere Therapieergebnisse erzielen. Da momentan davon ausgegangen wird, dass Synchronisationsphänomene ihre positive Wirkung nur haben, wenn sie unbewusst passieren, könnte damit überprüft werden, ob es überhaupt möglich wäre, die Fähigkeit zu synchronisieren explizit zu schulen.

Referenzen

- Abbott, K. A., Kocovski, N. L., & Obhi, S. S. (2018). Impact of Social Anxiety on Behavioral Mimicry During a Social Interaction With a Confederate. *Journal of Social and Clinical Psychology, 37*(1), 22-43.
- Acarturk, C., Cuijpers, P., Van Straten, A., & De Graaf, R. (2009). Psychological treatment of social anxiety disorder: a meta-analysis. *Psychological Medicine, 39*(2), 241-254.
- Alden, L. E., & Taylor, C. T. (2004). Interpersonal processes in social phobia. *Clinical Psychology Review, 24*(7), 857-882. doi:10.1016/j.cpr.2004.07.006
- Allsop, J. S., Vaitkus, T., Marie, D., & Miles, L. K. (2016). Coordination and collective performance: Cooperative goals boost interpersonal synchrony and task outcomes. *Frontiers in Psychology, 7*. doi:10.3389/fpsyg.2016.01462
- Altmann, U. (2010). Interrater-Reliabilität= 1 in Videostudien? Automatisierte Erhebung von Nonverbalität in einem Experiment zur Kooperation von Schülern (Automated coding of nonverbal behavior in an experiment on the cooperation of students). *Erziehungswissenschaftliche Forschung–nachhaltige Bildung. Beiträge zur, 5*, 261-267.
- Altmann, U. (2011). Investigation of movement synchrony using windowed crosslagged regression. *Analysis of Verbal and Nonverbal Communication and Enactment: The Processing Issue, 344-354*. doi:10.1007/978-3-642-25775-9_31
- Altmann, U. (2013). *Synchronisation nonverbalen Verhaltens: Weiterentwicklung und Anwendung zeitreihenanalytischer Identifikationsverfahren. [Synchronization of nonverbal behavior: Development and Application of time series analysis methods]*: Springer-Verlag.
- Altmann, U., Schoenherr, D., Paulick, J., Knitter, L., Worrack, S., Schiefele, A.-K., . . . Strauss, B. (2019). Introduction, Practical Guide, and Validation Study for Measuring Body Movements Using Motion Energy Analysis. *submitted*.
- American Psychiatric Association. (2018). *Diagnostisches und statistisches manual psychischer Störungen–DSM-5®*: Hogrefe Verlag.
- Anscombe, F. J. (1948). The transformation of Poisson, binomial and negative-binomial data. *Biometrika, 35*(3/4), 246-254. doi:10.2307/2332343
- Ashton-James, C. E., & Chartrand, T. L. (2009). Social cues for creativity: The impact of behavioral mimicry on convergent and divergent thinking. *Journal of Experimental Social Psychology, 45*(4), 1036-1040.
- Bailenson, J. N., & Yee, N. (2005). Digital chameleons: Automatic assimilation of nonverbal gestures in immersive virtual environments. *Psychological Science, 16*(10), 814-819. doi:10.1111/j.1467-9280.2005.01619.x
- Bandelow, B., & Wedekind, D. (2014). Soziale Phobie. *Der Nervenarzt, 85*(5), 635-647.
- Bandelow, B., Wiltink, J., Alpers, G. W., Benecke, C., Deckert, J., Eckhardt-Henn, A., . . . M.E., B. (2014). S3-Leitlinie. Behandlung von Angststörungen. Retrieved from https://www.awmf.org/uploads/tx_szleitlinien/051-028k_S3_Angstst%C3%B6rungen_2014-05_1_01.pdf
- Bänniger-Huber, E., & Huber, E. (2017). Nonverbal Processes in Psychotherapeutic Interaction. *Journal of Psychology and Psychotherapie Research, 4* (1), 1-8. doi:10.12974/2313-1047.2017.04.01.1
- Beebe, B., Lachmann, F. M., Markese, S., Buck, K. A., Bahrack, L. E., Chen, H., . . . Jaffe, J. (2012). On the origins of disorganized attachment and internal working models: Paper II. An empirical microanalysis of 4-month mother–infant interaction. *Psychoanalytic dialogues, 22*(3), 352-374.

- Behrends, A., Müller, S., & Dziobek, I. (2012). Moving in and out of synchrony: A concept for a new intervention fostering empathy through interactional movement and dance. *The Arts in Psychotherapy, 39*(2), 107-116.
- Benecke, C., & Krause, R. (2005a). Facial affective relationship offers of patients with panic disorder. *Psychotherapy Research, 15*(3), 178-187.
- Benecke, C., & Krause, R. (2005b). Initiales mimisch-affektives Verhalten und Behandlungszufriedenheit in der Psychotherapie von Patientinnen mit Panikstörungen/ Initial Affective Facial Behaviour and Outcome Satisfaction in the Psychotherapy of Patients with Panic Disorder. *Zeitschrift für Psychosomatische Medizin und Psychotherapie, 51*(4), 346-359. doi:10.13109/zptm.2005.51.4.346
- Bernieri, F. J., Davis, J. M., Rosenthal, R., & Knee, C. R. (1994). Interactional synchrony and rapport: Measuring synchrony in displays devoid of sound and facial affect. *Personality and Social Psychology Bulletin, 20*(3), 303-311.
- Bernieri, F. J., Reznick, J. S., & Rosenthal, R. (1988). Synchrony, pseudosynchrony, and dissynchrony: Measuring the entrainment process in mother-infant interactions. *Journal of Personality and Social Psychology, 54*(2), 243. doi:10.1037/0022-3514.54.2.243
- Boker, S. M., & Rotondo, J. L. (2002). Symmetry building and symmetry breaking in synchronized movement. *Mirror neurons and the evolution of brain and language, 42*, 163.
- Boker, S. M., Rotondo, J. L., Xu, M., & King, K. (2002). Windowed cross-correlation and peak picking for the analysis of variability in the association between behavioral time series. *Psychological Methods, 7*(3), 338. doi:10.1037/1082-989X.7.3.338
- Box, G. E., & Cox, D. R. (1964). An analysis of transformations. *Journal of the Royal Statistical Society. Series B (Methodological), 211-252*.
- Brandt, T., Kugler, G., Schniepp, R., Wuehr, M., & Huppert, D. (2015). Acrophobia impairs visual exploration and balance during standing and walking. *Annals of the New York Academy of Sciences, 1343*(1), 37-48.
- Brockwell, P. J., & Davis, R. A. (2013). *Time series: theory and methods*: Springer Science & Business Media.
- Burgoon, J. K., Guerrero, L. K., & Floyd, K. (2009). *Nonverbal Communication*. Boston: MA: Allyn and Bacon.
- Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological bulletin, 56*(2), 81.
- Chartrand, T. L., & Bargh, J. A. (1999). The chameleon effect: The perception-behavior link and social interaction. *Journal of Personality and Social Psychology, 76*(6), 893. doi:10.1037/0022-3514.76.6.893
- Chartrand, T. L., & Lakin, J. L. (2013). The antecedents and consequences of human behavioral mimicry. *Annual review of psychology, 64*, 285-308.
- Chetouani, M., Delaherche, E., Dumas, G., & Cohen, D. (2017). 15 Interpersonal Synchrony: From Social Perception to Social Interaction. *Social Signal Processing, 202*.
- Clark, D. M. (2005). A cognitive perspective on social phobia. *The essential handbook of social anxiety for clinicians, 193-218*.
- Clark, D. M., Ehlers, A., McManus, F., Hackmann, A., Fennell, M., Campbell, H., . . . Louis, B. (2003). Cognitive therapy versus fluoxetine in generalized social phobia: a randomized placebo-controlled trial. *Journal of Consulting and Clinical Psychology, 71*(6), 1058.
- Clark, D. M., & Wells, A. (1995). A cognitive model of social phobia. *Social phobia: Diagnosis, assessment, and treatment, 41*(68), 00022-00023.

- Coco, M. I., & Dale, R. (2014). Cross-recurrence quantification analysis of categorical and continuous time series: an R package. *Frontiers in Psychology, 5*. doi:10.3389/fpsyg.2014.00510
- Codrons, E., Bernardi, N. F., Vandoni, M., & Bernardi, L. (2014). Spontaneous group synchronization of movements and respiratory rhythms. *PLoS One, 9*(9), e107538.
- Condon, W. S., & Ogston, W. D. (1966). Sound film analysis of normal and pathological behavior patterns. *The Journal of nervous and mental disease, 143*(4), 338-347.
- Cornejo, C., Cuadros, Z., Morales, R., & Paredes, J. (2017). Interpersonal Coordination: Methods, Achievements, and Challenges. *Front Psychol, 8*, 1685. doi:10.3389/fpsyg.2017.01685
- Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological bulletin, 52*(4), 281. doi:10.1037/h0040957
- Dahan, A., Noy, L., Hart, Y., Mayo, A., & Alon, U. (2016). Exit from synchrony in joint improvised motion. *PLoS One, 11*(10), e0160747.
- Dales, S., & Jerry, P. (2008). Attachment, affect regulation and mutual synchrony in adult psychotherapy. *American Journal of Psychotherapy, 62*(3), 283-312. doi:10.1176/appi.psychotherapy.2008.62.3.283
- Dean, R. T., & Dunsmuir, W. T. (2016). Dangers and uses of cross-correlation in analyzing time series in perception, performance, movement, and neuroscience: The importance of constructing transfer function autoregressive models. *Behavior Research Methods, 1-20*. doi:10.3758/s13428-015-0611-2
- Del-Monte, J., Raffard, S., Capdevielle, D., Schmidt, R. C., Varlet, M., Salesse, R. N., . . . Marin, L. (2016). 15 Impact of mental disorders on social motor behaviours. *Interpersonal Coordination and Performance in Social Systems, 211*.
- Del-Monte, J., Raffard, S., Salesse, R. N., Marin, L., Schmidt, R. C., Varlet, M., . . . Capdevielle, D. (2013). Nonverbal expressive behaviour in schizophrenia and social phobia. *Psychiatry research, 210*(1), 29-35.
- Delaherche, E., & Chetouani, M. (2010). *Multimodal coordination: exploring relevant features and measures*. Paper presented at the Proceedings of the 2nd international workshop on Social signal processing.
- Delaherche, E., Chetouani, M., Mahdhaoui, A., Saint-Georges, C., Viaux, S., & Cohen, D. (2012). Interpersonal synchrony: A survey of evaluation methods across disciplines. *IEEE Transactions on Affective Computing, 3*(3), 349-365. doi:10.1109/T-AFFC.2012.12
- Dijk, C., Fischer, A. H., Morina, N., van Eeuwijk, C., & van Kleef, G. A. (2018). Effects of Social Anxiety on Emotional Mimicry and Contagion: Feeling Negative, but Smiling Politely. *Journal of Nonverbal Behavior, 42*(1), 81-99.
- Duffy, K. A., & Chartrand, T. L. (2015). Mimicry: causes and consequences. *Current Opinion in Behavioral Sciences, 3*, 112-116.
- Dumas, G., Laroche, J., & Lehmann, A. (2014). Your body, my body, our coupling moves our bodies. *Frontiers in human neuroscience, 8*, 1004.
- Duran, N. D., & Fusaroli, R. (2017). Conversing with a devil's advocate: Interpersonal coordination in deception and disagreement. *PLoS One, 12*(6), e0178140. doi:10.1371/journal.pone.0178140
- Ekman, P. (1965). Communication through nonverbal behavior: A source of information about an interpersonal relationship. In S. S. Tomkins & C. E. Izard (Eds.), *Affect, cognition and personality* (pp. 390-442). New York: Springer.
- Eskildsen, A., Hougaard, E., & Rosenberg, N. K. (2010). Pre-treatment patient variables as predictors of drop-out and treatment outcome in cognitive behavioural therapy for

- social phobia: A systematic review. *Nordic journal of psychiatry*, 64(2), 94-105. doi:10.3109/08039480903426929
- Farabee, D. J., Ramsey, S. L., & Cole, S. G. (1993). Social anxiety and speaker gaze in a persuasive atmosphere. *Journal of Research in Personality*, 27(4), 365-376.
- Fedoroff, I. C., & Taylor, S. (2001). Psychological and pharmacological treatments of social phobia: a meta-analysis. *Journal of clinical psychopharmacology*, 21(3), 311-324.
- Feldman, R. (2012). Parent–infant synchrony: A biobehavioral model of mutual influences in the formation of affiliative bonds. *Monographs of the Society for Research in Child Development*, 77(2), 42-51.
- Feldman, R. (2015). Mutual influences between child emotion regulation and parent–child reciprocity support development across the first 10 years of life: implications for developmental psychopathology. *Development and psychopathology*, 27(4pt1), 1007-1023.
- Feldman, R., Greenbaum, C. W., & Yirmiya, N. (1999). Mother–infant affect synchrony as an antecedent of the emergence of self-control. *Developmental Psychology*, 35(1), 223.
- Feldman, R., Magori-Cohen, R., Galili, G., Singer, M., & Louzoun, Y. (2011). Mother and infant coordinate heart rhythms through episodes of interaction synchrony. *Infant Behavior and Development*, 34(4), 569-577.
- Feniger-Schaal, R., Noy, L., Hart, Y., Koren-Karie, N., Mayo, A. E., & Alon, U. (2016). Would you like to play together? Adults' attachment and the mirror game. *Attach Hum Dev*, 18(1), 33-45. doi:10.1080/14616734.2015.1109677
- Fogel, A. (1993). Two principles of communication: Co-regulation and framing. *New perspectives in early communicative development*, 9-22.
- Fraenkel, D. L. (1983). The relationship of empathy in movement to synchrony, echoing, and empathy in verbal interactions. *American Journal of Dance Therapy*, 6(1), 31-48. doi:10.1007/bf02579518
- Galbusera, L., Finn, M. T., & Fuchs, T. (2016). Interactional synchrony and negative symptoms: An outcome study of body-oriented psychotherapy for schizophrenia. *Psychother Res*, 1-13. doi:10.1080/10503307.2016.1216624
- Gilboa-Schechtman, E., & Shachar-Lavie, I. (2013). More than a face: a unified theoretical perspective on nonverbal social cue processing in social anxiety. *Frontiers in human neuroscience*, 7, 904.
- Gottman, J. M., & Ringland, J. T. (1981). The analysis of dominance and bidirectionality in social development. *Child Development*, 393-412. doi:10.2307/1129157
- Grammer, K., Honda, M., Juette, A., & Schmitt, A. (1999). Fuzziness of nonverbal courtship communication unblurred by motion energy detection. *Journal of Personality and Social Psychology*, 77(3), 487. doi:10.1037/0022-3514.77.3.487
- Greene, M. G., Adelman, R. D., Friedmann, E., & Charon, R. (1994). Older patient satisfaction with communication during an initial medical encounter. *Social Science & Medicine*, 38(9), 1279-1288. doi:10.1016/0277-9536(94)90191-0
- Gueguen, N. (2011). The mimicker is a mirror of myself: Impact of mimicking on self-consciousness and social anxiety. *Social Behavior and Personality: an international journal*, 39(6), 725-728.
- Hale, J., & Hamilton, A. F. d. C. (2016). Cognitive mechanisms for responding to mimicry from others. *Neuroscience & Biobehavioral Reviews*, 63, 106-123.
- Hall, N. R., Millings, A., & Bouças, S. B. (2012). Adult Attachment Orientation and Implicit Behavioral Mimicry. *Journal of Nonverbal Behavior*, 36(4), 235-247. doi:10.1007/s10919-012-0136-7

- Hans, E., & Hiller, W. (2013). A meta-analysis of nonrandomized effectiveness studies on outpatient cognitive behavioral therapy for adult anxiety disorders. *Clinical Psychology Review, 33*(8), 954-964. doi:10.1037/a0025749
- Hart, Y., Czerniak, E., Karnieli-Miller, O., Mayo, A. E., Ziv, A., Biegon, A., . . . Alon, U. (2016). Automated Video Analysis of Non-verbal Communication in a Medical Setting. *Frontiers in Psychology, 7*(1130). doi:10.3389/fpsyg.2016.01130
- Hautzinger, M., & Thies, E. (2009). *Klinische Psychologie: Psychische Störungen kompakt: mit Online-Materialien*: Beltz.
- Håvås, E., Svartberg, M., & Ulvenes, P. (2015). Attuning to the unspoken: The relationship between therapist nonverbal attunement and attachment security in adult psychotherapy. *Psychoanalytic psychology, 32*(2), 235. doi:10.1037/a0038517
- Heerey, E. A., & Kring, A. M. (2007). Interpersonal consequences of social anxiety. *Journal of abnormal psychology, 116*(1), 125.
- Ho, D. E., Imai, K., King, G., & Stuart, E. A. (2011). MatchIt: nonparametric preprocessing for parametric causal inference. *Journal of Statistical Software, 42*(8), 1-28. doi:10.18637/jss.v042.i08
- Hofmann, S. G. (2007). Cognitive factors that maintain social anxiety disorder: A comprehensive model and its treatment implications. *Cognitive behaviour therapy, 36*(4), 193-209.
- Hogeveen, J., Chartrand, T. L., & Obhi, S. S. (2014). Social mimicry enhances mu-suppression during action observation. *Cerebral Cortex, 25*(8), 2076-2082.
- Horley, K., Williams, L. M., Gonsalvez, C., & Gordon, E. (2003). Social phobics do not see eye to eye:: A visual scanpath study of emotional expression processing. *Journal of anxiety disorders, 17*(1), 33-44.
- Hove, M. J., & Risen, J. L. (2009). It's all in the timing: Interpersonal synchrony increases affiliation. *Social Cognition, 27*(6), 949.
- Iacoboni, M. (2009). Imitation, empathy, and mirror neurons. *Annual review of psychology, 60*, 653-670. doi:10.1146/annurev.psych.60.110707.163604
- Isabella, R. A., Belsky, J., & von Eye, A. (1989). Origins of infant-mother attachment: An examination of interactional synchrony during the infant's first year. *Developmental Psychology, 25*(1), 12.
- Jacob, R. G., Lilienfeld, S. O., Furman, J. M., Durrant, J. D., & Turner, S. M. (1989). Panic disorder with vestibular dysfunction: further clinical observations and description of space and motion phobic stimuli. *Journal of anxiety disorders, 3*(2), 117-130.
- Johnston, L. (2002). Behavioral mimicry and stigmatization. *Social Cognition, 20*(1), 18-35.
- Kang, S.-H., & Gratch, J. (2012). *Understanding the nonverbal behavior of socially anxious people during intimate self-disclosure*. Paper presented at the International Conference on Intelligent Virtual Agents.
- Kimura, M., & Daibo, I. (2006). Interactional synchrony in conversations about emotional episodes: A measurement by "the between participants pseudosynchrony experimental paradigm". *Journal of Nonverbal Behavior, 30*(3), 115-126.
- Kinreich, S., Djalovski, A., Kraus, L., Louzoun, Y., & Feldman, R. (2017). Brain-to-Brain Synchrony during Naturalistic Social Interactions. *Sci Rep, 7*(1), 17060. doi:10.1038/s41598-017-17339-5
- Kleinbub, J. R. (2017). State of the Art of Interpersonal Physiology in Psychotherapy: A Systematic Review. *Frontiers in Psychology, 8*. doi:10.3389/fpsyg.2017.02053
- Kleinbub, J. R., Messina, I., Bordin, D., Voci, A., Calvo, V., Sambin, M., & Palmieri, A. (2012). *Synchronization of skin conductance levels in therapeutic dyads* (Vol. 86).
- Koehne, S., Hatri, A., Cacioppo, J. T., & Dziobek, I. (2016). Perceived interpersonal synchrony increases empathy: insights from autism spectrum disorder. *Cognition, 146*, 8-15.

- Koole, S. L., & Tschacher, W. (2016). Synchrony in Psychotherapy: A Review and an Integrative Framework for the Therapeutic Alliance. *Frontiers in Psychology*, 7(862). doi:10.3389/fpsyg.2016.00862
- Koss, T., & Rosenthal, R. (1997). Interactional synchrony, positivity and patient satisfaction in the physician-patient relationship. *Medical Care*, 35(11), 1158-1163.
- Kot, S., & Kulesza, W. (2016). The chameleon as a leech: The costs of mimicry for the mimicker. *Polish Psychological Bulletin*, 47(1), 131-135.
- Kouzakova, M., van Baaren, R., & van Knippenberg, A. (2010). Lack of behavioral imitation in human interactions enhances salivary cortisol levels. *Hormones and behavior*, 57(4), 421-426. doi:10.1016/j.yhbeh.2010.01.011
- Kreyenbrink, I., Joraschky, P., Konstantinidis, I., Neumann, N., & Lausberg, H. (2017). Nonverbales Verhalten von Patienten mit sozialen Phobien und ihren Therapeuten in psychodynamischen Psychotherapien (Teilprojekt SOPHO-NET). *Zeitschrift für Psychosomatische Medizin und Psychotherapie*, 63(3), 297-313.
- Kühn, S., Müller, B. C., van Baaren, R. B., Wietzker, A., Dijksterhuis, A., & Brass, M. (2010). Why do I like you when you behave like me? Neural mechanisms mediating positive consequences of observing someone being imitated. *Social Neuroscience*, 5(4), 384-392.
- Kulesza, W., Dolinski, D., & Wicher, P. (2016). Knowing that you mimic me: the link between mimicry, awareness and liking. *Social Influence*, 11(1), 68-74.
- Kupper, Z., Ramseyer, F., Hoffmann, H., Kalbermatten, S., & Tschacher, W. (2010). Video-based quantification of body movement during social interaction indicates the severity of negative symptoms in patients with schizophrenia. *Schizophrenia Research*, 121(1), 90-100. doi:10.1016/j.schres.2010.03.032
- Kupper, Z., Ramseyer, F., Hoffmann, H., & Tschacher, W. (2015). Nonverbal Synchrony in Social Interactions of Patients with Schizophrenia Indicates Socio-Communicative Deficits. *PLoS One*, 10(12), e0145882. doi:10.1371/journal.pone.0145882
- La France, M., Hecht, M. A., & Paluck, E. L. (2003). The contingent smile: a meta-analysis of sex differences in smiling. *Psychological bulletin*, 129(2), 305.
- La France, M., & Ickes, W. (1981). Posture mirroring and interactional involvement: Sex and sex typing effects. *Journal of Nonverbal Behavior*, 5(3), 139-154. doi:10.1007/BF00986131
- Lakens, D. (2010). Movement synchrony and perceived entitativity. *Journal of Experimental Social Psychology*, 46(5), 701-708.
- Lakens, D., & Stel, M. (2011). If they move in sync, they must feel in sync: Movement synchrony leads to attributions of rapport and entitativity. *Social Cognition*, 29(1), 1.
- Lakin, J. L., & Chartrand, T. L. (2003). Using nonconscious behavioral mimicry to create affiliation and rapport. *Psychological Science*, 14(4), 334-339.
- Laukka, P., Linnman, C., Åhs, F., Pissioti, A., Frans, Ö., Faria, V., . . . Furmark, T. (2008). In a nervous voice: Acoustic analysis and perception of anxiety in social phobics' speech. *Journal of Nonverbal Behavior*, 32(4), 195-214.
- Leander, N. P., Chartrand, T. L., & Bargh, J. A. (2012). You give me the chills: Embodied reactions to inappropriate amounts of behavioral mimicry. *Psychological Science*, 23(7), 772-779.
- Leander, N. P., Chartrand, T. L., & Wood, W. (2011). Mind your mannerisms: Behavioral mimicry elicits stereotype conformity. *Journal of Experimental Social Psychology*, 47(1), 195-201.
- Leichsenring, F., Salzer, S., Beutel, M. E., Herpertz, S., Hiller, W., Hoyer, J., . . . Poehlmann, K. (2013). Psychodynamic therapy and cognitive-behavioral therapy in social anxiety

- disorder: a multicenter randomized controlled trial. *American Journal of Psychiatry*. doi:10.1176/appi.ajp.2013.12081125
- Leighton, J., Bird, G., Orsini, C., & Heyes, C. (2010). Social attitudes modulate automatic imitation. *Journal of Experimental Social Psychology*, *46*(6), 905-910.
- Levitan, M., Crippa, J., Bruno, L., Pastore, D., Freire, R., Arrais, K., . . . Nardi, A. (2012). Postural balance in patients with social anxiety disorder. *Brazilian Journal of Medical and Biological Research*, *45*(1), 38-42.
- Lindsey, E. W., & Caldera, Y. M. (2015). Shared affect and dyadic synchrony among secure and insecure parent–toddler dyads. *Infant and Child Development*, *24*(4), 394-413.
- Lopes, F. L., Azevedo, T. M., Imbiriba, L. A., Freire, R. C., Valença, A. M., Caldirola, D., . . . Nardi, A. E. (2009). Freezing reaction in panic disorder patients associated with anticipatory anxiety. *Depression and anxiety*, *26*(10), 917-921.
- Louwerse, M. M., Dale, R., Bard, E. G., & Jeuniaux, P. (2012). Behavior matching in multimodal communication is synchronized. *Cognitive science*, *36*(8), 1404-1426. doi:10.1111/j.1551-6709.2012.01269.x
- Lovkvist, M. (2013). *Somatic empathic resonance: Subjective and intersubjective experiences of the psychotherapeutic dyad*. The Chicago School of Professional Psychology.
- Luborsky, L. (1995). Einführung in die analytische Psychotherapie: ein Lehrbuch. 2. Aufl. Göttingen, Zürich: Vandenhoeck und Ruprecht.
- Lumsden, J., Miles, L. K., & Macrae, C. N. (2014). Sync or sink? Interpersonal synchrony impacts self-esteem. *Frontiers in Psychology*, *5*, 1064.
- Lumsden, J., Miles, L. K., Richardson, M. J., Smith, C. A., & Macrae, C. N. (2012). Who syncs? Social motives and interpersonal coordination. *Journal of Experimental Social Psychology*, *48*(3), 746-751.
- Lutz, W., Schiefele, A.-K., Wucherpennig, F., Rubel, J., & Stulz, N. (2016). Clinical effectiveness of cognitive behavioral therapy for depression in routine care: A propensity score based comparison between randomized controlled trials and clinical practice. *Journal of Affective Disorders*, *189*, 150-158. doi:10.1016/j.jad.2015.08.072
- Machin, A. J., & Dunbar, R. I. (2011). The brain opioid theory of social attachment: a review of the evidence. *Behaviour*, *148*(9-10), 985-1025.
- Macrae, C. N., Duffy, O. K., Miles, L. K., & Lawrence, J. (2008). A case of hand waving: Action synchrony and person perception. *Cognition*, *109*(1), 152-156.
- Maddux, W. W., Mullen, E., & Galinsky, A. D. (2008). Chameleons bake bigger pies and take bigger pieces: Strategic behavioral mimicry facilitates negotiation outcomes. *Journal of Experimental Social Psychology*, *44*(2), 461-468.
- Marsh, K. L., Richardson, M. J., & Schmidt, R. C. (2009). Social connection through joint action and interpersonal coordination. *Topics in Cognitive Science*, *1*(2), 320-339.
- Marwan, N., & Kurths, J. (2002). Nonlinear analysis of bivariate data with cross recurrence plots. *Physics Letters A*, *302*(5), 299-307.
- Marwan, N., Romano, M. C., Thiel, M., & Kurths, J. (2007). Recurrence plots for the analysis of complex systems. *Physics reports*, *438*(5), 237-329. doi:10.1016/j.physrep.2006.11.001
- Miles, L. K., Griffiths, J. L., Richardson, M. J., & Macrae, C. N. (2009). Too late to coordinate: Contextual influences on behavioral synchrony. *European Journal of Social Psychology*, n/a-n/a. doi:10.1002/ejsp.721
- Miles, L. K., Lumsden, J., Richardson, M. J., & Macrae, C. N. (2011). Do birds of a feather move together? Group membership and behavioral synchrony. *Experimental Brain Research*, *211*(3-4), 495-503.

- Mogan, R., Fischer, R., & Bulbulia, J. A. (2017). To be in synchrony or not? A meta-analysis of synchrony's effects on behavior, perception, cognition and affect. *Journal of Experimental Social Psychology*, *72*, 13-20. doi:10.1016/j.jesp.2017.03.009
- Moulder, R. G., Boker, S. M., Ramseyer, F., & Tschacher, W. (2018). Determining synchrony between behavioral time series: An application of surrogate data generation for establishing falsifiable null-hypotheses. *Psychol Methods*. doi:10.1037/met0000172
- Nagaoka, C., & Komori, M. (2008). Body movement synchrony in psychotherapeutic counseling: A study using the video-based quantification method. *IEICE transactions on information and systems*, *91*(6), 1634-1640. doi:10.1093/ietisy/e91-d.6.1634
- Namy, L. L., Nygaard, L. C., & Sauerteig, D. (2002). Gender differences in vocal accommodation: The role of perception. *Journal of Language and Social Psychology*, *21*(4), 422-432. doi:10.1177/026192702237958
- Novembre, G., Knoblich, G., Dunne, L., & Keller, P. E. (2017). Interpersonal synchrony enhanced through 20 Hz phase-coupled dual brain stimulation. *Social Cognitive and Affective Neuroscience*, *12*(4), 662-670. doi:10.1093/scan/nsw172
- Nummenmaa, L., Lahnakoski, J., & Glerean, E. (2018). Sharing the social world via intersubject neural synchronization. *Current opinion in psychology*.
- Oullier, O., De Guzman, G. C., Jantzen, K. J., Lagarde, J., & Scott Kelso, J. (2008). Social coordination dynamics: Measuring human bonding. *Social Neuroscience*, *3*(2), 178-192.
- Palmieri, A., Kleinbub, J. R., Calvo, V., Benelli, E., Messina, I., Sambin, M., & Voci, A. (2018). Attachment-security prime effect on skin-conductance synchronization in psychotherapists: An empirical study. *Journal of Counseling Psychology*.
- Paulick, J., Deisenhofer, A.-K., Ramseyer, F., Tschacher, W., Boyle, K., Rubel, J., & Lutz, W. (2017). Nonverbal Synchrony: A new approach to better understand psychotherapeutic processes and drop-out. *Journal of Psychotherapy Integration*. doi:10.1037/int0000099
- Paulick, J., Rubel, J., Deisenhofer, A.-K., Schwartz, B., Thielemann, D., Altmann, U., . . . Lutz, W. (2018). Diagnostic features of nonverbal synchrony in psychotherapy: Comparing depression and anxiety. *Cognitive Therapy and Research*, Advanced online publication. doi:10.1007/s10608-018-9914-9
- Paxton, A. (2015). *Coordination: Theoretical, Methodological, and Experimental Perspectives*. University of California, Merced.
- Paxton, A., & Dale, R. (2013a). Argument disrupts interpersonal synchrony. *The Quarterly Journal of Experimental Psychology*, *66*(11), 2092-2102. doi:10.1080/17470218.2013.853089
- Paxton, A., & Dale, R. (2013b). Frame-differencing methods for measuring bodily synchrony in conversation. *Behavior Research Methods*, *45*(2), 329-343. doi: 10.3758/s13428-012-0249-2
- Paxton, A., & Dale, R. (2017). Interpersonal Movement Synchrony Responds to High- and Low-Level Conversational Constraints. *Front Psychol*, *8*, 1135. doi:10.3389/fpsyg.2017.01135
- Perna, G., Dario, A., Caldirola, D., Stefania, B., Cesarani, A., & Bellodi, L. (2001). Panic disorder: the role of the balance system. *Journal of Psychiatric Research*, *35*(5), 279-286.
- Pfammatter, M., Junghan, U. M., & Tschacher, W. (2012). Allgemeine Wirkfaktoren der Psychotherapie: Konzepte, Widersprüche und eine Synthese. *Psychotherapie*, *17*(1), 17-31.
- Ramseyer, F. (2008). *Synchronisation nonverbaler Interaktion in der Psychotherapie*. Selbstverl.

- Ramseyer, F. (2011). Nonverbal synchrony in psychotherapy: embodiment at the level of the dyad. In W. Tschacher & C. Bergomi (Eds.), *The Implications of Embodiment: Cognition and Communication* (pp. 193-207). Exeter, United Kingdom: Imprint Academic.
- Ramseyer, F., & Tschacher, W. (2008). Synchrony in dyadic psychotherapy sessions. *Simultaneity: Temporal structures and observer perspectives*, 329-347.
- Ramseyer, F., & Tschacher, W. (2010). Nonverbal synchrony or random coincidence? How to tell the difference. *Development of multimodal interfaces: active listening and synchrony* (pp. 182-196): Springer.
- Ramseyer, F., & Tschacher, W. (2011). Nonverbal Synchrony in Psychotherapy: Coordinated Body Movement Reflects Relationship Quality and Outcome. *Journal of Consulting and Clinical Psychology*, 79(3), 284-295. doi:10.1037/a0023419
- Ramseyer, F., & Tschacher, W. (2014). Nonverbal synchrony of head- and body-movement in psychotherapy: different signals have different associations with outcome. *Frontiers in Psychology*, 5(979). doi:10.3389/fpsyg.2014.00979
- Ramseyer, F., & Tschacher, W. (2016). Movement Coordination in Psychotherapy: Synchrony of Hand Movements is Associated with Session Outcome. A Single-Case Study. *Nonlinear dynamics, psychology, and life sciences*, 20(2), 145-166.
- Rasting, M., & Beutel, M. E. (2005). Dyadic affective interactive patterns in the intake interview as a predictor of outcome. *Psychotherapy Research*, 15(3), 188-193. doi:10.1080/10503300512331335039
- Reich, C. M., Berman, J. S., Dale, R., & Levitt, H. M. (2014). Vocal synchrony in psychotherapy. *Journal of Social and Clinical Psychology*, 33(5), 481-494.
- Rennung, M., & Göritz, A. S. (2016). Prosocial Consequences of Interpersonal Synchrony. *Zeitschrift für Psychologie*.
- Richardson, M. J., Marsh, K. L., Isenhower, R. W., Goodman, J. R., & Schmidt, R. C. (2007). Rocking together: Dynamics of intentional and unintentional interpersonal coordination. *Human movement science*, 26(6), 867-891.
- Rizzolatti, G., Camarda, R., Fogassi, L., Gentilucci, M., Luppino, G., & Matelli, M. (1988). Functional organization of inferior area 6 in the macaque monkey. *Experimental Brain Research*, 71(3), 491-507.
- Rizzolatti, G., & Craighero, L. (2004). The mirror-neuron system. *Annu. Rev. Neurosci.*, 27, 169-192. doi:10.1146/annurev.neuro.27.070203.144230
- Ruscio, A. M., Brown, T. A., Chiu, W. T., Sareen, J., Stein, M. B., & Kessler, R. C. (2008). Social fears and social phobia in the USA: results from the National Comorbidity Survey Replication. *Psychological Medicine*, 38(1), 15-28.
- Santesteban, I., White, S., Cook, J., Gilbert, S. J., Heyes, C., & Bird, G. (2012). Training social cognition: from imitation to theory of mind. *Cognition*, 122(2), 228-235.
- Scharfstein, L. A., Beidel, D. C., Sims, V. K., & Rendon Finnell, L. (2011). Social skills deficits and vocal characteristics of children with social phobia or Asperger's disorder: a comparative study. *J Abnorm Child Psychol*, 39(6), 865-875. doi:10.1007/s10802-011-9498-2
- Schmidt, R., Christianson, N., Carello, C., & Baron, R. (1994). Effects of social and physical variables on between-person visual coordination. *Ecological Psychology*, 6(3), 159-183.
- Schmidt, R., Morr, S., Fitzpatrick, P., & Richardson, M. J. (2012). Measuring the dynamics of interactional synchrony. *Journal of Nonverbal Behavior*, 36(4), 263-279.
- Schneier, F. R., Rodebaugh, T. L., Blanco, C., Lewin, H., & Liebowitz, M. R. (2011). Fear and avoidance of eye contact in social anxiety disorder. *Comprehensive psychiatry*, 52(1), 81-87.

- Schönbrodt, F. D., & Perugini, M. (2013). At what sample size do correlations stabilize? *Journal of Research in Personality, 47*(5), 609-612. doi:10.1016/j.jrp.2013.05.009
- Shockley, K. (2005). Cross recurrence quantification of interpersonal postural activity. *Tutorials in contemporary nonlinear methods for the behavioral sciences, 142-177.*
- Shockley, K., Richardson, D. C., & Dale, R. (2009). Conversation and coordinative structures. *Topics in Cognitive Science, 1*(2), 305-319.
- Shockley, K., Santana, M.-V., & Fowler, C. A. (2003). Mutual interpersonal postural constraints are involved in cooperative conversation. *Journal of Experimental Psychology: Human Perception and Performance, 29*(2), 326.
- Stangier, U., Clark, D. M., & Ehlers, A. (2006). Soziale Phobie: Fortschritte der Psychotherapie. *Göttingen, Hogrefe-Verlag.*
- Stangier, U., Clark, D. M., Ginzburg, D. M., & Ehlers, A. (2016). *Soziale Angststörung* (Vol. 28): Hogrefe Verlag.
- Stel, M., & Vonk, R. (2010). Mimicry in social interaction: Benefits for mimickers, mimicked, and their interaction. *British Journal of Psychology, 101*(2), 311-323. doi:10.1348/000712609X465424
- Strauss, B., Schoenherr, D., Altmann, U., Worrack, S., Tschacher, W., Ramseyer, F., . . . Schumann, E. (2018). [Workshop zur nonverbalen Synchronisation].
- Sun, X., Nijholt, A., Truong, K. P., & Pantic, M. (2011). *Automatic visual mimicry expression analysis in interpersonal interaction.*
- Sun, X., Truong, K. P., Pantic, M., & Nijholt, A. (2011). *Towards visual and vocal mimicry recognition in human-human interactions.* Paper presented at the Systems, Man, and Cybernetics (SMC), 2011 IEEE International Conference on.
- Swift, J. K., Callahan, J., & Levine, J. C. (2009). Using clinically significant change to identify premature termination. *Psychotherapy: Theory, Research, Practice, Training, 46*(3), 328.
- Swift, J. K., & Greenberg, R. P. (2012). Premature discontinuation in adult psychotherapy: A meta-analysis. *Journal of Consulting and Clinical Psychology, 80*(4), 547-559. doi:10.1037/a0028226
- Swift, J. K., Greenberg, R. P., Tompkins, K. A., & Parkin, S. R. (2017). Treatment refusal and premature termination in psychotherapy, pharmacotherapy, and their combination: A meta-analysis of head-to-head comparisons. *Psychotherapy, 54*(1), 47-57. doi:10.1037/pst0000104
- Theiler, J., Eubank, S., Longtin, A., Galdrikian, B., & Farmer, J. D. (1992). Testing for nonlinearity in time series: the method of surrogate data. *Physica D: Nonlinear Phenomena, 58*(1-4), 77-94.
- Tronick, E. D., Als, H., & Brazelton, T. B. (1977). Mutuality in mother-infant interaction. *Journal of communication, 27*(2), 74-79. doi:10.1111/j.1460-2466.1977.tb01829.x
- Tschacher, W., Giersch, A., & Friston, K. (2017). Embodiment and Schizophrenia: A Review of Implications and Applications. *Schizophrenia Bulletin, 43*(4), 745-753.
- Tschacher, W., & Ramseyer, F. (2017). *Synchronie in dyadischer Interaktion: Verkörperte Kommunikation in Psychotherapie, Beratung, Paargesprächen: Interaktionen in Alltag, Therapie und Kunst.*
- Tschacher, W., Ramseyer, F., & Koole, S. L. (2017). Sharing the Now in the Social Present: Duration of Nonverbal Synchrony Is Linked With Personality. *Journal of personality.*
- Tschacher, W., Rees, G., & Ramseyer, F. (2014). Nonverbal synchrony and affect in dyadic interactions. *Frontiers in Psychology, 5*(1323). doi:10.3389/fpsyg.2014.01323
- Tschacher, W., Scheier, C., & Grawe, K. (1998). Order and pattern formation in psychotherapy. *Nonlinear dynamics, psychology, and life sciences, 2*(3), 195-215.

- Tschacher, W., & Storch, M. (2010). Embodiment und Körperpsychotherapie. 161-175. doi:10.1007/978-3-642-01060-6_14
- Van Baaren, R., Janssen, L., Chartrand, T. L., & Dijksterhuis, A. (2009). Where is the love? The social aspects of mimicry. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 364(1528), 2381-2389.
- van Straaten, I., Engels, R. C. M. E., Finkenauer, C., & Holland, R. W. (2008). Sex Differences in Short-term Mate Preferences and Behavioral Mimicry: A Semi-Naturalistic Experiment. *Archives of Sexual Behavior*, 37(6), 902-911. doi:10.1007/s10508-007-9179-y
- Varlet, M., Marin, L., Capdevielle, D., Del-Monte, J., Schmidt, R. C., Salesse, R. N., . . . Raffard, S. (2014). Difficulty leading interpersonal coordination: towards an embodied signature of social anxiety disorder. *Frontiers in behavioral neuroscience*, 8.
- Vicaria, I. M., & Dickens, L. (2016). Meta-analyses of the intra-and interpersonal outcomes of interpersonal coordination. *Journal of Nonverbal Behavior*, 40(4), 335-361. doi:10.1007/s10919-016-0238-8
- Vrijzen, J. N., Lange, W.-G., Becker, E. S., & Rinck, M. (2010). Socially anxious individuals lack unintentional mimicry. *Behaviour Research and Therapy*, 48(6), 561-564.
- Vrijzen, J. N., Lange, W.-G., Dotsch, R., Wigboldus, D. H., & Rinck, M. (2010). How do socially anxious women evaluate mimicry? A virtual reality study. *Cognition and Emotion*, 24(5), 840-847. doi:https://doi.org/10.1080/13854040902833652
- Watanabe, T. (1983). A study of motion-voice synchronization. *Bulletin of JSME*, 26(222), 2244-2250. doi:10.1299/jsme1958.26.2244
- Webb, C. E., Rossignac-Milon, M., & Higgins, E. T. (2017). Stepping forward together: Could walking facilitate interpersonal conflict resolution? *Am Psychol*, 72(4), 374-385. doi:10.1037/a0040431
- Wenzel, A., Graff-Dolezal, J., Macho, M., & Brendle, J. R. (2005). Communication and social skills in socially anxious and nonanxious individuals in the context of romantic relationships. *Behav Res Ther*, 43(4), 505-519. doi:10.1016/j.brat.2004.03.010
- Westland, G. (2015). *Verbal and non-verbal communication in psychotherapy*: WW Norton & Company.
- Wierzbicki, M., & Pekarik, G. (1993). A meta-analysis of psychotherapy dropout: American Psychological Association.
- Wiltshire, T. (2015). Team interaction dynamics during collaborative problem solving.
- Wintersteen, M. B., Mensinger, J. L., & Diamond, G. S. (2005). Do gender and racial differences between patient and therapist affect therapeutic alliance and treatment retention in adolescents? *Professional Psychology: Research and Practice*, 36(4), 400. doi:10.1037/0735-7028.36.4.400
- Yun, K., Watanabe, K., & Shimojo, S. (2012). Interpersonal body and neural synchronization as a marker of implicit social interaction. *Scientific reports*, 2.
- Zhao, Z., Salesse, R. N., Gueugnon, M., Schmidt, R. C., Marin, L., & Bardy, B. G. (2015). Moving attractive virtual agent improves interpersonal coordination stability. *Human movement science*, 41, 240-254.

Abbildungsverzeichnis

| | |
|---|----|
| Abbildung 1 Kognitives Modell der Sozialen Angststörung nach Clark und Wells (2001), adaptiert aus Clark (2005)..... | 18 |
| Abbildung 2 Venn-Diagramm mit verschiedenen Begriffen für interpersonelle Synchronie, ¹ in dieser Arbeit synonym verwendete Oberbegriffe, ² in der Literatur ebenfalls verwendete Oberbegriffe, die nicht eindeutig abgrenzbar sind..... | 23 |
| Abbildung 3 In-Sync-Modell nach Koole und Tschacher (2016), S. 8..... | 40 |
| Abbildung 4 Überblick über wichtige Prozessschritte der Erfassung von nonverbaler Synchronie..... | 41 |
| Abbildung 5 Flowchart über Studienfluss der SOPHO-NET Videos (erste manuelle Sichtung),..... | 56 |
| Abbildung 6 Design und Ablauf der Studie TIMPATHIN, sX = X. Sitzung..... | 57 |
| Abbildung 7 Ablaufschema für die Erfassung von Bewegungen aus einem Video mittels Motion Energy Analysis (MEA)..... | 60 |
| Abbildung 8 Korrelationen von Bewegungen aus verschiedenen horizontalen Kamerapositionen (Abb. aus Altmann et al. (2019))..... | 62 |
| Abbildung 9 Kontroll-Pixel im Hintergrund (rot)..... | 63 |
| Abbildung 10 Vergleich aller Synchronisationsintervalle und der Intervalle bei Anwendung eines R ² -Schwellenwertes (Beispielproband 1)..... | 70 |
| Abbildung 11 Vergleich aller Synchronisationsintervalle und der Intervalle bei Anwendung eines R ² -Schwellenwertes (Beispielproband 2)..... | 70 |
| Abbildung 12 Veranschaulichung der Ergebnisse der Cox Regressionen zu den anderen Zeitpunkten..... | 75 |
| Abbildung 13 Vergleich der Synchronie zu Beginn der Therapie und zum Ende der Therapie innerhalb der verschiedenen Therapierichtungen; t-Test für verbundene Stichproben: ** entspricht p < .01, * entspricht p < .05, Abbildung und t-tests zur Illustration ergänzt. | 81 |
| Abbildung 14 Vergleich der Leading-Anteile zu Beginn der Therapie und zum Ende der Therapie, * entspricht p < .05..... | 82 |
| Abbildung 15 Grammer-Maße Größe der Bewegung und Dauer der Bewegung an einem Zeitreihen-Beispiel..... | 83 |
| Abbildung 16 Strukturgleichungsmodelle für die Vorhersage der IIP Skala FG (introvertiert, sozial vermeidend) durch die Bewegungsparameter Häufigkeit und Geschwindigkeit vor der Therapie (IIP _{FG1}) und nach der Therapie (IIP _{FG2})..... | 85 |

Abbildung 17 Erweitertes In-Sync-Modell der Psychotherapie, + entspricht einem positiven Einfluss, - entspricht einem negativen Einfluss, ... deuten an, dass das Modell horizontal durch weitere Aspekte erweitert werden kann..... 94

Tabellenverzeichnis

| | |
|---|----|
| Tabelle 1 Diagnosekriterien einer Sozialen Phobie nach DSM-V und ICD-10 | 16 |
| Tabelle 2 Anzahl der Studien und Effekte, die Prädiktoren für Therapieabbruch bei Patienten mit Sozialer Angststörung untersuchten (nach Eskildsen, 2010) | 21 |
| Tabelle 3 Einfluss verschiedener Faktoren auf interpersonelle Synchronie | 26 |
| Tabelle 4 Konsequenzen von interpersoneller Synchronie..... | 31 |
| Tabelle 5 Studien mit ihren Ergebnissen zum Zusammenhang zwischen interpersoneller Synchronie und Therapieoutcome | 34 |
| Tabelle 6 Vor- und Nachteile verschiedener Erfassungsmethoden für nonverbale Merkmale | 43 |
| Tabelle 7 Anzahl vorhandener Videos aus SOPHO-NET, die für TIMPATHIN eingeschlossen wurden sowie Ausschlusskriterien der durchgeführten Sitzungen | 57 |
| Tabelle 8 Technische und inhaltliche Voraussetzungen der Motion Energy Analysis | 61 |
| Tabelle 9 Hazard ratios mit 95% Konfidenzintervallen für verschiedene Subgruppenanalysen | 73 |
| Tabelle 10 Vergleich der Synchronie-Koeffizienten mit und ohne Hinzunahme des HAQ als Prädiktor..... | 76 |
| Tabelle 11 Ergebnisse Korrelationen, t-test und ANOVA für die Abhängigkeiten der Bewegungsparameter von demografischen Variablen und den IIP-Skalen..... | 84 |
| Tabelle 12 Eigenanteile innerhalb der Publikationen | 86 |
| Tabelle 13 Überblick über die aus Publikation 1 gewonnenen Erkenntnisse – Validität verschiedener linearer Zeitreihenanalyse-Techniken | 88 |

Anhang

Publikation 1: Schoenherr, D., Paulick, J., Worrack, S., Strauss, B., Rubel, J., Schwartz, B., Deisenhofer, A.-K., Lutz, W., Stangier, U., Altmann, U. (2018). Quantification of nonverbal synchrony using linear time series analysis methods: Lack of convergent validity and evidence for facets of synchrony. *Behavior Research Methods*, 1-23. doi:10.3758/s13428-018-1139-z.

Publikation 2: Schoenherr, D., Paulick, J., Strauss, B., Deisenhofer, A.-K., Schwartz, B., Rubel, J., Lutz, W., Stangier, U., Altmann, U. (2019). Identification of movement synchrony: Validation of windowed cross-lagged correlation and –regression with peak-picking algorithm. *PLoS One*, 14(2). doi:10.1371/journal.pone.0211494

Publikation 3: Schoenherr, D., Paulick, J., Strauss, B., Deisenhofer, A.-K., Schwartz, B., Rubel, J., Lutz, W., Stangier, U., Altmann, U. (zur Publikation angenommen, 06.12.2018). Nonverbal synchrony predicts premature termination of psychotherapy for social anxiety disorder. *Psychotherapy*.

Publikation 4: Schoenherr, D., Strauss, B., Paulick, J., Deisenhofer, A.-K., Schwartz, B., Rubel, J., Lutz, W., Stangier, U., Altmann, U. (in Begutachtung, eingereicht am 07.03.2019). Movement synchrony and attachment related anxiety and avoidance in social anxiety disorder. *Journal of Counseling Psychology*.

Publikation 5: Altmann, U., Schoenherr, D., Paulick, J., Deisenhofer, A.-K., Schwartz, B., Rubel, J., Stangier, U., Lutz, W., Strauss, B. (in Begutachtung, eingereicht am 14.02.2019) Associations between movement synchrony and outcome in patients with social anxiety disorder: Evidence for treatment specific effects. *Psychotherapy Research*.

Deskriptive Daten Bewegungsanalysen N = 86



Quantification of nonverbal synchrony using linear time series analysis methods: Lack of convergent validity and evidence for facets of synchrony

Désirée Schoenherr¹ · Jane Paulick² · Susanne Worrack¹ · Bernhard M. Strauss¹ · Julian A. Rubel² · Brian Schwartz² · Anne-Katharina Deisenhofer² · Wolfgang Lutz² · Ulrich Stangier³ · Uwe Altmann¹

© Psychonomic Society, Inc. 2018

Abstract

Nonverbal synchrony describes coordination of the nonverbal behavior of two interacting partners. Additionally, it seems to be important in human interactions, such as during psychotherapy. Currently, there are several options for the automated determination of synchrony based on linear time series analysis methods (TSAMs). However, investigations into whether the different methods measure the same construct have been missing. In this study, $N = 84$ patient–therapist dyads were videotaped during psychotherapy sessions. Motion energy analysis was used to assess body movements. We applied seven different TSAMs and recorded multiple output scores (average synchrony, maximum synchrony, and frequency of synchrony; in total, $N = 16$ scores). Convergent validity was examined using correlations of the output scores and exploratory factor analysis. Additionally, two criterion-based validations were conducted: investigations of concordant validity with a more generalized nonlinear method, and of the predictive validity of the synchrony scores for improvement in interpersonal problems at the end of therapy. We found that the synchrony measures only partially correlated with each other. The factor analysis did not support a common-factor model. A three-factor model with a second-order synchrony variable showed the best fit for eight of the selected synchrony scores. Only some synchrony scores were able to predict improvement at the end of therapy. We concluded that the considered TSAMs do not measure the same synchrony construct, but different facets of synchrony: the strength of synchrony of the total interaction, the strength of synchrony during synchronization intervals, and the frequency of synchrony.

Keywords Nonverbal behavior · Movement synchrony · Motion energy analysis · Time series analysis · Convergent validity

Currently, body movements can be assessed fully automatically and with high time resolution (e.g., 25 times per second) using either motion-tracking, motion capture devices, or video-based algorithms (Delaherche et al., 2012). *Motion energy analysis* (MEA) is a method that quantifies the intensity of videotaped movements frame-wise (Grammer, Honda, Juetter, & Schmitt, 1999). By determining a region of interest (ROI) for each of

two videotaped individuals (e.g., a patient and therapist during a psychotherapy session), two time series can be generated displaying the time course of the individuals' body movements. This technique has several advantages: (1) it is less time-consuming than collecting human ratings; (2) it is highly objective, reliable, and valid; and (3) in comparison to motion capture devices, no high-resolution camera equipment is necessary, and no sensors are attached to the patient's body (Altmann, 2010; Ramseyer & Tschacher, 2011). Therefore, during the past few years, the use of MEA has become enormously widespread. In behavioral and social science, MEA has been used to assess movements in mother–child interactions (Watanabe, 1983, 1987), child friendships (Altmann, 2011, 2013), and courtship behavior (Grammer et al., 1999; Grammer, Kruck, & Magnusson, 1998); for the comparison of nonverbal behavior in different types of interactions, such as tasks that elicit truthful, deceptive, argumentative, cooperative, competitive, or during fun tasks (Allsop, Vaitkus, Marie, & Miles, 2016; Altmann, 2011, 2013; Duran & Fusaroli, 2017; Paxton & Dale, 2013a;

✉ Désirée Schoenherr
desiree.schoenherr@med.uni-jena.de

¹ Institute of Psychosocial Medicine and Psychotherapy, University Hospital Jena, Jena, Germany

² Department of Clinical Psychology and Psychotherapy, Trier University, Trier, Germany

³ Department of Clinical Psychology and Psychotherapy, Goethe University Frankfurt/Main, Frankfurt/Main, Germany

Tschacher, Rees, & Ramseyer, 2014); in ecological psychology (Davis, Kay, Kondepudi, & Dixon, 2016); for the diagnosis of psychological disorders (Dean, Samson, Newberry, & Mittal, 2018; Dutschke et al., 2018; Kupper, Ramseyer, Hoffmann, Kalbermatten, & Tschacher, 2010; Kupper, Ramseyer, Hoffmann, & Tschacher, 2015); and in psychotherapy (Galbusera, Finn, & Fuchs, 2018; Paulick et al., 2018; Ramseyer, 2011, 2013; Ramseyer & Tschacher, 2011, 2016).

After generating motion energy time series, the influence of several variables (e.g., psychopathologies, attachment styles) on movement behavior can be investigated. Additionally, it has been demonstrated that the perception of a behavior shown by one interacting partner increases the probability of the other interacting partner engaging in that behavior (Chartrand & Bargh, 1999). During these sequences, an observer has the impression that the behavior of both interactors is synchronized, aligned, coordinated, co-regulated, or timed (Altmann, 2013; Bernieri & Rosenthal, 1991). Examples of such phenomena during interpersonal interactions are the imitation of facial expressions or gestures, posture mirroring, synchronous movements, or the convergence of voice parameters. Bernieri and Rosenthal (1991) pointed out that these phenomena are performed in a nonrandom manner, either following specific patterns or showing formal and temporal synchrony.

However, multiple terms and conceptualizations of these synchronization phenomena can be found, which are synonymous, partly overlapping or distinct from one another (Altmann, 2013; Feldman, 2007; Harrist & Waugh, 2002; Paxton, 2015). Paxton attempted to disambiguate different terms of synchronization phenomena (e.g., coordination, alignment, mimicry, imitation, synchrony, etc.) by natural language processing, showing that terms are rather separated by the underlying research question and study area than the phenomena under investigation. Terms are therefore rather based on different research areas with unique terminological trends (Paxton, 2015). Thus, a psychologist and a linguist may describe equal synchronization phenomena with different words. However, *synchrony* was identified as a suitable superordinate term to describe different conceptualizations such as facial imitation, movement synchrony or speech convergence. Due to the fact that the behaviors are nonverbal, we refer to these phenomena with the term nonverbal synchrony.

One possibility of grouping the different synchronization phenomena is by using a time dimension and the level of measurement as proposed by Altmann (2013). With respect to the time level, three different facets can be differentiated: (1) perfectly synchronous, simultaneous behavior or matching (*without* a time lag); (2) synchronous behavior with a time delay, echoing, alignment, imitation, or mimicry and (*with* a time lag); and (3) convergence, increasing similarity, and adaptation (*increasing similarity* over time). Referring to the level of measurement, data can be categorical, such as facial expressions (e.g., smile vs. no smile), or metric such as voice

pitch or movement intensity (Altmann, 2013). An example of the combination of both dimensions would be the examination of exactly simultaneous facial expressions (matching of categorical data). In addition to these different groups of synchrony, different algorithms exist for measuring the construct. Nevertheless, all algorithms are used to measure the construct synchrony; systematic investigations of their validity and concordance are missing. However, these investigations are needed to make judgments about the comparability of study results.

Only one study compared human ratings of nonverbal synchrony and nonverbal synchrony obtained by cross-lagged correlation (Paxton & Dale, 2013b). The authors compared the results of the cross-lagged correlation of human second-by-second coding of the movements of two persons with the results of the cross-lagged correlation of a frame-differencing technique. This study provided evidence that movements rated by humans and by the algorithms led to comparable synchrony results. However, more research will be needed to disentangle the different nonverbal synchrony constructs and to examine which time series analysis methods (TSAMs) lead to comparable results.

Since rater reliability in human synchrony ratings is rather weak, or many raters are needed to obtain high values (Bernieri, 1988), we focused on objective and reliable algorithms. As illustrated above, the construct to which synchrony obtained by algorithms is compared is crucial. To date, no true value/construct of synchrony has been defined. Therefore, in line with Cronbach and Meehl (1955), who stated that investigations of construct validity (convergent and discriminant validity) are especially relevant if no direct criterion or “universe of content” is available, we examined the *convergent* validity of different algorithms. For this aim, we first describe different linear-based algorithms that assess nonverbal synchrony, and then apply these to an identical dataset containing motion energy time series. Furthermore, we tested whether all measures load on a common factor. If not, the adequate factor solution would be determined. In addition, we conducted two further validations: First, we compared linear-based TSAMs to a more generalized nonlinear approach (cross-recurrence quantification analysis). Second, the predictive validity of the TSAM output scores was examined by inspecting the synchrony-outcome association using data from psychotherapy sessions.

However, the aim of our article was not to identify the best algorithm to assess synchrony, but rather to test whether different algorithms are equivalent operationalizations of the construct of nonverbal synchrony.

Systematization of different linear-based algorithms assessing the nonverbal synchrony of two interacting persons

The process of co-regulation between two individuals can be understood as dynamic (Altmann, 2013; Boker, Rotondo, Xu,

& King, 2002; Fogel, 1993). In our study, nonverbal synchrony was operationalized as the degree of association between the nonverbal behaviors of two interacting persons. Approaches assuming a linear relationship between both interacting partners (correlations or regressions) are very common. Thereby, correlations or regressions are computed for the entire time series or windows of the time series. Subsequently, the obtained values are usually aggregated to an output score in order to quantify this association (also called “co-aggregation measures,” see Coco & Dale, 2014). Other approaches investigate the degree of association by using recurrence plots that show recurrence points if two systems are similar to each other with respect to their phase trajectories. Cross-recurrence quantification (CRQA) analysis can be used to analyze these points (for an overview, see Delaherche et al., 2012). This method is widely applied to analyze states of one time series and how the other revisits these states, which is especially relevant for the investigation of temporal patterns of nonverbal synchrony during an interactional process. Thereby, this method does not make the assumption of a linear relationship between two time series (Coco & Dale, 2014; Marwan & Kurths, 2002). Comparing linear methods with cross-recurrence analysis showed that a systematic association between both methods is sometimes missing (Wiltshire, 2015). However, the results from CRQA can be regarded as a generalization of cross-correlation methods (Marwan, Romano, Thiel, & Kurths, 2007). Spectral methods provide further options to analyze nonverbal synchrony (for more details, see Delaherche et al., 2012).

In the present article, we focus on methods that assume a linear relationship between time series trajectories, because (1) these methods are widely used in interactional research (Altmann, 2013; Kupper et al., 2015; Nelson, Grahe, Ramseyer, & Serier, 2014; Paulick et al., 2018; Ramseyer & Tschacher, 2011) and (2) including other approaches would result in a reduction of comparability of the different methods, since the statistical assumptions of the methods are very distinctive. As a validation, we will contrast these methods with one nonlinear method. We chose CRQA because this method is widely used in the context of interactional analyses (e.g., Shockley, 2005; Shockley, Richardson, & Dale, 2009).

Linear TSAMs can be differentiated by (a) assumptions about time lag (matching vs. echoing) and (b) the length of time series windows (global vs. local TSAMs).

Assumptions about time lag Two operationalizations of nonverbal synchrony can be distinguished: (a1) using no time delay (termed “*matching*”; see above) or (a2) using a specific time delay (termed “*echoing*”). Most algorithms that measure echoing also include matching. Currently, the selection of the maximum appropriate time lag is largely left up to the researcher. It characterizes the greatest interval separating the behavior of two persons, which is still considered to be

connected. Investigation of the coordination of skin conductance level showed a meaningful nonverbal synchrony with a maximum time lag of 7 s (Robinson, Herman, & Kaplan, 1982). Additionally, virtual agents that mimic persons with a time delay of exactly 4 s were rated more positively than nonmimickers (Bailenson & Yee, 2005). In accordance with that, Bilakhia, Petridis, Nijholt, and Pantic (2015) recommended a time lag of 0.04 to 4 s, based on the empirical investigation of episodes showing motor mimicry. Altmann (2011) used a maximum time lag of 2.5 s, whereas Ramseyer and Tschacher (2011) used a time lag of 5 s. Another study by Louwerse, Dale, Bard, and Jeuniaux (2012) showed that mimic expressions and head movements have a short time lag of approximately 1.5 s. However, there have also been attempts to evaluate the chosen time-lag based on the comparison with shuffled data, implying that the choice of best lag could be empirically determined.

Length of time series window Regarding the second dimension—the length of the assessed time series window—two types of linear TSAMs can be differentiated: (b1) global and (b2) local measures (Altmann, 2013). Two global methods are cross-lagged correlation (CLC) and cross-lagged regression (CLR; Gottman & Ringland, 1981; Paxton & Dale, 2013b). Global TSAMs calculate Pearson product-moment correlations or regressions using the two entire time series. The entire time series are incorporated with equal or different starting points. The distance between the starting points is referred to as the time lag (see above). Global methods have the advantage of being less time-consuming in terms of computational costs, as they are less complex. However, when using global methods, the assumption of global stationarity is made. Stationarity means that given the respective time series, the mean and variances stay constant over time. Additionally, by using global methods, it is assumed that person A influences person B or vice versa for the complete time course; that is, there is no changing interdependence between the interacting partners. This assumption is often violated in naturalistic contexts such as human communication (Boker et al., 2002). As a result, local TSAMs were developed.

Local TSAMs analyze the entire time series window-wise, conducting Pearson product-moment correlations (or regressions) between segments/windows of two time series. Examples are windowed cross-correlation (WCC; Tronick, Als, & Brazelton, 1977), windowed cross-lagged correlation (WCLC; Altmann, 2013; Paulick et al., 2018; Ramseyer & Tschacher, 2011), and windowed cross-lagged regression (WCLR; Altmann, 2011). By disassembling the entire time series into windows, the assumption of stationarity can be made locally, which is less restrictive. The length of the analyzed windows is referred to as the bandwidth or window width. The size of this window, in addition to the maximal time lag, is of great importance. Setting the window width too

small results in decreased reliability, whereas a large window width likely results in a violated assumption of local stationarity (Boker et al., 2002). Currently, the selection of an appropriate window width is largely up to the researcher. Theoretically, the window width must be large enough to completely capture the interrelated movements of two persons. From this perspective, Boker et al. (2002) recommended a window width of 4 s. From a methodological point of view, the window must incorporate enough values to determine a stable correlation. Schönbrodt and Perugini (2013) recommended a sample size of at least 65 values to stably detect high correlations. For weaker correlations, 250 values are needed to obtain stable correlations. Tronick et al. (1977) used a window size of 10 s (= 10 values). Cappella (1996) postulated that a window width of at minimum 50–70 values should be used and that this window width should be about 4–5 times larger than the maximal time lag used. However, the optimal trade-off, which comprises a high reliability of correlations, a stationary time series window and a theoretically plausible episode of interrelated nonverbal behaviors, has yet to be empirically determined.

Another important aspect is whether these windows are applied as moving/rolling windows that overlap or whether the time series is split into nonoverlapping windows. Splitting the time series into windows that do not overlap may result in synchrony events at the splitting point not being detected. Therefore, rolling windows with overlap are usually preferred. Regarding forecasting, an application of rolling origins is also preferable to fixed origins in processing time series, as they yield higher efficiency and reliability in out-of-sample tests (Tashman, 2000).

Human interaction is characterized by the interdependence of both interacting partners, meaning that there are periods during which the behavior of person A influences the behavior of person B (A = drive, B = driven), and vice versa (B = drive, A = driven). The drive is also known as the *zeitgeber*, or the person who sets the pace as well as the person who leads (Boker et al., 2002; Kupper et al., 2015; McGarva & Warner, 2003). When using global methods, the reciprocity of an interaction may not be captured, because by shifting one time series with a time lag and calculating the correlation with the second time series, the *zeitgeber* of the interaction is fixed. That is, person A always influences person B, or vice versa. If windows are used, the *zeitgeber* may vary for each window (Ramseyer & Tschacher, 2011). This change of leading may be an important aspect of nonverbal communication (Boker et al., 2002). Using a local method makes examining dynamic interaction changes possible (Boker et al., 2002).

Output scores By using different linear TSAMs, different values result as output scores. Paxton and Dale (2013b) reported that a variety of output scores, including the average synchrony score and the highest value, can be used to quantify

synchrony. Most measures assess the strength of synchrony. Altmann (2011, 2013) proposed the frequency of synchrony—that is, the ratio of synchronous time to total time—as an output score. The selection of an output score depends largely on the research question, the researcher (Paxton & Dale, 2013b), and the algorithm, since not every algorithm provides every output score (e.g., the frequency of synchrony can only be calculated using Altmann's, 2011, 2013, algorithms). With respect to the research question, it is possible to use the average score as an indicator of the overall interrelatedness of an interaction (e.g., Ramseyer & Tschacher, 2011). Additionally, the maximum score of a window can be used to investigate microprocesses in psychotherapy. The frequency of nonverbal synchrony may be used to evaluate interaction on a temporal level (i.e., what percent of the interaction was synchronous). This means that many differences exist with respect to linear TSAMs: Various algorithms have been used, and different output scores may be calculated. Some recent approaches are listed in Table 1 and introduced in the following sections.

Cross-lagged correlation (CLC) In comparison to a simple Pearson product-moment correlation, CLC additionally considers the time lag between two time series, so that the correlation between the time series is calculated for each time lag until a maximum value of the time lag is reached (Kato et al., 1983). With respect to its interpretation, CLC refers to echoing. The output of the CLC is a function of the correlations displaying the strength of the association between the time series with respect to every permitted time lag. The two most common outcome scores are the average of all CLC values and the maximum CLC value. By averaging the correlations of different lags, averaged degrees of global matching and echoing are obtained. Using the highest correlation yields the highest global matching or echoing in the sequence.

Cross-lagged regression (CLR) An important issue with regard to the analysis of time series data is autocorrelation. Referring to body movement time series, an autocorrelation means that the individual's current movement is influenced by their previous movements. Not considering this autocorrelation can result in spurious correlations (Altmann, 2011, 2013; Gottman & Ringland, 1981). In recent years, the global measure CLR was used to overcome this shortcoming (Cappella, 1996). When using the CLR, a maximum time lag must also be specified. One difference to CLC is that CLR uses a regression. In this regression, two predictors are incorporated to predict a person's later movement: (1) the previous movements of the person (autocorrelation) and (2) the previous movement of the interacting partner (cross-correlation). If the model including the autocorrelation and the cross-correlation cannot explain significantly more variance than the model including the autocorrelation only, the association

Table 1 Global and local TSAMs and their applications

| TSAM | Time lag | Windowed | Control for auto-corr | Outcome | Output score | | | For example, applied by |
|--------------------|----------|----------|-----------------------|----------------------|--------------|-----|------|--|
| | | | | | av | max | freq | |
| CLC | X | | | Correlation function | X | X | | Kato et al. (1983), Paxton and Dale (2013b) |
| CLR | X | | X | R^2 function | X | X | | Gottman and Ringland (1981) |
| WCC | | X | | Correlation function | X | X | | Tronick et al. (1977) |
| WCLC _{S1} | X | X | | Correlation function | X | X | | Paulick et al. (2018), Ramseyer and Tschacher (2011) |
| WCLC _{S2} | X | X | | Correlation function | X | X | | Kleinbub and Ramseyer (2018) |
| WCLC _F | X | X | | R^2 function | X | X | X | Altmann (2011, 2013) |
| WCLR | X | X | X | R^2 function | X | X | X | Altmann (2011, 2013) |

C = correlation; CLC = cross-lagged correlation; CLR = cross-lagged regression; WCC = windowed cross-correlation; WCLC_{S1} = windowed cross-lagged correlation (e.g., Paulick et al.); WCLC_{S2} = windowed cross-lagged correlation R (e.g., Kleinbub & Ramseyer); WCLC_F = windowed cross-lagged correlation (e.g., Altmann); WCLR = windowed cross-lagged regression; av = average output score; max = maximum output score; freq = frequency output score.

is categorized as nonsynchronous. Models can be compared using R^2 difference testing. Similar to the CLC, the CLR results in a function of R^2 values referring to every permitted time lag. The interpretation of possible output scores (average, maximum) is similar to the CLC.

Note that the distinction between correlational and regressive approaches refers to the way in which the association between both persons' time series is operationalized (Altmann, 2011, 2013). However, both approaches are similar, because a correlation is a one-predictor regression with standardized variables.

Windowed cross-correlation (WCC) The WCC represents the simplest local measure. Similar to the simple Pearson product-moment correlation, the correlation is calculated without considering a time lag. The only difference is that the time series is disassembled into windows, resulting in the correlation being computed window-wise (Tronick et al., 1977). The WCC can be applied with overlapping or nonoverlapping windows. Before computing the WCC, the bandwidth must be specified. As a result of the WCC, a function displaying the strength of the association between the two time series is obtained for each window. With respect to the WCC, a commonly used output score is the mean of all obtained cross-correlations (Bilakhia et al., 2015; Nagaoka & Komori, 2008). Accordingly, the resulting output score of the WCC can be interpreted as the averaged degree of local matching. The maximum score can also be determined as the maximum correlation of one window, operationalizing the maximum synchrony in the sequence.

Windowed cross-lagged correlation (WCLC) To our knowledge, the WCLC was first applied by Watanabe (1983) and became more popular after its introduction by Boker et al. (2002), who used automated motion capture time series and combined WCLC with a peak-picking algorithm. In WCLC, the association between the time series windows of two

individuals is calculated using either overlapping or nonoverlapping windows. The correlation is calculated window-wise—that is, as a first step the correlation between the first window of the time series of person A and the first window of the time series of person B is conducted (time lag = 0). Afterward, one time series window is shifted by one time lag and the correlation between the first window of the time series of person A and the second window of the time series of person B is conducted (time lag = 1). This procedure is repeated until the maximum time lag is reached and all windows of both time series have been used. After applying the WCLC an $m \times n$ matrix of values is obtained, where m characterizes the number of different time lags and n denotes the time point (in frames). Note that a positive time lag refers to predictions in which person A is the drive and person B is the driven, whereas a negative time lag means that person B is the drive and person A the driven.

Various versions of the WCLC can be distinguished. Versions differ mainly in their preprocessing and processing of the time series or how they use the matrix of correlations/ R^2 to generate an output score quantifying synchrony. In the following study, three versions of the WCLC are presented. The first version, called WCLC_{S1} here and developed by Paulick et al. (2018), is based on correlations and computes the strength of the association between the two time series as an outcome measure (subscript S stands for strength, and 1 indicates that it is the first WCLC strength measure). More precisely, the authors re-created the algorithm by Ramseyer and Tschacher (2011), which is the most popular method within the psychotherapeutic context. The corrected motion energy time series (z -transformed, with the threshold for minimal movement; see Grammer et al., 1999) are cross-correlated in nonoverlapping windows of a 1-min duration (for each window, cross-correlations were computed for positive and negative time lags of up to 5 s, in steps of 0.04 s). Subsequently, the cross-correlations of the matrix are

Fisher's- Z transformed and their absolute values are aggregated to an output score for the nonverbal synchrony of each video sequence.

The second version is $WCLC_{S2}$, developed by Kleinbub and Ramseyer (2018), which was recently published as the R package *rMEA* (again the subscript S is for strength, and 2 for the second WCLC strength measure). The function incorporates the following input variables: *mea* (MEA time series), *lagSec* (maximum time lag, in seconds), *winSec* (bandwidth, in seconds), *incSec* (step size, in seconds), *rZ* (application of Fisher's z -transformation), and *ABS* (transformation to absolute values). It results in a cross-correlation matrix that is aggregated to an output score measuring the strength of the association between two time series. A difference between the two presented $WCLC_S$ versions is the applied minimal movement threshold in $WCLC_{S1}$. Since Paulick et al.'s (2018) $WCLC_{S1}$ is based on Ramseyer and Tschacher's (2011) procedure, Grammer et al.'s (1999) threshold for minimal movements was applied before calculating the cross-correlations. However, in $WCLC_{S2}$ this threshold is not present anymore. Apart from this difference, $WCLC_{S1}$ and $WCLC_{S2}$ are comparable.

The third version is $WCLC_F$ (where the subscript F stands for *frequency*) by Altmann (2011, 2013). Altmann (2013) used the determination coefficient (squared correlation) as a result of the WCLC because it contains only positive values, and higher correlations are weighted higher than very low correlations. An example of a time series including a synchronization interval, surface and contour plots of a possible matrix are displayed in Fig. 1. The distinction between the algorithms measuring the strength ($WCLC_{S1}$, $WCLC_{S2}$) and this algorithm lies in the peak-picking algorithm, that is used to identify synchronization intervals. The peak-picking algorithm is described in the Peak-Picking Algorithm paragraph below. In this version of the WCLC, not the synchrony of the entire interaction is quantified, but the local synchronization intervals. These are characterized by the highest values/ R^2 's of synchrony in the local environment (Boker et al., 2002), which are arranged on a horizontal line (see Fig. 2). To compute an output score of synchrony referring to the $WCLC_F$, the duration of all synchronization intervals of a particular time series pair was summed up. Afterward the ratio of the time, synchronized to the total duration of the sequence, was calculated (Altmann, 2011, 2013). Thus, the output score indicates the percentage of synchronous interactions. Referring to the $WCLC_{S1}$ and $WCLC_{S2}$, correlations of the different time lags and time points were averaged to obtain an output score. Therefore, another difference between $WCLC_F$ and both versions of $WCLC_S$ is the output score, because this version assesses instead the frequency of the associations between the two time series as the primary output score. However, the algorithm also computes the strength of the association in the synchronization intervals. Therefore, the average and maximum output scores are also calculated. Note that in all

versions, only positive values are used. In $WCLC_{S1}$ and $WCLC_{S2}$, absolute values of the correlations are used, whereas in $WCLC_F$, R^2 is used. The procedures differ in that calculating the R^2 values results in weighting higher correlations more than lower correlations. WCLC conceptualizes both matching and echoing locally.

Windowed cross-lagged regression (WCLR) Altmann (2011, 2013) developed WCLR, which is a combination of CLR and WCLC, in order to benefit from the advantages of both methods. WCLR is similarly to WCLC. The main difference is that a regression is conducted that includes the previous behavior of an individual as a predictor (similar to the CLR). Interrelatedness is computed window-wise, whereby the position of the reference window and time-lag to the other window are varied. Altmann (2011, 2013) proposed the application of two models on one pair of windows:

$$\text{Model 1 : } X_{\text{self } t} = \alpha_0 + \alpha_1 X_{\text{self } s} + \varepsilon_{M1 t} \quad (1)$$

$$\text{Model 2 : } X_{\text{self } t} = \beta_0 + \beta_1 X_{\text{self } s} + \beta_2 X_{\text{partner } s} + \varepsilon_{M2 t} \quad (2)$$

Equation 1 shows Model 1, incorporating the previous behavior of an individual in order to predict the current behavior—that is, the regression accounts for possible autocorrelation. $X_{\text{self } t}$ refers to the individual's own behavior at time point t , and $X_{\text{self } s}$ refers to the individual's previous behavior at time point s (i.e., s is prior to t). α_0 denotes the intercept, α_1 the slope of the linear regression, and $\varepsilon_{M1 t}$ refers to the residual variance of Model 1 at a time point t . These terms are used analogously in Eq. 2. Additionally, a second predictor, $X_{\text{partner } s}$, is included, which denotes the previous behavior of the interacting partner at time point s .

Accordingly, if Model 2 fits significantly better than Model 1, synchrony between the two individuals is detected. Whether Model 2 fits significantly better than Model 1 can be investigated with an R^2 difference test (Altmann, 2011). Computing the WCLR results in an $m \times n$ matrix of R^2 difference values, where m characterizes the number of different time lags and n denotes the time point (in frames). To determine an output score, a peak-picking algorithm is applied to the matrix. The output of the WCLR (and the WCLC) is a list of intervals that were classified by the algorithm as synchronous (time-lagged) time series sequences. On the basis of this list, the ratio of synchronous time to the total duration of the sequence is calculated. Altmann (2013) showed that the WCLR is preferable to the WCLC by using simulated (cyclic) time series. The WCLR also conceptualizes local matching and echoing.

Peak-picking algorithm The output of the $WCLC_F$ and WCLR (implemented by Altmann, 2011, 2013) is an R^2 matrix, which stores the strength of window-wise associations between both persons' time series. To extract the synchronization intervals

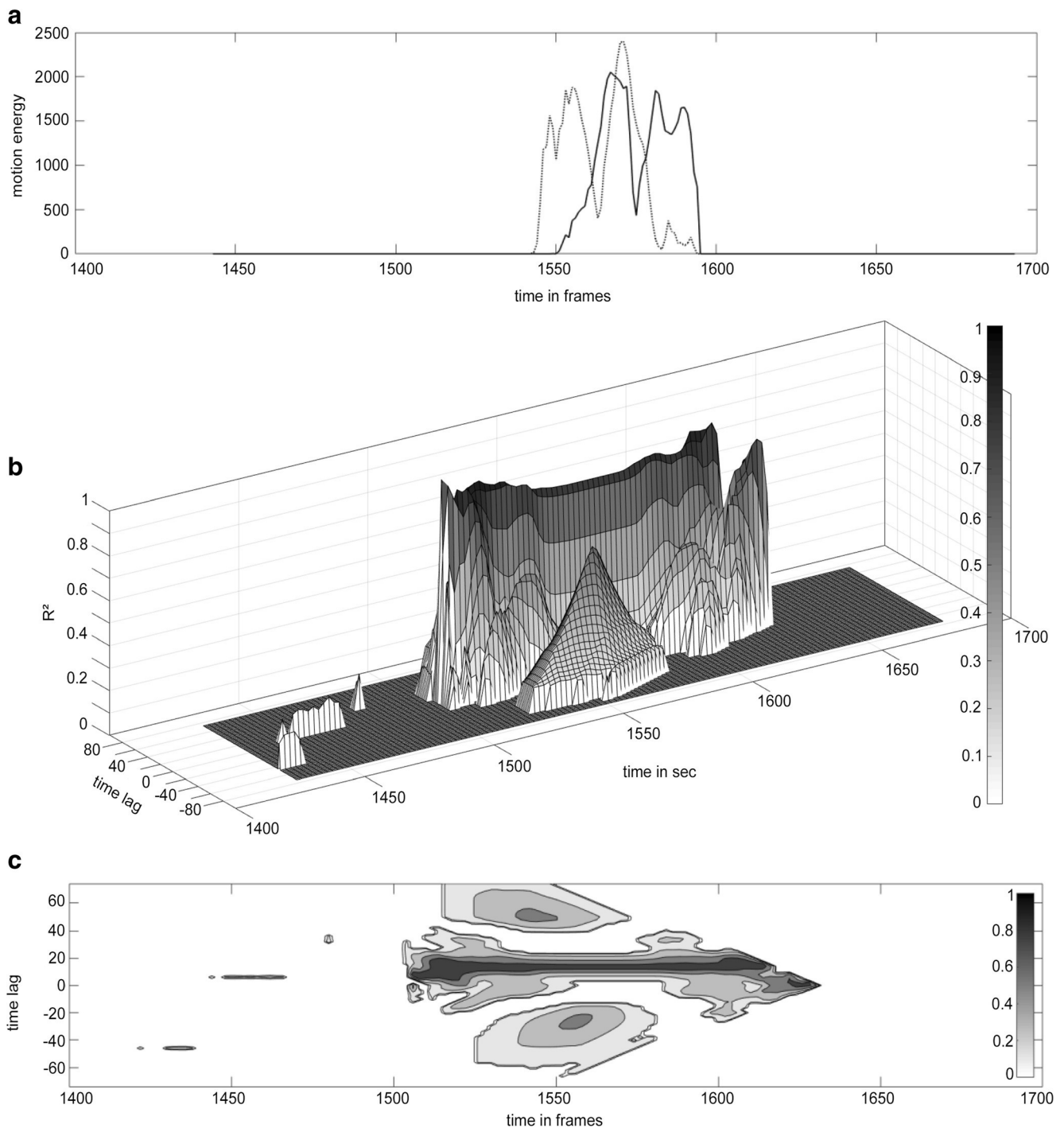


Fig. 1 Time series and two plots of the resulting R^2 matrix after applying $WCLC_F$. Panel A shows the two time series (dotted and solid) with an interval displaying synchrony (interval duration approximately 2 s,

applied bandwidth 3 s = 75 frames displayed in gray). Panel B shows a surface plot of the R^2 matrix, and panel C a contour plot of the R^2 matrix. The time lag is displayed in frames

from the matrix, peaks must be found. Boker et al. (2002) defined the maximum of the cross-correlation as a peak, where values decrease on either side of the particular peak in a specific local region. In their algorithm, the peak that has the closest distance to time lag = 0 is chosen as the best peak to identify the synchronization interval. Altmann (2011, 2013) created an alternative peak-picking algorithm (for an

illustration, see Fig. 2). First, all local maxima in the R^2 matrix (peaks) are identified. Next, neighboring peaks with an equal time lag are determined. The third step is a summary of neighboring peaks to peak crests (with a small time lag tolerance). Fourth, all intervals that last less than 0.4 s are removed, because this amount of time is considered too short to contain meaningful synchrony. Subsequently, an adjustment of the

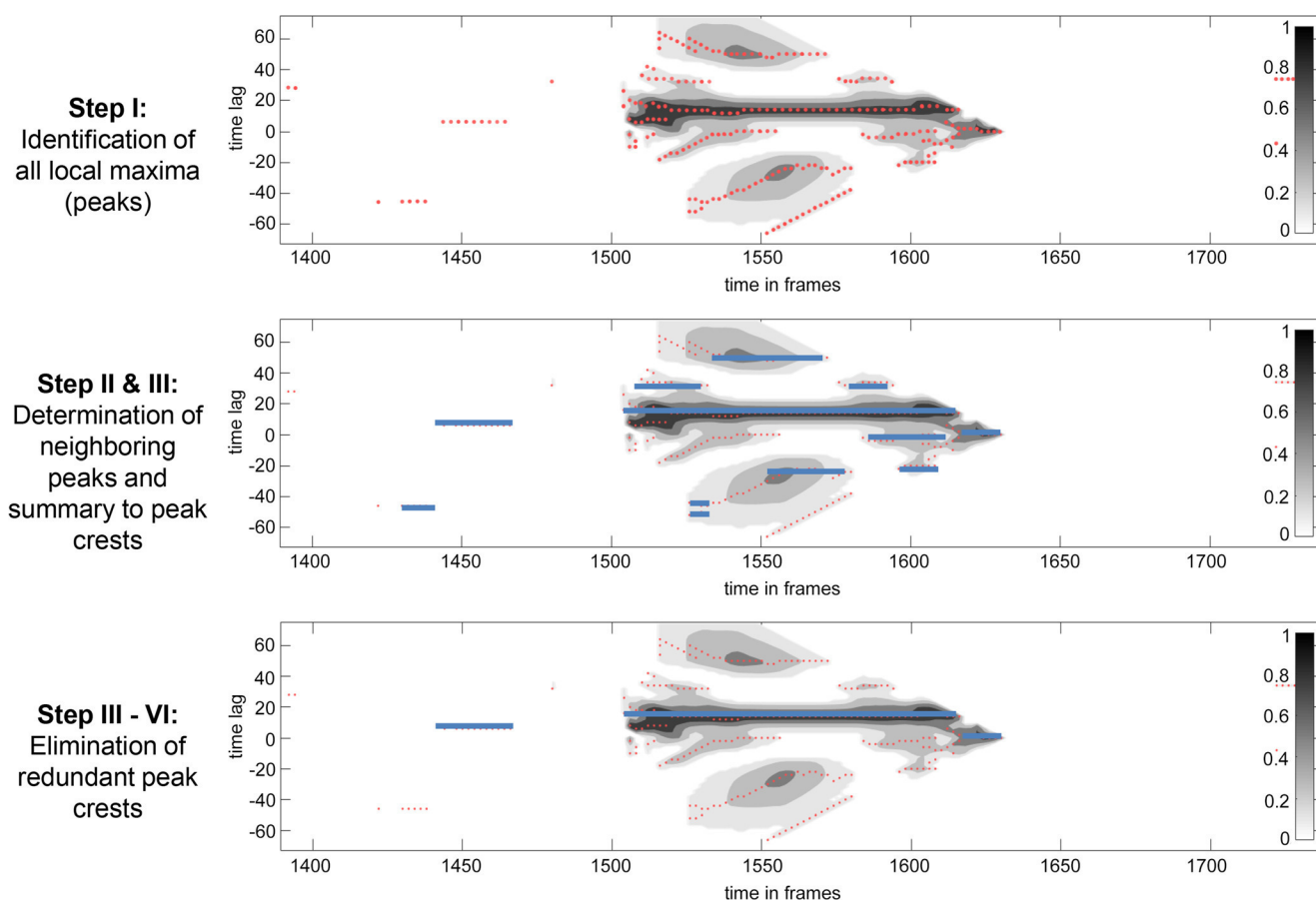


Fig. 2 Contour plot of the R^2 matrix, which is a result of the analysis of two time series with $WCLC_F$. Red dots indicate the R^2 peaks selected by the Boker et al. (2002) algorithm, whereas blue lines are the R^2 “mountain crests” selected by the Altmann algorithm

onset of an interval takes place using the time lag. Finally, overlapping peak crests are identified and peak crests with the lower R^2 average are removed (see Fig. 2).

From a theoretical point of view, the peak-picking algorithms by Boker et al. (2002) and Altmann (2013) have fundamentally different implications for the concept of nonverbal synchrony. The former implies that synchrony occurs over the entire duration of the interaction, with only the strength of synchrony varying over time. Therefore, the primary aim is not the detection of synchronization intervals but the quantification of synchrony. This implies that investigations of microprocesses within therapy cannot be conducted, because no synchronization intervals are determinable. The latter peak-picking algorithm assumes the existence of so-called *synchronization intervals* and that the time series are interrelated only during these intervals. On the basis of such an on–off pattern, the frequency of synchrony can be quantified. However, within a synchronization interval, Altmann’s (2013) peak-picking algorithm also computes the strength of synchrony in terms of the average R^2 values of all identified synchronization intervals. This implies the assumption that an interaction contains in-synchrony intervals and also intervals during which no synchrony can be found. It allows for an examination of microprocesses within psychotherapy, in

which the content of therapy may be related to different synchronization intervals and in which specific moments of change (e.g., repairs of alliance ruptures) may be investigated further.

Research questions

The aim of the present study was twofold: to test construct validity and investigate criterion validity with two different criteria. Since authors have used different parameter settings (e.g., bandwidth or overlapping vs. nonoverlapping windows), we conducted all analyses on a set of parameters recommended by the authors and on a completely equal set of parameters, for comparability both between studies and within the present study.

Convergent validity First, we examined the convergent validity of linear TSAMs used in behavioral and social science. We applied these TSAMs to motion energy time series, resulting in movement synchrony output scores. Thereby, we used various output scores (average, highest value, and the ratio of synchronous time). We hypothesized that all correlations would be moderate to high because all TSAMs conceptualize nonverbal synchrony and are widely applied to motion energy time series

in order to assess movement synchrony. Additionally, we tested whether a common factor model fits the data. Should this not be the case, an exploratory factor analysis would be conducted to systemize the output scores of different TSAMs.

Criterion validity (1) We described CRQA as a widely applied nonlinear approach to investigating synchrony that is more generalized than the linear TSAMs presented. Thus, we also conducted correlations with this measure and included it in the factor analysis in order to validate the results of the linear TSAMs. (2) Our sample comprised patients suffering from social anxiety disorder (SAD), who are considered to show a high level of impairment in interpersonal contexts (Wenzel, Graff-Dolezal, Macho, & Brendle, 2005). Furthermore, with respect to therapy outcome in a disorder-heterogeneous sample, research has shown that psychotherapy enhances interpersonal abilities and reduces interpersonal problems, especially when movement synchrony during the early stage of therapy is high (Ramseyer & Tschacher, 2011). Therefore, we hypothesized that output synchrony scores would be negatively related to interpersonal problems at the end of therapy. We conducted partial correlations (controlling for initial impairment) to examine the predictive validity of the presented TSAMs.

Method

Background

Our data were gleaned from video recordings from the multi-center randomized controlled SOPHONET treatment study, conducted from 2007 to 2009 by outpatient clinics at universities in Bochum, Dresden, Jena, Mainz, and Goettingen, Germany (Leichsenring et al., 2013, 2014).

Setting and material

The therapies included five preparatory sessions that are compulsory in the German health care system (for more details, see Leichsenring et al., 2013, 2014), as well as 25 individual 50-min treatment sessions. Due to the fact that the videos were recorded in different centers, camera positions varied slightly (lateral view $\sim 90^\circ$, angular view $\sim 45^\circ$). The psychotherapist and patient were always recorded with one video camera. Furthermore, only video files with a constant camera position, camera settings, and light conditions were considered. To be included in the study, videos had to show both persons (therapist and patient) for at least 15 min during the first half of the therapy session. The starting point of the considered sequence was set after welcoming, administrative questions, and the filling out of questionnaires. The latest possible start time was set to 10 min. Videos that showed the filling in of questionnaires for more than 10 min were excluded. Further exclusion criteria

included: the presence of another person apart from the therapist and patient, one person leaving his/her chair during the 15 min, and one person changing his/her position so that he/she was no longer visible in the video. To ensure the comparability of video files, they were converted to equivalent formats using a size of 640×480 , a frame rate of 25 fps, and a bit rate of 2,000 Kbit per second (Any Video Converter 3.0, AVC, 2009).

Study subjects

On the basis of the inclusion criteria of the SOPHONET study, the patients had to be 18–70 years old, have a diagnosis of SAD according to the German version of the Structured Clinical Interview for DSM-IV (Wittchen, Wunderlich, Gruschwitz, & Zaudig, 1997), a score > 30 on the Liebowitz Social Anxiety Scale (LSAS; Mennin et al., 2002), and a primary diagnosis of SAD with respect to the Anxiety Disorders Interview Schedule (Brown, Barlow, & Di Nardo, 1994). Comorbid disorders less severe than SAD were allowed, with the exception of psychotic disorders, acute substance-related disorders, and cluster A and B personality disorders. Additional exclusion criteria were a risk of self-harm, organic mental disorders, severe medical conditions, and concurrent psychotherapeutic or psychopharmacological treatments (see Leichsenring et al., 2013, 2014).

In the present study, we examined a subsample of the SOPHONET sample. The selection was driven by the following inclusion criteria for video quality: (1) patients had to be visible for the first 15 min of the third therapy session, (2) video quality was sufficient without severe video errors, and (3) no third person was present. After applying the inclusion criteria to the video files, 84 of 495 study subjects remained for the present investigation (study flow is presented in Fig. 3). Regarding the present investigation, 53 patients were female, 52 had a high school diploma or higher, and 37 were living in a current relationship. The mean therapy duration per patient was 23.87 sessions ($SD = 6.93$). The mean pretreatment severities were 73.44 ($SD = 21.79$) on the LSAS, 10.41 ($SD = 6.51$) on the Beck Depression Inventory (BDI; Beck, Ward, Mendelson, Mock, & Erbaugh, 1961), and 1.74 ($SD = 0.47$) on the Inventory of Interpersonal Problems (IIP; Horowitz, Strauss, Thomas, & Kordy, 2016).

Assessment of movement

We used the first 15 min of the third therapy session (\pm one session) to conduct motion energy time series analysis, because it has been demonstrated that movement synchrony of the first 15 min is representative of the entire 50 min of a therapy session (Paulick et al., 2018; Ramseyer & Tschacher, 2011). To assess the time course of individual body movements, we used motion energy analysis (MEA), implemented in MATLAB 2016 (The MathWorks, Inc, 2016; Altmann, 2013).

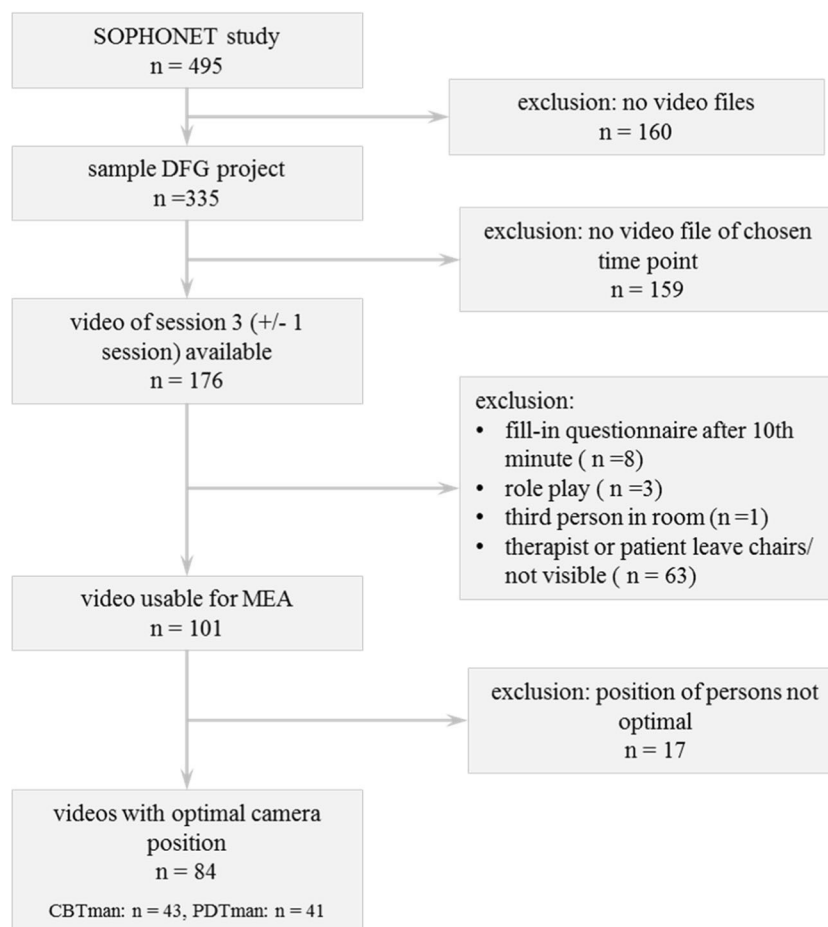


Fig. 3 Study flow. DFG = German Research Foundation, CBT = cognitive-behavioral therapy, PDT = psychodynamic-oriented therapy, man = manual-guided

The ROI covered the upper body from the chair's seating-base upward.

We determined the cutoff value for meaningful pixel changes empirically (according to Altmann, 2013). For this purpose, we used $n = 20$ video files. We specified 30 different pixels in the background as control pixels, which should not show any grayscale intensity changes as a result of movements. Time series of these 30 pixels were calculated, showing the intensity change from one frame to another. Afterward, we determined the 99% quantile of all values, which equaled 12—that is, 99% of the intensity changes were lower than 12. Thus, the cutoff value was set to 12. Furthermore, we applied a 2-D median filter in order to reduce video noise (i.e., irregularities in the video sequences), as described in Altmann, Schoenherr, Paulick, Knitter, et al. (submitted). In addition, we transformed all motion energy time series to a value range of 0–100 by dividing each value by the number of pixels in the ROI and multiplying it by 100. The value of 100 means that 100% of the ROI was activated; the value of 1 means that 1% of the ROI was activated. Therefore, the values of different persons were comparable, even when they did not have the same ROI size. Moreover, we filtered video errors by

applying the MATLAB script described in Altmann, Schoenherr, Paulick, Knitter, et al. (submitted), available at <https://github.com/10101-00001/MEA>.

Parameter settings for the present analysis

When choosing parameter settings, we wanted to have maximum comparability to the literature and maximum comparability within the article. Therefore, we applied the algorithms first as described in the literature (heterogeneous parameter settings), and second with entirely equal parameter settings and output metric (homogeneous parameter settings). If no standard was available in the literature, we chose parameter settings that facilitated comparability of the algorithms. Since there is no common standard with respect to data transformation and smoothing procedure, we neither smoothed nor transformed the time series, to avoid information loss. Referring to the maximal time lag considered, we used a homogeneous maximal time lag of 5 s (in accordance to Ramseyer, 2011).

Heterogeneous parameter settings For the CLC or CLR, we applied a simple cross-correlation or cross-regression with

step size 0.04 s (= 1 frame). The CLC results in an absolute global correlation, the CLR in an R^2 value. For the WCC, we used a bandwidth of 10 s (= 250 frames) with a step size of 1 s (= 25 frames; Tronick et al., 1977). As a global score, an averaged/maximum global absolute correlation was used. With respect to $WCLC_{S1}$ and $WCLC_{S2}$, we applied a 1-min window as bandwidth (= 1,500 frames) with nonoverlapping windows (step size = 1 min = 1,500 frames; Paulick et al., 2018; Ramseyer & Tschacher, 2011). Since a moving median had already been applied to the raw data, we omitted the moving average of the Ramseyer/Paulick procedure in order to apply the algorithms to the same raw data. Before calculating the correlations, a z -transformation was applied. In addition, $WCLC_{S1}$ applies a minimal threshold for movement (see Grammer et al., 1999; Ramseyer & Tschacher, 2011). Both $WCLC_{S1}$ and $WCLC_{S2}$ result in an absolute averaged/maximum correlation as the output score. For $WCLC_F$ and WCLR, we used a bandwidth of 5 s (= 125 frames) with a step size of 0.04 s (= 1 frame), as well as overlapping windows (Altmann, Schoenherr, Paulick, Deisenhofer, et al., 2018). In the present study, Altmann's (2013) peak-picking algorithm was used for the $WCLC_F$ and WCLR with a time lag tolerance of one frame (which equates to 0.04 s). All synchronization intervals that lasted less than 0.4 s were removed because they were too short to display meaningful synchrony (for details, see Altmann, 2013). All algorithms had the same time series as input. All time-lagged methods had a maximum time lag of 5 s.

Homogeneous parameter setting To enhance comparability, all algorithms were used with equal settings: same time series, maximum time lag of 5 s (= 125 frames), bandwidth of 5 s (= 125 frames), and overlapping windows with a step size of 0.04 s (= 1 frame). To allow comparison of the correlations and R^2 values of the different methods and obtain a homogeneous metric, we squared the simple correlations.

We calculated various output scores for each method, when possible (average absolute correlation = average strength of synchrony, highest absolute correlation = maximum strength of synchrony, ratio of synchronous time to total duration = frequency of synchrony). The global output scores used with respect to the different TSAMs are displayed in Fig. 4.

For a criterion-based validation with a nonlinear method, we also computed synchrony scores using CRQA, implemented in R by Coco and Dale (2014). We preprocessed our time series via z -transformation and determined the optimal radius and embedded dimensions by using the `optimizeParam` function. The mean radius of all time series was 8.85, and the optimal value of the embedded dimensions was 3. We calculated recurrence rate (RR) profiles by applying the `runCrqa` function for continuous data with a window size of 125 frames (= 5 s). Afterward, we averaged the RR profiles to obtain an output score (average) representing nonlinear movement synchrony. Additionally, we calculated the mean of the maximal RR (maxrec) to obtain an output maximum score.

Statistical analysis

After video analysis with MEA and computation of synchrony indices with multiple TSAMs, we examined the association between different synchrony output scores. Descriptive statistics (mean, standard deviation) were investigated. Then we calculated bivariate Pearson product-moment correlations between the synchrony output scores in order to investigate convergent validity. Correlations lower than .35 were considered low, those between .35 and .67 moderate, and those between .68 and 1.0 high (Taylor, 1990). P values of the correlations were corrected for multiple testing by Bonferroni correction. Correlations with the CRQA score were also conducted with respect to all three output scores (average, max, freq). Then we tested the theoretical assumption that a common factor model holds. Thereby, we first tested the average and ratio/frequency

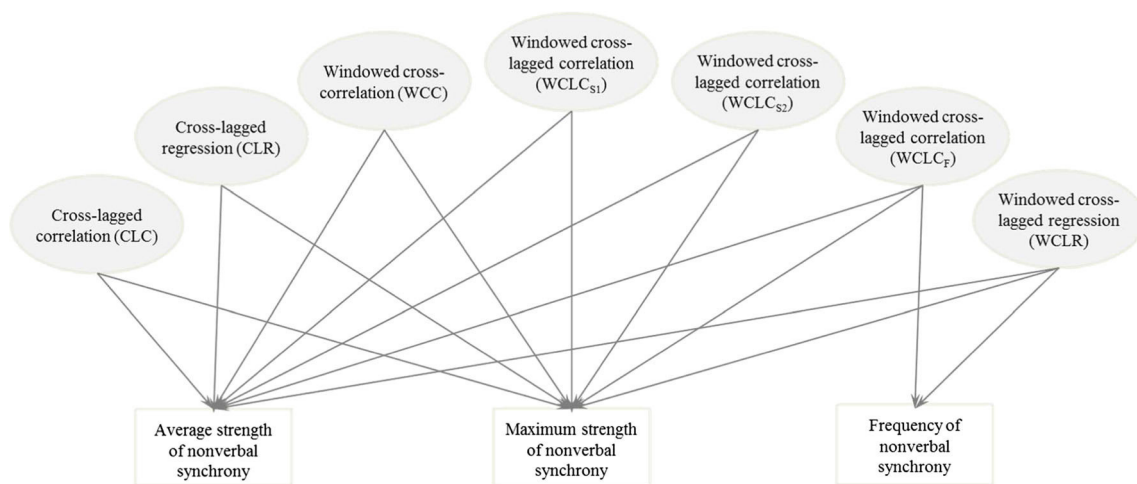


Fig. 4 Output measures for the different linear TSAMs

scores, and second we examined the common factor structure of the maximum scores, because including the correlated average and maximum scores in one analysis might have biased the factor analysis. The root mean square error of approximation (RMSEA) should be $< .05$, the comparative-fit index (CFI) $> .9$, and the Tucker–Lewis index (TLI) $> .9$ to suggest a satisfactory fit. Because the results showed that a common factor model was not appropriate, we conducted an exploratory factor analysis (maximum likelihood estimator, rotation oblimin), first for the average and ratio scores and second for the maximum scores, in *Mplus 7.4*. Thereby, we specified that the factor solution lay between one and six factors and tested different factor solutions (e.g., two-factor, three-factor, etc.) for data fit. In our exploratory factor analysis, we allowed for correlating factors.

Finally, we examined whether the output scores were equally related to a construct: interpersonal problems at the end of therapy. Interpersonal problems were measured with the IIP (Horowitz et al., 2016), which is composed of eight scales, including dominance, self-sacrifice, and social inhibition, that measure interpersonal problems. Each scale consists of four items based on a 5-point Likert scale. We conducted partial correlations while controlling for interpersonal problems at the beginning of therapy.

It should be noted that we performed each analysis on two datasets. The first dataset ($N = 84$ time series pairs) were synchrony measures based on the parameter settings suggested by the authors of the algorithms. The second dataset were synchrony measures computed when the same parameter settings were applied to each algorithm. This was done to maximize both comparability with other studies (heterogeneous settings) and comparability within our own study (homogeneous settings).

Results

Descriptive statistics

The means and standard deviations of the examined synchrony output scores and time lags are displayed in Table 2.

Correlation analysis

The correlations of different synchrony output scores are displayed in Table 3 with respect to the average strength and frequency of synchrony. The table captures correlations in which each algorithm was applied both as recommended in

Table 2 Means and standard deviations of synchrony indices and their time lag variables

| | Synchrony indices | | | Time lag | | |
|----------------------------------|-------------------|-------------|-------------|---------------|----------------|--------------|
| | Mean (SD) | | | Mean (SD) | | |
| TSAM | Average score | Max score | Ratio score | Average score | Max score | Ratio score |
| Heterogeneous parameter settings | | | | | | |
| CLC | .036 (.023) | .075 (.039) | – | 0 | – 8.05 (76.85) | – |
| CLR | .005 (.014) | .013 (.039) | – | 0 | 108.52 (23.97) | – |
| WCC | .115 (.029) | .704 (.121) | – | 0 | 0 | – |
| WCLC _{S1} | .047 (.015) | .111 (.041) | – | 0 | 0 | – |
| WCLC _{S2} | .084 (.015) | .750 (.245) | – | 0 | 0 | – |
| WCLC _F | .269 (.046) | .871 (.059) | .442 (.108) | 64.18 (2.97) | – 9.57 (73.70) | 64.18 (2.97) |
| WCLR | .252 (.041) | .854 (.055) | .444 (.108) | 65.71 (3.01) | – 2.29 (72.35) | 65.71 (3.01) |
| CRQA | .168 (.095) | .177 (.095) | – | 0 | 0 | – |
| Homogeneous parameter settings | | | | | | |
| CLC | .002 (.003) | .007 (.009) | – | 0 | – 8.05 (76.85) | – |
| CLR | .005 (.014) | .013 (.039) | – | 0 | 108.52 (23.97) | – |
| WCC | .018 (.008) | .738 (.148) | – | 0 | 0 | – |
| WCLC _{S1} | .008 (.002) | .042 (.032) | – | 0 | 0 | – |
| WCLC _{S2} | .038 (.008) | .997 (.006) | – | 0 | 0 | – |
| WCLC _F | .269 (.046) | .871 (.059) | .442 (.108) | 64.18 (2.97) | – 9.57 (73.70) | 64.18 (2.97) |
| WCLR | .252 (.041) | .854 (.055) | .444 (.108) | 65.71 (3.01) | – 2.29 (72.35) | 65.71 (3.01) |
| CRQA | .168 (.095) | .177 (.095) | – | 0 | 0 | – |

Heterogeneous parameter settings according to the literature. TSAM = time series analysis method; SD = standard deviation; CLC = cross-lagged correlation; CLR = cross-lagged regression; WCC = windowed cross-correlation; WCLC = windowed cross-lagged correlation (WCLC_S, with respect to strength of synchrony; WCLC_F, with respect to frequency of synchrony); WCLR = windowed cross-lagged regression. $N = 84$, all values in frames

Table 3 Correlations between *average* output scores of nonverbal synchrony assessed with various TSAMs

| | CLC | CLR | WCC | WCLC _{S1} | WCLC _{S2} | WCLC _F | WCLR | CRQA | WCLC _F | WCLR |
|--------------------|-------|-------|-------|--------------------|--------------------|-------------------|-------|-------|-------------------|-------|
| | | | | | | | | | freq | freq |
| CLC | .960* | .001 | .250 | .194 | .256 | .132 | .110 | .272 | .193 | .236 |
| CLR | .004 | 1 | -.117 | .067 | -.026 | -.033 | -.011 | -.110 | -.155 | -.155 |
| WCC | .268 | -.107 | .954* | .271 | .866* | .770* | .755* | .768* | .874* | .881* |
| WCLC _{S1} | .349* | .037 | .607* | .713* | .292 | .217 | .190 | .172 | .137 | .146 |
| WCLC _{S2} | .528* | .085 | .734* | .692* | .781* | .791* | .772* | .681* | .733* | .749* |
| WCLC _F | .147 | -.033 | .827* | .492* | .588* | 1 | .984* | .604* | .753* | .749* |
| WCLR | .123 | -.011 | .803* | .460* | .565* | .984* | 1 | .601* | .767* | .757* |
| CRQA | .269 | -.110 | .738* | .454* | .555* | .604* | .601* | 1 | .769* | .777* |
| WCLC _F | .204 | -.155 | .880* | .443* | .550* | .753* | .767* | .769* | 1 | .987* |
| freq | | | | | | | | | | |
| WCLR | .239 | -.155 | .892* | .464* | .579* | .749* | .757* | .777* | .987* | 1 |
| freq | | | | | | | | | | |

R^2 upper triangle (gray shaded): correlations of different synchrony indices (R^2) based on homogeneous parameter settings; lower triangle: correlations of different synchrony indices (corelations/ R^2) based on heterogeneous parameter settings from literature. CLC = cross-lagged correlation; CLR = cross-lagged regression; WCC = windowed cross-correlation; WCLC = windowed cross-lagged correlation (WCLC_S, with respect to strength of synchrony; WCLC_F, with respect to frequency of synchrony); WCLR = windowed cross-lagged regression; CRQA = Cross recurrence quantification analysis. * Significant correlations, Bonferroni-adjusted alpha level = .001, freq = frequency of synchrony/ratio

the literature (lower triangle) and with entirely equal parameter settings (upper triangle, gray-shaded). For the sake of clarity, we only present results referring to the average scores and ratio score, here. The other correlations (maximum output scores) are displayed in the [Appendix](#). High correlations between the scores of the more local methods indicate a substantial association between parameter settings, with one exception: The correlations with WCLC_{S1} using equal parameter settings are nonsignificantly different from zero. The global methods are neither correlated with each other nor highly correlated with the scores of the local methods. Two associations are illustrated by scatterplots (Fig. 5).

With respect to the maximum scores, most of the output scores were nonsignificantly correlated. Significant associations were only shown referring to the frequency measures (WCLC_F, WCLR) and between the WCC and frequency measures and CRQA. Additionally, WCLC_{S1} (max) and CLC (max) were

associated (see the Appendix, Table 6). The maximum and average scores for each algorithm were associated with each other mostly, but ranged from $r = -.210$ to $r = .999^*$ (see the Appendix, Tables 7 and 8; here and below, asterisks indicate results significant at the alpha level given in particular tables).

Exploratory factor analysis

With respect to our two parameter settings, we conducted two factor analyses (A: output scores resulting with heterogeneous parameter settings [according to literature], B: output scores resulting with homogeneous parameter settings). We included average scores and ratio/frequency scores only because the highly correlated average and maximum scores might have biased the factor analysis. For neither of the settings, a common factor model fit the data (A: $\chi^2 = 380.25$, $df = 35$, $p < .01$,

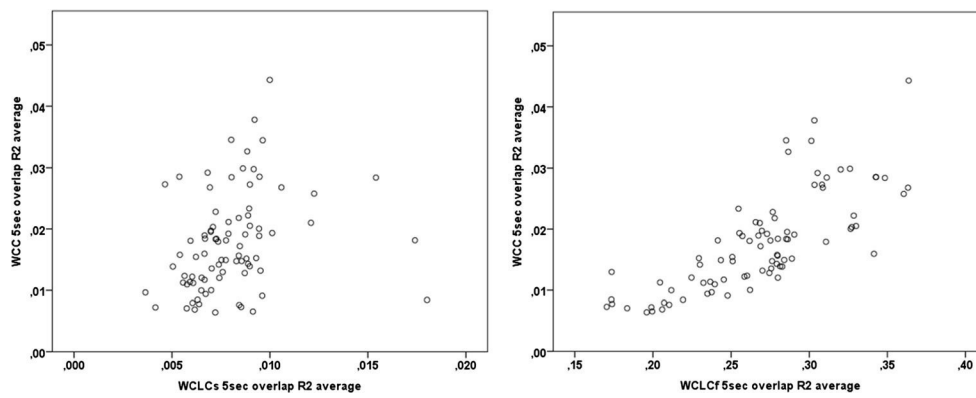


Fig. 5 Scatterplots showing the correlations between WCC (average) and $WCLC_{S1}$ (average), on the left, and $WCLC_F$ (average), on the right

RMSEA = 0.34, CFI = 0.67, TLI = 0.58; B: $\chi^2 = 331.87$, $df = 35$, $p < .01$, RMSEA = 0.32, CFI = 0.71, TLI = 0.63).

For the exploratory analyses, the results showed that a three-factor model fit the data for both parameter settings (A: $\chi^2 = 19.24$, $df = 18$, $p = .38$, RMSEA = 0.029, CFI = 0.999, TLI = 0.997, B: $\chi^2 = 9.80$, $df = 18$, $p = .94$, RMSEA < 0.00, CFI = 1.00, TLI = 1.02). The factor loadings of the three-factor solution are displayed in Table 4. The correlations were, between Factors 1 and 2, $r = .742^*$ for the dataset generated with heterogeneous parameter settings for the algorithms (respectively, $r = .726^*$ for homogeneous parameter settings), between Factors 1 and 3, $r = .576^*$ (respectively, $r = .737^*$), and between

Factors 2 and 3, $r = .564^*$ (respectively, $r = .702^*$). For maximum scores, no adequate factor solution was found.

Post-hoc analysis On the basis of the three-factor solution of the exploratory factor analysis (EFA) and our main hypothesis (common factor for synchrony), we also conducted a confirmatory factor analysis to test whether the three factors loaded on a single factor. Therefore, we specified three latent endogenous variables based on the factor loadings and one superior latent synchrony variable. Referring to the heterogeneous parameter settings, we found an adequate model by excluding the WCC and CLR (which had double loadings in EFA). The

Table 4 Factor loadings (geomin-rotated loadings) of the exploratory factor analysis

| | Heterogeneous parameter settings | | | Homogeneous parameter settings | | |
|------------------|----------------------------------|----------|----------|--------------------------------|----------|----------|
| | Factor 1 | Factor 2 | Factor 3 | Factor 1 | Factor 2 | Factor 3 |
| CLC (av) | -.041 | -.128 | .630* | .034 | -.228 | .448* |
| CLR (av) | -.296* | -.010 | .263* | -.364* | .251 | .028 |
| WCC (av) | .561* | .232* | .283* | .422* | .017 | .592* |
| $WCLC_{S1}$ (av) | .026 | .097 | .638* | -.266 | .002 | .538* |
| $WCLC_{S2}$ (av) | .019 | .006 | .967* | -.003 | .224 | .767* |
| $WCLC_F$ (av) | -.013 | 1.038* | .006 | .003 | .882* | .126* |
| $WCLC_F$ (freq) | 1.018* | .000 | -.042 | .969* | .041 | .000 |
| WCLR (av) | .140 | .852* | -.002 | .037 | .990* | -.009 |
| WCLR (freq) | .996* | -.009 | .007 | .920* | .011 | .080 |
| CRQA (av) | .689* | -.005 | .163 | .542* | -.062 | .370* |

Heterogeneous parameter settings according to the literature. CLC = cross-lagged correlation; CLR = cross-lagged regression; WCC = windowed cross-correlation; $WCLC$ = windowed cross-lagged correlation ($WCLC_S$, with respect to strength of synchrony; $WCLC_F$, with respect to frequency of synchrony); WCLR = windowed cross-lagged regression. The highest factor loadings are shaded in gray. *Significant at a 5% alpha level, freq = frequency

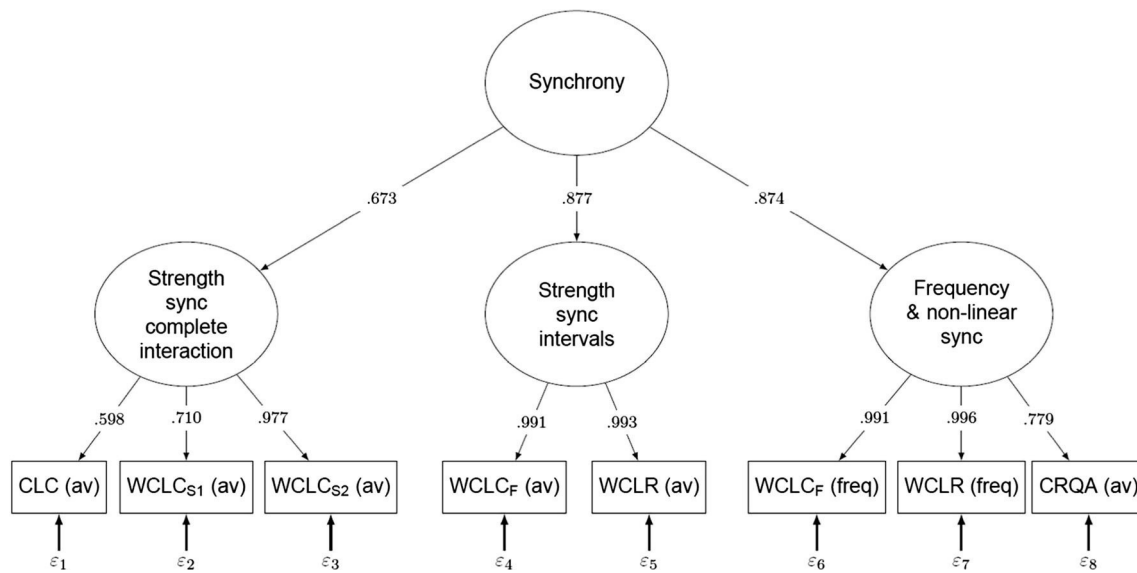


Fig. 6 Common factor model synchrony (heterogeneous parameter settings)

model showed an excellent fit ($\chi^2 = 7.99$, $df = 18$, $p = .98$, $RMSEA < 0.001$, $CFI = 1.000$, $TLI = 1.018$). The model with significant (standardized) path coefficients is displayed in Fig. 6. Regarding the homogeneous parameter settings, no adequate and converging model was found.

Criterion-based validation: Correlations with interpersonal problems

Correlations of the output scores with interpersonal problems at the end of therapy (while controlling for initial interpersonal problems) are displayed in Table 5. The hypothesized negative association between nonverbal synchrony and interpersonal problems posttherapy was found with respect to WCC (average), WCLCF_F

(ratio), and WCLR (ratio), which showed significant correlations. Additionally, marginally significant associations were found with regard to WCLCF (average) and WCLR (average).

Discussion

The aim of the present study was to examine the convergent validity of linear TSAMs for the assessment of nonverbal synchrony. Besides the diversity of TSAMs, most of these algorithms provide more than one output score—for instance, an average and a maximum score. Therefore, various output scores per TSAM were calculated when possible. In the literature, these different output scores and TSAMs are all used to

Table 5 Partial correlations between output scores and interpersonal problems at the end of therapy, with p values

| Output score | Heterogeneous parameter settings | | Homogeneous parameter settings | |
|------------------------------|----------------------------------|-----------|--------------------------------|-----------|
| | r IIP _{Post} | p Value | r IIP _{Post} | p Value |
| CLC (average) | -.029 | .799 | -.090 | .419 |
| CLR (average) | .076 | .497 | .076 | .497 |
| WCC (average) | -.268* | .015 | -.305* | .005 |
| WCLCS ₁ (average) | -.179 | .108 | -.086 | .444 |
| WCLCS ₂ (average) | -.123 | .272 | -.147 | .187 |
| WCLCF (average) | -.204 | .066 | -.204 | .066 |
| WCLCF (ratio) | -.262* | .018 | -.262* | .018 |
| WCLR (average) | -.205 | .065 | -.205 | .065 |
| WCLR (ratio) | -.282* | .010 | -.282* | .010 |
| CRQA (average) | -.138 | .215 | -.138 | .215 |

Heterogeneous parameter settings according to the literature. r = correlation; CLC = cross-lagged correlation; CLR = cross-lagged regression; WCC = windowed cross-correlation; WCLC = windowed cross-lagged correlation (WCLCS₁; with respect to strength of synchrony; WCLCF_F; with respect to frequency of synchrony); WCLR = windowed cross-lagged regression; CRQA = Cross recurrence quantification analysis. * Significant correlations at the 5% alpha level

assess nonverbal synchrony. Since it is not clear if all output scores measure the same construct, all TSAMs were applied to an identical dataset of time series pairs. We conducted all analyses with a set of heterogeneous parameter settings (according to literature) and homogeneous parameter settings (highest comparability between algorithms within this study).

The present study was able to demonstrate that not all output scores that are used to calculate synchrony are correlated. Especially, global and local TSAMs measure different facets of synchrony. Whereas global TSAMs assume the interrelatedness of both interacting partners to be stable (i.e., person A always influences person B, or vice versa), local TSAMs operationalize interrelatedness dynamically with varying leading and pacing. Therefore, it is plausible for global and local TSAMs not to be associated with each other. The correlations largely support this inference. Furthermore, there is no evidence that TSAMs that have the same methodical approach (e.g., correlational vs. regressive methods) would inevitably assess the same construct. Additionally, we found no common construct underlying all TSAMs with either of the two parameter settings. The results of the exploratory factor analyses suggest a three-factor solution. However, we were able to show a common factor as a latent second-order variable of these three factors by excluding WCC and CLR for the heterogeneous parameter settings.

Multifactor structure of synchrony measures

The examined three factors of the EFA incorporated with the results of the correlations are described in the following paragraphs.

Synchrony Factor 1 The first factor is formed by the $WCLC_F$ (ratio), $WCLR$ (ratio), and $CRQA$ (average) measures. In line with that, all three output scores are highly correlated (ranging from $r = .77$ to $r = .99$). Both linear measures are frequency measures that capture the ratio of time that was synchronized to the total duration of the sequence. Correlations are very high indicating nearly equal scores. The facet of synchrony that is measured differs enormously in comparison to the other (strength) measures. Apparently, the frequency measure is highly associated with the nonlinear output score ($r = .77$). The result of the $CRQA$ is the recurrence rate indicating the frequency of revisiting states of both phase trajectories. The construct can be described as the *frequency of synchrony*.

Synchrony Factor 2 This factor incorporates the average output scores of the $WCLC_F$ and $WCLR$. With respect to the $WCLC_F$ and $WCLR$, a peak-picking algorithm is used, which identifies the start and end points of synchronization intervals as well as synchrony strength within the identified intervals. Their output values are based on the identified synchronization intervals. Intervals that do not show synchrony are

neglected. The construct can best be described as the *strength of synchrony within the identified synchronization intervals*. The strength of synchrony in intervals (Factor 2) and the frequency of synchrony (Factor 1) are both measured by $WCLC_F$ and $WCLR$ and are therefore highly associated; however, they are not equivalent. If synchrony is highly frequent, it is not, by default, always strong. The strength of the association is partially determined by the peak height of both time series. If both are high and similar, high strength is identified. In the context of psychotherapy, this would mean that both persons gesture in a large or a very space-consuming fashion. If the magnitude of the persons' movements is very different within their space, strength will be lower, although the interval will also be identified as showing synchrony.

The results showed that the $WCLC_F$ and $WCLR$ measures correlated very highly with each other ($r = .98$). Since $WCLR$ was developed on the basis of cyclical data (Altmann, 2011, 2013), it is probably advantageous for this type of data as compared to $WCLC_F$. With respect to the noncyclical data in the present study, the scores were nearly equivalent. Therefore, $WCLC_F$ seems to be preferable, because the computational expense of $WCLC_F$ is much lower than that of $WCLR$.

Synchrony Factor 3 Factor 3 of the EFAs incorporates the CLC output score and the $WCLC_{S1}$ and $WCLC_{S2}$ output scores. In comparison to the associations within the other factors, the correlations between CLC and the other scores of this factor are only moderate. This may be explained because CLC is a global measure and both forms of $WCLC_S$ are local measures. However, their belonging to this factor may be explained by the large bandwidth of 1 min used in both forms of $WCLC_S$. Since the total duration of the sequence was 15 min, the results of $WCLC_S$ are similar to those of the CLC and are assigned to the same factor. Both $WCLC_S$ measures are highly associated ($r > .7$), which is plausible because both measures differ only in the used the Grammer et al. (1999) minimal movement threshold (see the [Method](#) section). The scores of Factor 3 quantify the strength of synchrony and aggregate the value with respect to the total interaction. Thus, the r or R^2 values of nonsignificant sequences (sequences without movement synchrony) and movement synchrony intervals are included in the aggregation. The construct is best described as the *strength of synchrony of the total interaction*. The main difference between this factor and the $WCLC_F$ and $WCLR$ measures is the peak-picking algorithm. The clear advantage of $WCLC_F$ and $WCLR$ lies in their determination of synchronization intervals, which also clearly distinguishes Factors 1 and 2 from Factor 3.

Special cases CLR and WCC Our results showed that the CLR is not related to any other measure, suggesting that CLR maps a completely different construct than do the other TSAMs. This is emphasized by the finding that no converging and adequate

model was found by incorporating CLR. We do not recommend using CLR as a synchrony measure.

The measure WCC is related to the output scores of all local TSAMs. This corresponds to the results of the factor analysis showing significant factor loadings for more than one factor. Apparently, WCC is related to each facet of synchrony—especially to the frequency of synchrony and the strength of synchrony of the total interaction. This might be plausible because sequences may have similar proportions of matching and echoing. Therefore, the assessed matching of the WCC correlates highly with the sum of echoing and matching of the other algorithms. This does not necessarily imply that it is beneficial not to include a time lag; rather, it presents the opportunity to estimate the general level of synchrony of an interaction using WCC.

The existence of a diversity of procedures and low concordance of scales is also found in other areas of psychological research, such as attachment research. The investigation of the attachment–outcome relation is complicated by the fact that measures of attachment are quite heterogeneous (Bouthillier, Julien, Dubé, Bélanger, & Hamelin, 2002; Kirchmann, Fenner, & Strauß, 2007; Manes et al., 2016; Roisman et al., 2007) and show weak, if any, convergence. Thus, the comparability of studies is reduced. Evidence related to attachment as a predictor of outcome in specific psychological treatments remains unclear (Manes et al., 2016). Recently, it has been assumed that (i) the methods measure different aspects of the construct and (ii) attachment is therefore not unidimensional but, rather, multidimensional. Referring to the results of the present study, the same can be said of nonverbal synchrony: Nonverbal synchrony does not seem to be a unidimensional but, rather, a multidimensional construct. However, the different facets can be related to a superior construct of nonverbal synchrony.

Criterion-based validations

In the current literature, it is often discussed that nonlinear methods better reflect interpersonal interactions. Therefore, we also used the most commonly applied nonlinear method: cross-recurrence quantification analysis. Different linear output scores are differently associated with the output scores of CRQA: The average score of the CRQA shows high correlations with the local TSAMs. Only the $WCLC_{S1}$ score with homogeneous parameter settings seems to be an exception. Therefore, the $WCLC_{S1}$ should not be applied with small bandwidth and overlapping windows. To summarize, in comparison to some linear models, nonlinear models do not necessarily result in completely different synchrony indices. This result can be considered a validation of the local TSAMs with the parameter settings that are recommended in the literature.

When examining empirical research questions with TSAMs, which are related to different synchrony constructs,

inconsistent findings may be found. Tronick et al. (1977), for instance, showed significant nonverbal synchrony between mother–child dyads using WCC, whereby Gottman and Ringland (1981) did not find this association when reanalyzing the very same dataset using CLR. We also conducted partial correlations between ten nonverbal synchrony output scores and therapy outcome (IIP). The assumed significant negative relationship between synchrony and the outcome was only observable when using three of the ten output scores. This would mean that a high level of synchrony at the beginning of therapy is associated with fewer interpersonal problems at the end of the therapy. The scores showing this relationship can be assigned to the same synchrony facet (frequency of synchrony), supporting the idea of different synchrony facets. The association between the IIP results and synchrony was also present marginally significantly with respect to synchrony in the intervals. However, total synchrony was only descriptively and nonsignificantly associated with interpersonal problems. This is inconsistent with the findings of Ramseyer and Tschacher (2011), who reported a significant association. However, Paulick et al. (2018) did not find a linear association between IIP and total synchrony. Assuming an association between synchrony and interpersonal problems, the frequency measures of the $WCLC_F$ and $WCLR$ and the WCC (average) are the only valid output scores.

Practical implications

Since the output scores do not measure the same nonverbal synchrony construct, the question arises, which of the TSAMs and their output scores best measure nonverbal synchrony. The present study is unable to answer this question completely, because we have no “true value” of nonverbal synchrony to which we can compare the output scores of the applied TSAMs. In the future it may be beneficial to create a database that includes sequences with and without nonverbal synchrony. To date, only two brief studies exist that have compared output scores with simulated (Altmann, 2013) or human-rated (Paxton & Dale, 2013b) nonverbal synchrony. Therefore, extensive studies will be needed. Additionally, the criterion to which the results of an algorithms are compared is important: If the aim is to measure an aspect of synchrony that is associated with interpersonal problems, the results of algorithms that correlate highly with interpersonal problems are best. If the aim is to measure a synchrony facet that is similar to human-rated synchrony, human-rated sequences will be needed as a criterion against which to compare the algorithms. In addition, there have also been attempts to define synchrony empirically in comparison to pseudosynchrony. Therefore, the choice of the best-suited algorithms is inherently connected to the criterion.

In this context, however, it should be emphasized that the results of the TSAMs presented are dependent on the parameter settings. These include, for example, the degree of

smoothing, transformation, and bandwidth. In the present study, standard settings from the literature and equal parameter settings were used. The results were more consistent using the parameter settings from the literature. Therefore, future studies should investigate which parameter settings really are best suited to optimally capture nonverbal synchrony. Since, for example, bandwidth has a high impact on results, which is why we do not recommend the direct comparison of results that were conducted using different bandwidths, even when identical TSAMs were used. Nevertheless, in line with Delaherche and Chetouani (2010), we recommend the application of local methods, because they are more likely to cope with statistical challenges such as nonstationarity, and also take zeitgeber changes or a changing time lag into account.

With respect to the present study, it can be said that $WCLC_F$ or $WCLR$ can be used to assess the strength of synchrony in predefined intervals or the frequency of synchrony in an interaction. CLC , $WCLC_{S1}$, and $WCLC_{S2}$ can be applied to measure the strength of synchrony of a total interaction. CLR measures are not recommended, because they are not comparable to the other TSAMs assessing nonverbal synchrony. If one is solely interested in the amount of synchrony (not in pacing, leading, or time-lag-related variables), the WCC (average) is a valid output score with which to estimate the amount of interrelatedness between two time series. Nevertheless, that a common construct of synchrony may be found for parameter settings from the literature, future studies should explicitly define and characterize the facet of synchrony they intend to measure.

Strengths and limitations

A limitation is that the TSAM parameters were not systematically varied (e.g., in terms of time lag or bandwidth). This indicates that our results cannot be generalized to algorithms using other parameter settings. Since parameter settings are an important source of variance between the results of different TSAMs, we can only draw conclusions about the presented algorithms *with* their parameter settings. In addition, no correction of spurious correlations, apart from the correction for autocorrelation, was applied. Prewhitening can be performed to reduce bias due to autocorrelation, as was shown by R. T. Dean and Dunsmuir (2016). However, autocorrelation should have influenced all methods to a comparable extent (except for the methods that include autocorrelation within their model specification, such as CLR and $WCLR$). Furthermore, we did not control for any further spurious correlations that might have been caused by randomly occurring synchrony. Different methods exist by which spurious correlations can be controlled for. One possibility is the use of surrogate/virtual pairs—that is, time series pairs that are split and randomly recombined (Louwerse et al., 2012; Moulder, Boker, Ramseyer, & Tschacher, 2018). Synchrony determined within

these surrogate pairs can be used as a baseline to evaluate the meaningfulness of genuine synchrony. Another opportunity to build a baseline is to randomly shuffle the data points (Louwerse et al., 2012) or windows (Ramseyer & Tschacher, 2010) of one time series. Altmann (2013) and Gottman and Ringland (1981) proposed a parametric test to solve this issue. Another possibility is to increase the cutoff for distinguishing between randomly occurring and meaningful synchrony (R. T. Dean & Dunsmuir, 2016). However, the lack of control for spurious correlations should also have a comparable effects on all TSAMs. Additionally, our study only compared algorithms assuming a linear relationship between the interacting persons. In further analyses, it may also be interesting to investigate other dependencies, by using spectral analysis, for example. Additionally, the inconsistent results with respect to the synchrony–outcome association can also be attributed to the limited sample size or to varying parameter settings. Valid parameter settings for each TSAM have to be examined in order to optimally map the corresponding synchrony construct. Overall, it can be assumed that a method effect exists, which explains the contrary results. For future studies, it will be important to specify exactly which method (including which parameter settings—e.g., for bandwidth) was used and which facet of synchrony is being addressed. Regarding the confirmatory factor analysis, it should be noted that typically exploratory and confirmatory analyses are run on different samples. However, due to our limited sample size, this was not possible, which should also be addressed in future studies.

One of the strengths of this study is that different algorithms were used to measure the linear dependency of two time series using an identical dataset. On the basis of the different algorithms and output scores, we systemized the approaches to different synchrony facets. Thereby, we used (1) the parameter settings recommended in the literature and (2) equal parameter settings across all TSAMs. In addition, the study draws attention to the fact that the choice of the algorithm and parameters is essential to the subsequent analyses (e.g., association analysis of nonverbal synchrony and therapeutic success). We applied the methods to a sample from a psychotherapeutic setting—that is, to real-world data—to ensure comparability of the results and their statistical challenges to previous studies (Galbusera et al., 2018; Paulick et al., 2018; Ramseyer & Tschacher, 2011). Thus, we are confident of the generalizability of our results to this field of research. However, a replication of our findings based on another dataset is advisable.

Conclusion and future directions

The present study shows that confidence in the convergent validity of TSAMs based on the apparent fit of the construct and method is critical. In the literature, different methods have been presented to determine nonverbal synchrony. Thus, the facet of synchrony that is being measured depends on the

algorithm applied and the output score used. Therefore, comparing study results calculated with TSAMs that measure different facets is critical. However, with the parameter settings used in the literature, we found a superior latent factor synchrony for most of the TSAMs presented.

In the future, the construct to be measured should be explicitly defined, and the chosen method should also be checked to determine whether the defined construct is actually being measured. Further validation studies should be conducted to measure content validity. Human ratings of synchrony should be compared with the results obtained by these

methods. Furthermore, which method or parameter combinations lead to valid results will need to be investigated.

Author note This study was funded by the German Research Foundation (DFG project “Timing of Nonverbal Patient-Therapist-Interaction and Therapeutic Success of Social Phobic Patients (TIMPATHIN),” funding number GZ: STR 306/28-1 and LU 660/8-1) and used video recordings from the SOPHONET treatment study, funded by the German Federal Ministry of Education and Research (BMBF, FKZ 01GV0607). We cordially thank U. Willutzki, S. Herpertz, J. Hoyer, P. Joraschky, W. Hiller, E. Leibing, M. Beutel, and F. Leichenring for providing video recordings for the project. We thank K. Boyle, a native speaker of English, for proof reading.

Appendix: Results—Maximum output score

The correlations of the maximum output scores are displayed in Table 6. Tables 7 and 8 show the correlations between maximum and average scores, respectively ratio scores (Table 7 with heterogeneous parameters, Table 8 with homogeneous parameter settings).

Table 6 Correlations between *maximum* output scores of nonverbal synchrony assessed with various TSAMs

| | CLC | CLR | WCC | WCLC _{S1} | WCLC _{S2} | WCLC _F | WCLR | CRQA |
|--------------------|-------|-------|-------|--------------------|--------------------|-------------------|-------|--------|
| CLC | .967* | .012 | .088 | .123 | -.095 | .065 | .043 | .155 |
| CLR | .018 | 1 | .100 | -.067 | .046 | -.176 | -.161 | -.107 |
| WCC | .073 | .157 | .595* | .126 | -.076 | .406* | .309 | .411* |
| WCLC _{S1} | .590* | -.081 | .248 | .255* | .014 | .008 | -.027 | .193 |
| WCLC _{S2} | .170 | .101 | -.014 | .165 | .081 | -.011 | -.068 | -.558* |
| WCLC _F | .085 | -.176 | .293 | .125 | -.154 | 1 | .779* | .306 |
| WCLR | .094 | -.161 | .155 | .142 | -.170 | .779* | 1 | .261 |
| CRQA | .216 | -.107 | .355* | .215 | -.190 | .306 | .261 | 1 |

R^2 values in the upper triangle (gray-shaded) are based on algorithms with exactly equal parameter settings, and correlations in the lower triangle are based on algorithms with the settings recommended from the literature. CLC = cross-lagged correlation, CLR = cross-lagged regression, WCC = windowed cross-correlation, WCLC = windowed cross-lagged correlation (WCLC_S, with respect to the strength of synchrony; WCLC_F, with respect to the frequency of synchrony), WCLR = windowed cross-lagged regression, CRQA = cross-recurrence quantification analysis. * Significant correlations, Bonferroni-adjusted alpha level = .001

Table 7 Correlations between the *average* output scores and the *maximum* output scores of nonverbal synchrony assessed with various TSAMs—Heterogeneous parameter settings

| | CLC (av) | CLR (av) | WCC (av) | WCLC _{S1} (av) | WCLC _{S2} (av) | WCLC _F (av) | WCLR (av) | CRQA (av) | WCLC _F (ratio) | WCLR (ratio) |
|---------------------------|-------------|-------------|-------------|----------------------------|----------------------------|---------------------------|--------------|--------------|------------------------------|-----------------|
| CLC (max) | .817* | .010 | .169 | .368 | .419* | .109 | .078 | .216 | .144 | .161 |
| CLR (max) | .013 | .999* | -.101 | .038 | .095 | -.029 | -.007 | -.105 | -.152 | -.150 |
| WCC (max) | .023 | .154 | .515* | .450* | .389* | .433* | .366 | .347 | .364 | .380* |
| WCLC _{S1} (max) | .410* | -.080 | .252 | .596* | .502* | .162 | .123 | .219 | .199 | .186 |
| WCLC _{S2} (max) | .009 | .104 | -.437* | -.040 | -.210 | -.420* | -.433* | -.175 | -.442* | -.448* |
| WCLC _F (max) | .085 | -.181 | .506 | .325 | .369 | .631 | .626 | .298 | .437 | .417 |
| WCLR (max) | .203 | -.167 | .418* | .318 | .378* | .573* | .569* | .254 | .362 | .355 |
| CRQA (max) | .271 | -.113 | .752* | .457* | .567* | .616* | .613* | .999* | .780* | .789* |
| WCLC _F (ratio) | .204 | -.155 | .880* | .443* | .550* | .753* | .767* | .769* | 1 | .987* |
| WCLR (ratio) | .239 | -.155 | .892* | .464* | .579* | .749* | .757* | .777* | .987* | 1 |

CLC = cross-lagged correlation, CLR = cross-lagged regression, WCC = windowed cross-correlation, WCLC = windowed cross-lagged correlation (WCLC_S, with respect to the strength of synchrony; WCLC_F, with respect to the frequency of synchrony), WCLR = windowed cross-lagged regression, CRQA = cross-recurrence quantification analysis. * Significant correlations, Bonferroni-adjusted alpha level = .001

Table 8 Correlations between the *average* output scores and the *maximum* output scores of nonverbal synchrony assessed with various TSAMs—Homogeneous parameter settings: bandwidth 5 s, overlapping windows

| | CLC (av) | CLR (av) | WCC (av) | WCLC _{S1} (av) | WCLC _{S2} (av) | WCLC _F (av) | WCLR (av) | CRQA (av) | WCLC _F (ratio) | WCLR (ratio) |
|---------------------------|-------------|-------------|-------------|----------------------------|----------------------------|---------------------------|--------------|--------------|------------------------------|-----------------|
| CLC (max) | .687* | .004 | .071 | .183 | .111 | .056 | .025 | .157 | .101 | .115 |
| CLR (max) | .010 | .999* | -.113 | .072 | -.020 | -.029 | -.007 | -.105 | -.152 | -.150 |
| WCC (max) | .037 | .095 | .431* | .266 | .400* | .503* | .479* | .405* | .448* | .448* |
| WCLC _{S1} (max) | .059 | -.067 | .207 | .283 | .211 | .141 | .113 | .199 | .132 | .143 |
| WCLC _{S2} (max) | -.274 | .050 | -.386* | .047 | -.216 | -.126 | -.147 | -.566* | -.374* | -.368 |
| WCLC _F (max) | .037 | -.181 | .410* | .289 | .463* | .631* | .626* | .298 | .437* | .417* |
| WCLR (max) | .181 | -.167 | .359 | .272 | .428* | .573* | .569* | .254 | .362 | .355 |
| CRQA (max) | .273 | -.113 | .778* | .167 | .694* | .616* | .613* | .999* | .780* | .789* |
| WCLC _F (ratio) | .193 | -.155 | .874* | .137 | .733* | .753* | .767* | .769* | 1 | .987* |
| WCLR (ratio) | .236 | -.155 | .881* | .146 | .749* | .749* | .757* | .777* | .987* | 1 |

CLC = cross-lagged correlation, CLR = cross-lagged regression, WCC = windowed cross-correlation, WCLC = windowed cross-lagged correlation (WCLC_S, with respect to the strength of synchrony; WCLC_F, with respect to the frequency of synchrony), WCLR = windowed cross-lagged regression, CRQA = cross-recurrence quantification analysis. * Significant correlations, Bonferroni-adjusted alpha level = .001

References

- Allsop, J. S., Vaitkus, T., Marie, D., & Miles, L. K. (2016). Coordination and collective performance: Cooperative goals boost interpersonal synchrony and task outcomes. *Frontiers in Psychology*, 7, 1462. <https://doi.org/10.3389/fpsyg.2016.01462>
- Altmann, U. (2010). Interrater-Reliabilität = 1 in Videostudien? Automatisierte Erhebung von Nonverbalität in einem Experiment zur Kooperation von Schülern [Automated coding of nonverbal behavior in an experiment on the cooperation of students]. *Beiträge zur Erziehungswissenschaftliche Forschung-Nachhaltige Bildung*, 5, 261–267.
- Altmann, U. (2011). Investigation of movement synchrony using windowed cross-lagged regression. In A. Esposito, A. Vinciarelli, K. Vicsi, C. Pelachaud, & A. Nijholt (Eds.), *Analysis of verbal and nonverbal communication and enactment: The processing issues* (pp. 335–345). Berlin, Germany: Springer. https://doi.org/10.1007/978-3-642-25775-9_31
- Altmann, U. (2013). *Synchronisation nonverbalen Verhaltens: Weiterentwicklung und Anwendung zeitreihenanalytischer Identifikationsverfahren*. Berlin, Germany: Springer.
- Altmann, U., Schoenherr, D., Paulick, J., Deisenhofer, A.-K., Schwartz, B., Rubel, J., ... Strauss, B. M. (2018). *Timing of nonverbal patient-therapist interaction and therapeutic success of social phobic patients*. Manuscript in preparation.
- Altmann, U., Schoenherr, D., Paulick, J., Knitter, L., Worrack, S., Schiefele, A.-K., ... Strauss, B. M. (submitted). *Introduction, practical guide, and validation study for measuring body movements using motion energy analysis*. Manuscript submitted for publication.
- Bailenson, J. N., & Yee, N. (2005). Digital chameleons: Automatic assimilation of nonverbal gestures in immersive virtual environments. *Psychological Science*, 16, 814–819. <https://doi.org/10.1111/j.1467-9280.2005.01619.x>
- Beck, A. T., Ward, C. H., Mendelson, M., Mock, J., & Erbaugh, J. (1961). An inventory for measuring depression. *Archives of General Psychiatry*, 4, 561–571. <https://doi.org/10.1001/archpsyc.1961.01710120031004>
- Bernieri, F. J. (1988). Coordinated movement and rapport in teacher-student interactions. *Journal of Nonverbal Behavior*, 12, 120–138. <https://doi.org/10.1007/BF00986930>
- Bernieri, F. J., & Rosenthal, R. (1991). Interpersonal coordination: Behavior matching and interactional synchrony. In R. S. Feldman & B. Rimé (Eds.), *Fundamentals of nonverbal behavior* (pp. 401–432). New York: Cambridge University Press.
- Bilakhia, S., Petridis, S., Nijholt, A., & Pantic, M. (2015). The MAHNOB Mimicry Database: A database of naturalistic human interactions. *Pattern Recognition Letters*, 66, 52–61. <https://doi.org/10.1016/j.patrec.2015.03.005>
- Boker, S. M., Rotondo, J. L., Xu, M., & King, K. (2002). Windowed cross-correlation and peak picking for the analysis of variability in the association between behavioral time series. *Psychological Methods*, 7, 338–355. <https://doi.org/10.1037/1082-989X.7.3.338>
- Bouthillier, D., Julien, D., Dubé, M., Bélanger, I., & Hamelin, M. (2002). Predictive validity of adult attachment measures in relation to emotion regulation behaviors in marital interactions. *Journal of Adult Development*, 9, 291–305. <https://doi.org/10.1023/A:1020291011587>
- Brown, T. A., Barlow, D. H., & Di Nardo, P. A. (1994). *Anxiety disorders interview schedule for DSM-IV (ADIS-IV): Client interview schedule*. Albany, NY: Graywind.
- Cappella, J. N. (1996). Dynamic coordination of vocal and kinesic behavior in dyadic interaction: Methods, problems, and interpersonal outcomes. In J. H. Watt & C. A. VanLear (Eds.), *Dynamic patterns in communication processes* (pp. 353–386). Thousand Oaks, CA, US: Sage.
- Chartrand, T. L., & Bargh, J. A. (1999). The chameleon effect: The perception-behavior link and social interaction. *Journal of Personality and Social Psychology*, 76, 893–910. <https://doi.org/10.1037/0022-3514.76.6.893>
- Coco, M. I., & Dale, R. (2014). Cross-recurrence quantification analysis of categorical and continuous time series: An R package. *Frontiers in Psychology*, 5, 510. <https://doi.org/10.3389/fpsyg.2014.00510>
- Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, 52, 281–302. <https://doi.org/10.1037/h0040957>
- Davis, T. J., Kay, B. A., Kondepudi, D., & Dixon, J. A. (2016). Spontaneous interentity coordination in a dissipative structure. *Ecological Psychology*, 28, 23–36. <https://doi.org/10.1080/10407413.2016.1121737>
- Dean, D. J., Samson, A. T., Newberry, R., & Mittal, V. A. (2018). Motion energy analysis reveals altered body movement in youth at risk for psychosis. *Schizophrenia Research*, 200, 35–41. <https://doi.org/10.1016/j.schres.2017.05.035>
- Dean, R. T., & Dunsmuir, W. T. M. (2016). Dangers and uses of cross-correlation in analyzing time series in perception, performance, movement, and neuroscience: The importance of constructing transfer function autoregressive models. *Behavior Research Methods*, 48, 783–802. <https://doi.org/10.3758/s13428-015-0611-2>
- Delaherche, E., & Chetouani, M. (2010, October). *Multimodal coordination: Exploring relevant features and measures*. Paper presented at the 2nd international Workshop on Social Signal Processing, Florence, Italy.
- Delaherche, E., Chetouani, M., Mahdhaoui, A., Saint-Georges, C., Viaux, S., & Cohen, D. (2012). Interpersonal synchrony: A survey of evaluation methods across disciplines. *IEEE Transactions on Affective Computing*, 3, 349–365. <https://doi.org/10.1109/T-AFFC.2012.12>
- Duran, N. D., & Fusaroli, R. (2017). Conversing with a devil's advocate: Interpersonal coordination in deception and disagreement. *PLoS ONE*, 12, e0178140. <https://doi.org/10.1371/journal.pone.0178140>
- Dutschke, L. L., Stegmayer, K., Ramseyer, F., Bohlhalter, S., Vanbellinghen, T., Strik, W., & Walther, S. (2018). Gesture impairments in schizophrenia are linked to increased movement and prolonged motor planning and execution. *Schizophrenia Research*, 200, 42–49. <https://doi.org/10.1016/j.schres.2017.07.012>
- Feldman, R. (2007). Parent-infant synchrony and the construction of shared timing: Physiological precursors, developmental outcomes, and risk conditions. *Journal of Child Psychology and Psychiatry*, 48, 329–354. <https://doi.org/10.1111/j.1469-7610.2006.01701.x>
- Fogel, A. (1993). Two principles of communication: Co-regulation and framing. In J. Nadel & L. Camaioni (Eds.), *New perspectives in early communicative development* (pp. 9–22). London, UK: Routledge.
- Galbusera, L., Finn, M. T., & Fuchs, T. (2018). Interactional synchrony and negative symptoms: An outcome study of body-oriented psychotherapy for schizophrenia. *Psychotherapy Research*, 28, 457–469. <https://doi.org/10.1080/10503307.2016.1216624>
- Gottman, J. M., & Ringland, J. T. (1981). The analysis of dominance and bidirectionality in social development. *Child Development*, 52, 393–412. <https://doi.org/10.2307/1129157>
- Grammer, K., Honda, M., Juette, A., & Schmitt, A. (1999). Fuzziness of nonverbal courtship communication unblurred by motion energy detection. *Journal of Personality and Social Psychology*, 77, 487–508. <https://doi.org/10.1037/0022-3514.77.3.487>
- Grammer, K., Kruck, K. B., & Magnusson, M. S. (1998). The courtship dance: Patterns of nonverbal synchronization in opposite-sex encounters. *Journal of Nonverbal Behavior*, 22, 3–29. <https://doi.org/10.1023/A:1022986608835>
- Harrist, A. W., & Waugh, R. M. (2002). Dyadic synchrony: Its structure and function in children's development. *Developmental Review*, 22, 555–592. [https://doi.org/10.1016/S0273-2297\(02\)00500-2](https://doi.org/10.1016/S0273-2297(02)00500-2)

- Horowitz, L. M., Strauss, B., Thomas, A., & Kordy, H. (2016). *IIP-D Inventar zur Erfassung Interpersonaler Probleme—Deutsche Version*. [Inventory for the assessment of interpersonal problems—German version] (3rd ed.). Göttingen, Germany: Hogrefe.
- Kato, T., Takahashi, E., Sawada, K., Kobayashi, N., Watanabe, T., & Ishh, T. (1983). A computer analysis of infant movements synchronized with adult speech. *Pediatric Research*, *17*, 625–628. <https://doi.org/10.1203/00006450-198308000-00004>
- Kirchmann, H., Fenner, A., & Strauß, B. (2007). Konvergenz des Erwachsenen-Bindungsprototypen-Ratings (EBPR) mit verschiedenen Selbstbeschreibungsinstrumenten zur Erfassung von Bindungsmerkmalen. *PPmP—Psychotherapie: Psychosomatik-Medizinische Psychologie*, *57*, 334–342. <https://doi.org/10.1055/s-2007-970860>
- Kleinbub, K. R., & Ramseyer, F. (2018). rMEA: Synchrony in motion energy analysis (MEA) time-series (R package version 1.0.0.9012). <https://cran.r-project.org/web/packages/rMEA/index.html>
- Kupper, Z., Ramseyer, F., Hoffmann, H., Kalbermatten, S., & Tschacher, W. (2010). Video-based quantification of body movement during social interaction indicates the severity of negative symptoms in patients with schizophrenia. *Schizophrenia Research*, *121*, 90–100. <https://doi.org/10.1016/j.schres.2010.03.032>
- Kupper, Z., Ramseyer, F., Hoffmann, H., & Tschacher, W. (2015). Nonverbal synchrony in social interactions of patients with schizophrenia indicates socio-communicative deficits. *PLoS ONE*, *10*, e0145882. <https://doi.org/10.1371/journal.pone.0145882>
- Leichsenring, F., Salzer, S., Beutel, M. E., Herpertz, S., Hiller, W., Hoyer, J., ... Poehlmann, E. (2013). Psychodynamic therapy and cognitive-behavioral therapy in social anxiety disorder: A multicenter randomized controlled trial. *American Journal of Psychiatry*, *170*, 759–767. <https://doi.org/10.1176/appi.ajp.2013.12081125>
- Leichsenring, F., Salzer, S., Beutel, M. E., Herpertz, S., Hiller, W., Hoyer, J., ... Poehlmann, E. (2014). Long-term outcome of psychodynamic therapy and cognitive-behavioral therapy in social anxiety disorder. *American Journal of Psychiatry*, *171*, 1074–1082. <https://doi.org/10.1176/appi.ajp.2014.13111514>
- Louwerse, M. M., Dale, R., Bard, E. G., & Jeuniaux, P. (2012). Behavior matching in multimodal communication is synchronized. *Cognitive Science*, *36*, 1404–1426. <https://doi.org/10.1111/j.1551-6709.2012.01269.x>
- Manes, S., Nodop, S., Altmann, U., Gawlytta, R., Dinger, U., Dymel, W., ... Petrowski, B. (2016). Social anxiety as a potential mediator of the association between attachment and depression. *Journal of Affective Disorders*, *205*, 264–268. <https://doi.org/10.1016/j.jad.2016.06.060>
- Marwan, N., & Kurths, J. (2002). Nonlinear analysis of bivariate data with cross recurrence plots. *Physics Letters A*, *302*, 299–307. [https://doi.org/10.1016/S0375-9601\(02\)01170-2](https://doi.org/10.1016/S0375-9601(02)01170-2)
- Marwan, N., Romano, M. C., Thiel, M., & Kurths, J. (2007). Recurrence plots for the analysis of complex systems. *Physics Reports*, *438*, 237–329. <https://doi.org/10.1016/j.physrep.2006.11.001>
- Matlab (2016). The MathWorks, Inc., Natick, Massachusetts, United States.
- McGarva, A. R., & Warner, R. M. (2003). Attraction and social coordination: Mutual entrainment of vocal activity rhythms. *Journal of Psycholinguistic Research*, *32*, 335–354. <https://doi.org/10.1023/a:1023547703110>
- Mennin, D. S., Fresco, D. M., Heimberg, R. G., Schneier, F. R., Davies, S. O., & Liebowitz, M. R. (2002). Screening for social anxiety disorder in the clinical setting: Using the Liebowitz Social Anxiety Scale. *Journal of Anxiety Disorders*, *16*, 661–673. [https://doi.org/10.1016/S0887-6185\(02\)00134-2](https://doi.org/10.1016/S0887-6185(02)00134-2)
- Moulder, R. G., Boker, S. M., Ramseyer, F., & Tschacher, W. (2018). Determining synchrony between behavioral time series: An application of surrogate data generation for establishing falsifiable null-hypotheses. *Psychological Methods*. Advance online publication. <https://doi.org/10.1037/met0000172>
- Nagaoka, C., & Komori, M. (2008). Body movement synchrony in psychotherapeutic counseling: A study using the video-based quantification method. *IEICE Transactions on Information and Systems*, *91*, 1634–1640. <https://doi.org/10.1093/ietisy/e91-d.6.1634>
- Nelson, A., Grahe, J., Ramseyer, F., & Serier, K. (2014). Psychological data from an exploration of the rapport/synchrony interplay using motion energy analysis. *Journal of Open Psychology Data*, *2*, e5. <https://doi.org/10.5334/jopd.ae>
- Paulick, J., Deisenhofer, A.-K., Ramseyer, F., Tschacher, W., Boyle, K., Rubel, J., & Lutz, W. (2018). Nonverbal synchrony: A new approach to better understand psychotherapeutic processes and drop-out. *Journal of Psychotherapy Integration*, *28*, 367–384. <https://doi.org/10.1037/int0000099>
- Paxton, A. (2015). *Coordination: Theoretical, methodological, and experimental perspectives* (Doctoral dissertation). University of California, Merced.
- Paxton, A., & Dale, R. (2013a). Argument disrupts interpersonal synchrony. *Quarterly Journal of Experimental Psychology*, *66*, 2092–2102. <https://doi.org/10.1080/17470218.2013.853089>
- Paxton, A., & Dale, R. (2013b). Frame-differencing methods for measuring bodily synchrony in conversation. *Behavior Research Methods*, *45*, 329–343. <https://doi.org/10.3758/s13428-012-0249-2>
- Ramseyer, F. (2011). Nonverbal synchrony in psychotherapy: Embodiment at the level of the dyad. In W. Tschacher & C. Bergomi (Eds.), *The implications of embodiment: Cognition and communication* (pp. 193–207). Exeter, UK: Andrews.
- Ramseyer, F. (2013, June). *Synchronized movement in social interaction*. Paper presented at the 2013 Inputs–Outputs Conference: An Interdisciplinary Conference on Engagement in HCI and Performance, Brighton, UK.
- Ramseyer, F., & Tschacher, W. (2010). Nonverbal synchrony or random coincidence? How to tell the difference. In A. Esposito (Ed.), *Development of multimodal interfaces: Active listening and synchrony* (pp. 182–196). Berlin, Germany: Springer.
- Ramseyer, F., & Tschacher, W. (2011). Nonverbal synchrony in psychotherapy: Coordinated body movement reflects relationship quality and outcome. *Journal of Consulting and Clinical Psychology*, *79*, 284–295. <https://doi.org/10.1037/a0023419>
- Ramseyer, F., & Tschacher, W. (2016). Movement coordination in psychotherapy: Synchrony of hand movements is associated with session outcome. A single-case study. *Nonlinear Dynamics, Psychology, and Life Sciences*, *20*, 145–166.
- Robinson, J. W., Herman, A., & Kaplan, B. J. (1982). Autonomic responses correlate with counselor–client empathy. *Journal of Counseling Psychology*, *29*, 195–198. <https://doi.org/10.1037/0022-0167.29.2.195>
- Roisman, G. I., Holland, A., Fortuna, K., Fraley, R. C., Clausell, E., & Clarke, A. (2007). The Adult Attachment Interview and self-reports of attachment style: An empirical rapprochement. *Journal of Personality and Social Psychology*, *92*, 678–697. <https://doi.org/10.1037/0022-3514.92.4.678>
- Schönbrodt, F. D., & Perugini, M. (2013). At what sample size do correlations stabilize? *Journal of Research in Personality*, *47*, 609–612. <https://doi.org/10.1016/j.jrp.2013.05.009>
- Shockley, K. (2005). Cross recurrence quantification of interpersonal postural activity. In M. A. Riley & G. C. Van Orden (Eds.), *Tutorials in contemporary nonlinear methods for the behavioral sciences* (pp. 142–177). Arlington, VA: National Science Foundation.
- Shockley, K., Richardson, D. C., & Dale, R. (2009). Conversation and coordinative structures. *Topics in Cognitive Science*, *1*, 305–319. <https://doi.org/10.1111/j.1756-8765.2009.01021.x>
- Tashman, L. J. (2000). Out-of-sample tests of forecasting accuracy: An analysis and review. *International Journal of Forecasting*, *16*, 437–450. [https://doi.org/10.1016/S0169-2070\(00\)00065-0](https://doi.org/10.1016/S0169-2070(00)00065-0)

- Taylor, R. (1990). Interpretation of the correlation coefficient: A basic review. *Journal of Diagnostic Medical Sonography*, 6, 35–39. <https://doi.org/10.1177/875647939000600106>
- Tronick, E. D., Als, H., & Brazelton, T. B. (1977). Mutuality in mother–infant interaction. *Journal of Communication*, 27, 74–79. <https://doi.org/10.1111/j.1460-2466.1977.tb01829.x>
- Tschacher, W., Rees, G. M., & Ramseyer, F. (2014). Nonverbal synchrony and affect in dyadic interactions. *Frontiers in Psychology*, 5, 1323. <https://doi.org/10.3389/fpsyg.2014.01323>
- Watanabe, T. (1983). A study of motion–voice synchronization. *Bulletin of JSME*, 26, 2244–2250. <https://doi.org/10.1299/jsme1958.26.2244>
- Watanabe, T. (1987). A quantitative analyzing method for the synchrony between a speaker's voice and a listener's nodding in interpersonal communication. *JSME International Journal*, 30, 2040–2043. <https://doi.org/10.1299/jsme1987.30.2040>
- Wenzel, A., Graff-Dolezal, J., Macho, M., & Brendle, J. R. (2005). Communication and social skills in socially anxious and nonanxious individuals in the context of romantic relationships. *Behaviour Research and Therapy*, 43, 505–519. <https://doi.org/10.1016/j.brat.2004.03.010>
- Wiltshire, T. (2015). Team interaction dynamics during collaborative problem solving (Doctoral dissertation). University of Central Florida, Orlando, FL.
- Wittchen, H.-U., Wunderlich, U., Gruschwitz, S., & Zaudig, M. (1997). *SKID-I: Strukturiertes klinisches Interview für DSM-IV*. Göttingen, Germany: Hogrefe.

RESEARCH ARTICLE

Identification of movement synchrony: Validation of windowed cross-lagged correlation and -regression with peak-picking algorithm

Désirée Schoenherr^{1*}, Jane Paulick², Bernhard M. Strauss¹, Anne-Katharina Deisenhofer², Brian Schwartz², Julian A. Rubel², Wolfgang Lutz², Ulrich Stangier³, Uwe Altmann¹

1 Jena University Hospital, Institute of Psychosocial Medicine and Psychotherapy, Jena, Germany, **2** Trier University, Department of Clinical Psychology and Psychotherapy, Trier, Germany, **3** Goethe University Frankfurt/Main, Department of Clinical Psychology and Psychotherapy, Frankfurt/Main, Germany

* desiree.schoenherr@med.uni-jena.de



OPEN ACCESS

Citation: Schoenherr D, Paulick J, Strauss BM, Deisenhofer A-K, Schwartz B, Rubel JA, et al. (2019) Identification of movement synchrony: Validation of windowed cross-lagged correlation and -regression with peak-picking algorithm. PLoS ONE 14(2): e0211494. <https://doi.org/10.1371/journal.pone.0211494>

Editor: Dennis Tay, The Hong Kong Polytechnic University, HONG KONG

Received: September 7, 2018

Accepted: January 15, 2019

Published: February 11, 2019

Copyright: © 2019 Schoenherr et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: The data underlying this study have been uploaded to GitHub and are accessible using the following link: https://github.com/DesireeSchoenherr/data_validation_study.

Funding: This work was supported by the German Research Foundation (DFG project “Timing of nonverbal patient-therapist-interaction and therapeutic success of social phobic patients (TIMPATHIN) - <http://www.dfg.de/en/index.jsp>”, under funding number GZ BMS: STR 306/28-1 and

Abstract

In psychotherapy, movement synchrony seems to be associated with higher patient satisfaction and treatment outcome. However, it remains unclear whether movement synchrony rated by humans and movement synchrony identified by automated methods reflect the same construct. To address this issue, video sequences showing movement synchrony of patients and therapists (N = 10) or not (N = 10), were analyzed using motion energy analysis. Three different synchrony conditions with varying levels of complexity (naturally embedded, naturally isolated, and artificial) were generated for time series analysis with windowed cross-lagged correlation/ -regression (WCLC, WCLR). The concordance of ratings (human rating vs. automatic assessment) was computed for 600 different parameter configurations of the WCLC/WCLR to identify the parameter settings that measure movement synchrony best. A parameter configuration was rated as having a good identification rate if it yields high concordance with human-rated intervals (Cohen’s kappa) and a low amount of over-identified data points. Results indicate that 76 configurations had a good identification rate (IR) in the least complex condition (artificial). Two had an acceptable IR with regard to the naturally isolated condition. Concordance was low with regard to the most complex (naturally embedded) condition. A valid identification of movement synchrony strongly depends on parameter configuration and goes beyond the identification of synchrony by human raters. Differences between human-rated synchrony and nonverbal synchrony measured by algorithms are discussed.

Introduction

The term *nonverbal behavior* is used to describe various behaviors such as gaze, gestures, facial expressions, body postures and movements [1]. The investigation of *nonverbal synchrony*

WL: LU 660/8-1) and uses video recordings from the SOPHO-NET treatment study funded by the German Federal Ministry of Education and Research (BMBF, FKZ 01GV0607). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

within a dyad has gained increased consideration within areas of research such as physician-patient interactions [2], psychotherapist-patient interactions [3–5], mother-child communication [6–8], human-machine interactions [9], interactions within friendships [10, 11] and courtship behavior [12]. Nonverbal synchrony refers to nonverbal behaviors of interacting individuals that are connected to each other on a temporal level [11, 13]. In contrast to posture mirroring as showing in the same static posture, movement synchrony as a specific type of nonverbal synchrony, refers to the dynamic aspect, which is the temporal connection between motions of the interacting persons, independent of a particular body part or direction [14]. This nonverbal synchrony can be observed in short sequences that we call synchronization intervals [10]. In psychotherapy context, movement synchrony would be observed, for example, if the therapist is nodding and the patient is changing his/her body position with a short time delay or when both interaction partners are nodding simultaneously.

In comparison to human ratings, methods that automatically record nonverbal behavior and determine nonverbal synchrony are economic, less time-consuming and require fewer resources. A current method that automatically generates time series that represent movements of individual persons in a video is motion energy analysis (MEA) [5, 6, 11]. If multiple persons have been recorded, time series can be computed for each person within a video by determining a region of interest (ROI). After computation, motion energy time series can be used to identify movement synchrony. For this purpose, typical automated methods such as cross-recurrence quantification analyses, spectral analyses, or correlative and regressive time series analyses methods are applied (TSAM) [15].

In the social and behavioral sciences, the use of correlative and regressive TSAM predominantly prevails. Two prominent examples of correlative and regressive TSAM are windowed cross-lagged correlation (WCLC) [16] and windowed cross-lagged regression (WCLR) [10, 11]. The application of TSAM on movement time series is based on a movement synchrony definition of interacting partners' synchronous or time-lagged synchronous movements. The underlying methodical principle can be described as follows: The correlation (or regression) between the first segment of the time series of person A and the first segment of the time series of person B is calculated. Next, this local association is tested for significance (e.g., R^2 -difference-test). Afterwards, the segment of the time series of person B is shifted and procedure is repeated until the entire time series and all reasonable time lags between A and B have been tested. The result of such windowed cross-lagged correlation (WCLC) or windowed cross-lagged regression (WCLR) is a "landscape" (R^2 -matrix), which shows at which time point the behavior of A is significantly associated with the simultaneous or time-lagged behavior of B. Subsequently, different algorithms can be used to analyze this matrix and obtain a global synchrony score. For example, Altmann [10], Altmann [11] used a modification of Boker et al.'s (2002) peak-picking algorithm to determine synchronization intervals. These intervals were characterized by the time lag and a specific start and end point. Different indices may be calculated to obtain a global score, for example, the averaged strength of the association (averaged over time and different time lags) as used by Ramseyer and Tschacher [5]. We abbreviate this method and the corresponding index with $WCLC_S$ (subscript S stands for strength). Another possibility is the $WCLC_F$ (subscript F stands for frequency) or the WCLR, where the frequency of synchrony is calculated as the ratio between synchronized time and total time (for more details, see Schoenherr, Paulick [17]). Fig 1 illustrated the computational steps for the movement synchrony identification.

Previous studies have shown that, based on identical time series, different correlative and regressive TSAM lead to different synchrony indices [11, 17]. Additionally, calculating synchrony for different baseline-surrogate datasets with varying parameter settings (e.g., window size = bandwidth) lead to different results [18]. This indicates that the computed value of

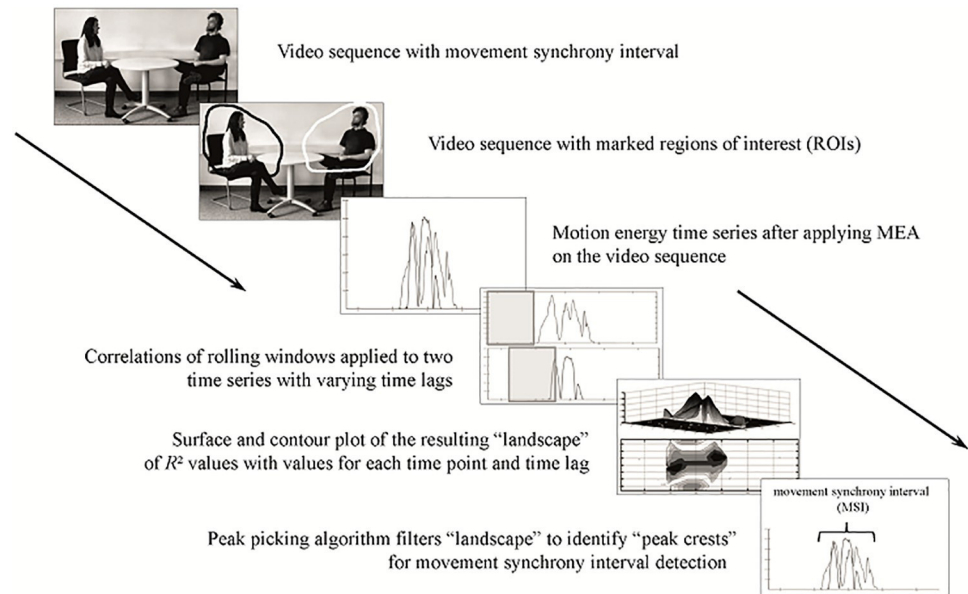


Fig 1. Computational steps from video sequence to identification of a movement synchrony interval.

<https://doi.org/10.1371/journal.pone.0211494.g001>

synchrony strongly depends on the algorithm *and* parameter settings applied. However, to date, there is no consensus on which parameter settings lead to optimal results. [Table 1](#) displays examples of the use of WCLC, WCLR and related procedures as well as the selected parameter settings.

Smoothing procedure

In order to eliminate video noise, many authors recommend smoothing the time series before applying TSAM. Grammer, Honda [33], Paulick, Deisenhofer [3] and Ramseyer and Tschacher [5], Ramseyer and Tschacher [27] used a moving average method. Nagaoka and Komori [25] utilized wavelets. Altmann [10], Altmann [11] recommended smoothing splines and Paxton and Dale [26] used a Butterworth low-pass filter. Apart from the different methods, the critical issue is the amount of smoothing applied. To date, however, it remains unclear which amount of smoothing leads to improved synchrony detection. [Fig 2](#) illustrates the impact of smoothing on the shape of the time series.

Transformation

Motion energy time series are non-stationary [11]. This means that the variance and the expectation value of the time series are not stable over time. A linear trend or cyclic behavior in both time series under consideration could lead to spurious correlations, which may cause an over-identification of synchronization intervals [10]. Authors have attempted to control non-stationarity by transforming the distribution of values into an approximately normal distribution [11]. By means of an Anscombe transformation [34] or Box-Cox transformation [35], a normal distribution can be approximated. Moreover, these transformations have a variance-stabilizing effect. Nagaoka and Komori [25] used a Box-Cox transformation with the parameter $\lambda = 0$, which corresponds to a logarithmic transformation, whereas Altmann [11] recommends the Anscombe transformation. Furthermore, z-standardization is used to standardize the value

Table 1. Overview of applied parameter setting.

| | Smoothing procedure | Transformation | Method | Bandwidth | Step | Measures per second |
|--|---|------------------------------------|-------------------------|---|-----------------------|---------------------|
| Altmann [10], Altmann [11] | Smoothing splines ($\lambda = .9995$) | Anscombe transformed | WCLR, WCLC _F | 100 frames (ca. 4 sec) | 2 frames (0.08 sec) | 24 |
| Ashenfelter, Boker [19] | - | - | WCLC | 160 frames (2 sec) | 10 frames (0.125 sec) | 80 |
| Bilakhia, Petridis [20] | - | - | WCLC | - | - | 58 |
| Boker, Rotondo [16] | - | - | WCLC | 320 frames (4 sec) | 8 frames (0.1 sec) | 80 |
| Boker and Rotondo [21] | - | - | WCLC | 160 frames (2 sec) | - | 80 |
| Bozkurt, Yemez [22] | - | - | WCLC adapted (CCA) | 180 frames (6 sec) | - | 30 |
| Campbell [23] | Low pass filtered | - | WCC | 1 frame | - | - |
| Delaherche and Chetouani [15] | - | Ramseyer's normalization procedure | WCLC | 1 sec | - | - |
| Paulick, Deisenhofer [3] | Moving average 10 | z-transformed | WCLC _S | 60 frames (1 min) | 60 frames (1 min) | 10 |
| Messinger, Mahoor [7] | - | z-transformed | WCLC | 3 sec | - | Ca. 30 |
| Michelet, Karp [24] | - | - | WCC | 75 frames | - | 25 |
| Nagaoka and Komori [25] | - | Box-Cox transformed | WCC | 18000 frames (10 min) | 150 frames (5 sec) | 30 |
| Paxton and Dale [26] | Butterworth low-pass filter | - | CLC | - | - | 8 |
| Ramseyer and Tschacher [5], Ramseyer and Tschacher [27], Ramseyer and Tschacher [28] | Moving average 10 | z-transformed | WCLC _S | 600 frames (1 min) | 1 frame (0.1 sec) | 10 |
| Sun, Nijholt [29] | - | - | WCC | Between 20–280 frames | - | - |
| Tschacher, Rees [30] | Moving average 10 | z-transformed | WCLC _S | 300 frames (30 sec) | 0.1 sec | 10 |
| Tronick, Als [31] | - | - | WCC | 10 frames (10 sec) | - | 1 |
| Yang, Wang [32] | - | - | WCLC | Between 64–256 frames (0.64–2.56 sec) Best: 128 frames (1.28 sec) | 32 frames (0.32 sec) | 100 |
| Watanabe [6] | - | - | WCLC | 480 frames (8 sec) | - | 60 |

Note. Subscript letters indicate the type of global output parameter (for details see [17]). WCLC: windowed cross-lagged correlation, WCC: windowed cross-correlation, CLC: cross-lagged correlation, WCLR: windowed cross-lagged regression, CCA: canonical correlation analysis, λ refers to the roughness penalty parameter, '-' indicates that information was not reported.

<https://doi.org/10.1371/journal.pone.0211494.t001>

range of time series from different persons [5, 27]. However, this transformation does not lead to any change of the distribution or variance. Another option to standardize the value range is to transform the motion energy time series (METS) with respect to the ROI size (size standardization—for details see [Methods](#)). As a result, values of this linear transformation vary from zero (no motion) to 1 (entire ROI is active). [Fig 3](#) displays the different transformations and their impact on the distribution of values.

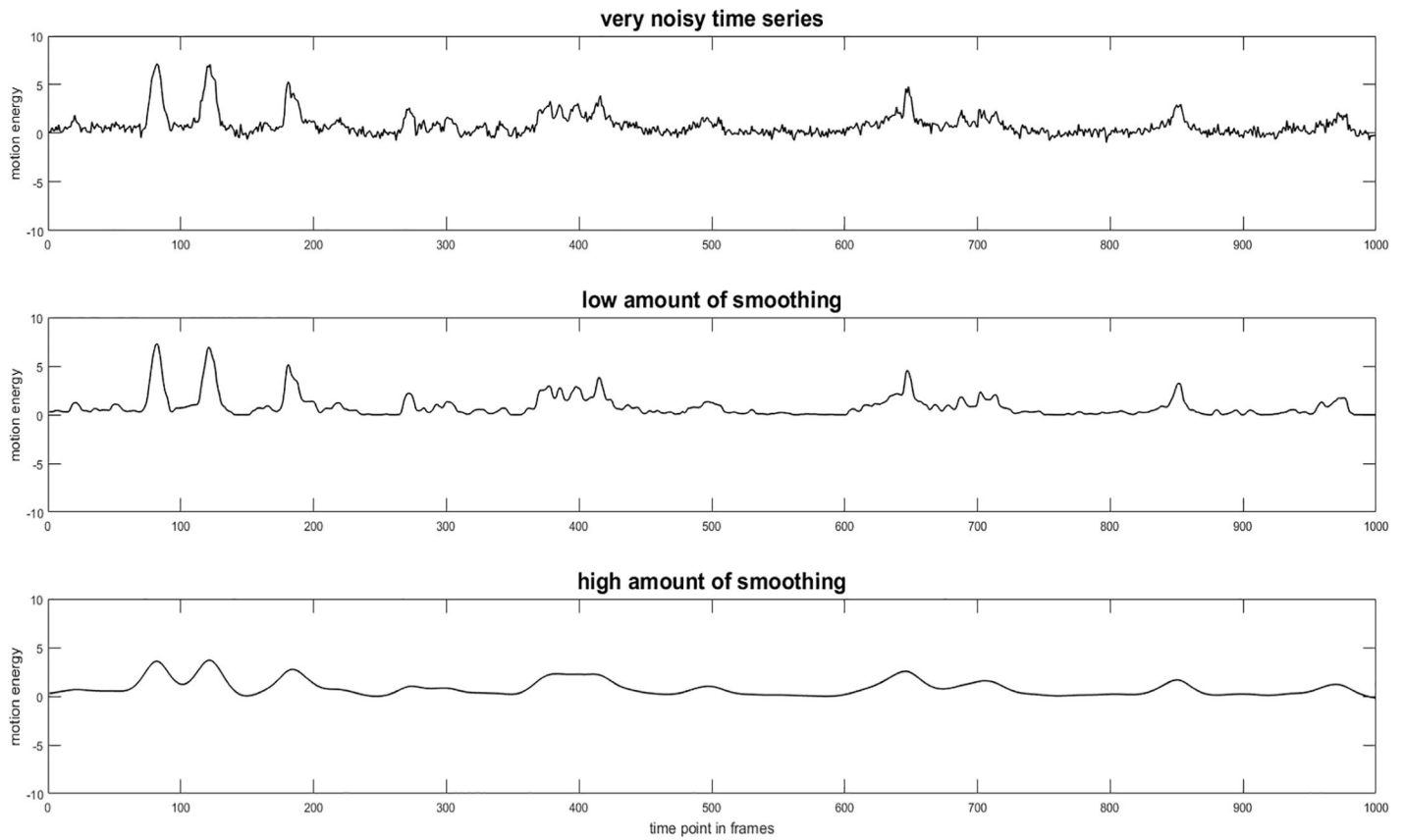


Fig 2. Different amounts of smoothing applied to motion energy time series (METS).

<https://doi.org/10.1371/journal.pone.0211494.g002>

Method

Two methods that show high convergent validity and a local identification of movement synchronization intervals (MSI) are the WCLR and WCLC_F from Altmann [10], Altmann [11], [17]. The algorithms are publicly available at GitHub: https://github.com/10101-00001/sync_ident. The output of both methods is a table showing every MSI of the interaction under investigation. This enables us to compare these MSI with intervals rated by human raters. The WCLR was developed for cyclic time series to avoid spurious cross-correlations due to auto-correlation. We used these two methods because they identify synchronization intervals. These intervals may be used to investigate more local events that emerge on a short-time level as alliance ruptures [36] or good moments within therapy. Furthermore, these measures had the highest correlation with an external criteria in the validation study of Schoenherr, Paulick [17].

Bandwidth

An important parameter in the application of windowed correlative and regressive methods is bandwidth. Bandwidth specifies the length of the examined segment. The choice of bandwidth depends on content and methodological considerations [16]. If the bandwidth is set too long, the detected synchrony may be biased and the true dynamics of the interaction may not be optimally mapped. In contrast, with small bandwidths, it is also possible that synchronization intervals remain undetected. Additionally, if the selected segment is too small, the correlation

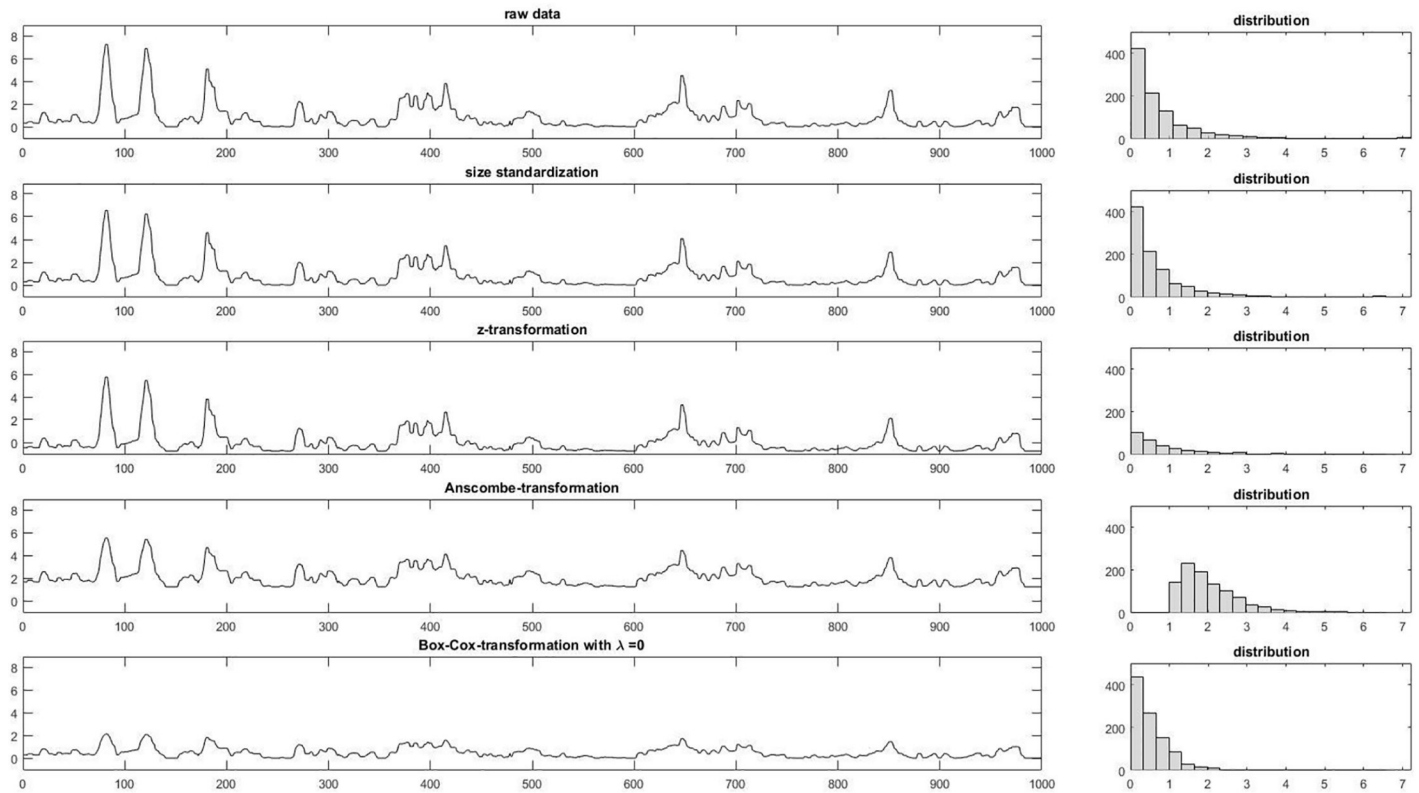


Fig 3. Different transformations and their impact on the distribution of motion energy values assessed with MEA.

<https://doi.org/10.1371/journal.pone.0211494.g003>

or regression cannot be reliably determined due to the low number of measured values. Schönbrodt and Perugini [37] recommend a sample size of 250 values for stable parameter estimation with respect to typical correlation analyses. This number, however, is dependent on the true correlation. For high correlations ($r = .7$), the authors found a sample size of 65 to be sufficient. An optimal trade-off for high reliability and high sensitivity has not yet been determined [16]. The choice of the bandwidth is also dependent on the number of values recorded per second. Currently, every researcher chooses the bandwidth with respect to the phenomena under investigation. If one wants to assess global synchrony that is established over period, higher bandwidths (30 seconds and higher) are used [5]; if the phenomena under investigation are local synchrony intervals, smaller bandwidths (2.5 seconds: Altmann [10]) are applied.

With respect to the problem of non-stationarity, Boker, Rotondo [16] proposed that the assumption of stationarity can be made locally within segments. Therefore, using windows may also be beneficial for statistical reasons.

Spurious correlations

Various approaches can be used to control for coincidentally identified synchronization intervals. Ramseyer and Tschacher [38] built on the concept of pseudosynchrony by Bernieri, Reznick [39]. They created an artificial database of interactions by means of random permutations of the segments of the real time series. In doing so, they created an empirical distribution to determine a cut-off value to discriminate between pseudosynchrony and genuine synchrony. Louwerse, Dale [40] also built surrogate pairs to compare genuine synchrony with pseudosynchrony by using random permutations of data points. Recently, Moulder, Boker

[18] compared different methods of surrogate dataset generation (data shuffling, segment shuffling, data sliding, participant shuffling) showing that not every baseline is equally conservative. Additionally, there seems to be an interaction with parameters chosen for applying the WCLC. In contrast to baselines which are based on surrogate datasets, Gottman and Ringland [41] and Altmann [10], Altmann [11] propose a parametrical approach to test genuine synchrony against randomly occurring synchrony, whereby the distribution of the proof statistic is pre-defined (e.g., χ^2 -distributed). While computing the correlation, they test the value against zero using the pre-defined distribution. The surrogate approach provides an empirically derived null distribution, whereas the approach of Gottman and Ringland [41] would require the selection of a pre-defined distribution that could be used to determine significance of computed correlations. Another approach is to increase the cut-off for the comparison of genuine and pseudosynchrony [42]. This may be realized by introducing a cut-off value for meaningful correlations or R^2 values.

Research questions

In the literature, most of the TSAM are applied with various parameter settings (e.g., transformation, smoothing, method, bandwidth, control for spurious correlations), which are largely based on the researchers' preferences or theoretical considerations. Apart from Altmann [10] and Paxton and Dale [26], validation studies are lacking evidence. Especially the validation of TSAM against human-rated synchrony would be an important contribution to the field. In the present study, we examine the quality of movement synchrony identification of different TSAM configurations depending on the parameters described, while using human-rated synchronization intervals and simulated time series build out of the human-rated intervals with varying complexity as reference. Aim of this paper is to exemplify a validation of TSAM regarding their capability to correctly identify synchronization intervals given an external criterion (in our case: simulated time series and human ratings of movement synchrony).

Methods

Step 1: Selection of video material and generation of time series

To create a dataset of video sequences with and without movement synchrony under naturalistic conditions, 40 sequences showing different amounts of movement synchrony (e.g., synchronous head nodding or body movements which provoke the impression to be interrelated) were selected by the first author. Sequences were extracted from video recordings of the 3rd psychotherapy session of social anxiety disorder patients treated with cognitive-behavioral therapy or psychodynamic therapy. The video recordings originated from the SOPHO-NET treatment study, which was conducted between 2007 and 2009 (for more details see [43, 44]). [for more details, see 43, 44]. Only videos with an optimal camera position and acceptable recording conditions were included. The sequences had a mean duration of 116.69 ($SD = 11.35$) seconds.

All 40 sequences were rated by three independent female raters with a master's degree in psychology, age range: 26 to 28 years. The raters were familiar with the concept of movement synchrony and were instructed about the general working procedure of the algorithms we use to measure movement synchrony. Before rating, all raters had to rate two example sequences correctly to ensure that they understood the instruction. On a dichotomous scale, the raters were asked to rate whether the sequence showed patient and therapist movements, which were related to each other in any possible way (synchrony: yes) or not (synchrony: no). Rater had no restrictions of the time lag; therefore, they could rate perfectly synchronous or time-lagged movements as synchrony. In addition, raters were allowed to watch the sequence as often as

they wanted to. Of all video sequences, 18 were rated consistently as sync-sequences and 11 as no sync-sequences. If a sequence was rated as containing movement synchrony, the interval in which this took place was additionally identified (= movement synchronization interval: MSI). Ten movement synchrony sequences and ten sequences showing no movement synchrony, which were rated equally by all raters (sync or no sync), were included for subsequent analysis.

The videos of selected sequences ($N = 10$ with synchrony + 10 without synchrony) were converted into a consistent video format (size of 640x480, frame rate of 25 fps, bit rate of 2000 Kb/s using Any Video Converter 3.0 [45]). Sequences had a mean duration of 118 ($SD = 12.94$) seconds. Patient and therapist movements displayed in the videos were assessed using MEA [46] MATLAB scripts are publicly available at: <https://github.com/10101-00001/MEA>. The ROI covered the upper body from the chair's seat upwards. Next, the starting and end points (in frames) of the specified MSI were identified using the plot of the time series. The starting point was defined as the first frame that showed a motion energy value greater than zero proximal to the human-rated starting point. The end point was set proximal to the human-rated end point, where the motion energy time series were equal to zero again. MSI had a mean duration of 145 frames (5.8 seconds), a standard deviation of 57.93 frames (2.37 seconds), a minimum duration of 26 frames (1.04 seconds) and a maximum duration of 282 frames (11.28 seconds). The individuals shown in Fig 1 in this manuscript have given written informed consent (as outlined in PLOS consent form) to publish their case details.

Step 2: Generation of synchrony conditions

We used three different experimental conditions since parameter settings should be tested with respect to varying complexity of the stimulus material: 1) natural synchrony rated by humans embedded in a normal interaction, 2) natural synchrony rated by humans isolated from normal interaction, 3) artificial synchrony. In the first condition, patient and therapist METS resulting from MEA were used (see Fig 4, lower figure, gray box indicates MSI within

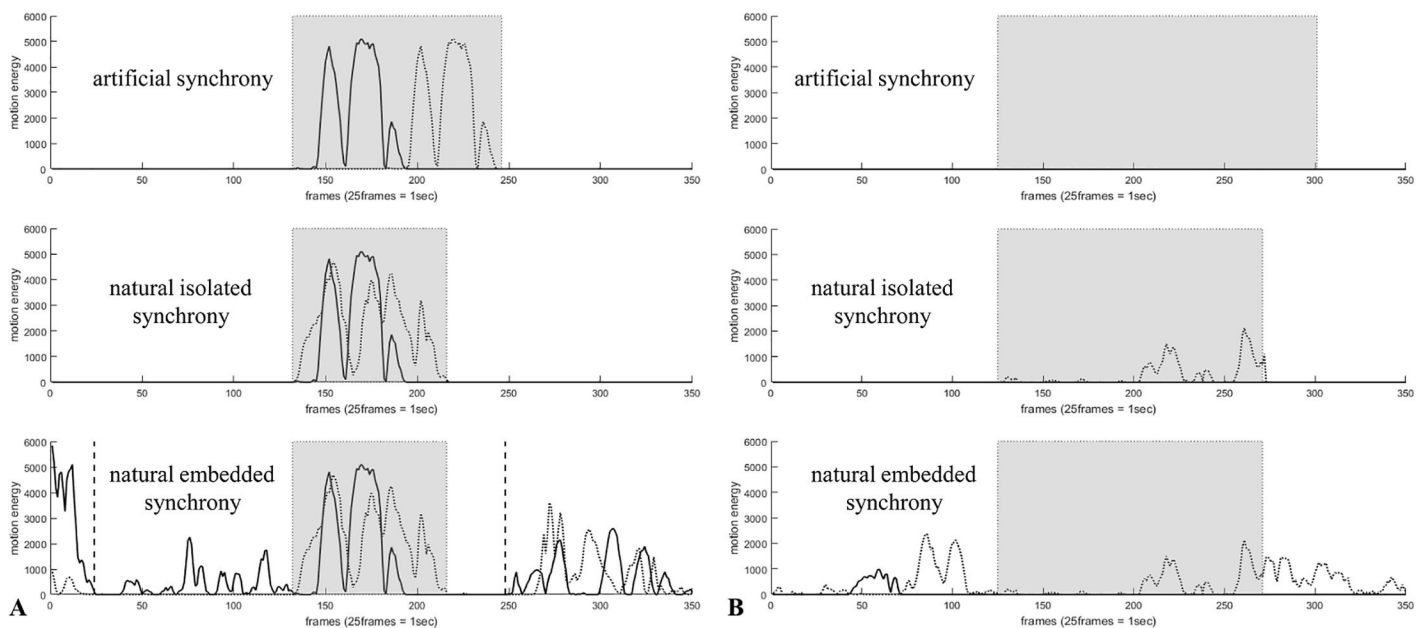


Fig 4. Examples of the three different conditions with an MSI (A) and a no MSI (B) using two time series (patient and therapist movements). Artificial synchrony (upper plot), naturally isolated synchrony (middle), naturally embedded synchrony (lower plot). Gray boxes indicate human rated MSI or no MSI; dashed vertical lines indicate the area investigated by the algorithms.

<https://doi.org/10.1371/journal.pone.0211494.g004>

the sequence). Due to the fact that the natural time series were very complex and included other movements apart from the MSI (or no MSI), we modified the time series to obtain a naturally isolated condition: All values of both time series (patient and therapist time series) were set to zero excluding the motion energy values of the MSI (or no MSI) identified by the human raters. Thus, this condition incorporates only the movements of both persons within the specified MSI (or no MSI). Thirdly, we generated an artificial condition: Given the time series of A and B from the naturally isolated condition, we replaced B's time series with a time-lagged time series of A (time lag of 50 frames). Therefore, we obtained an MSI showing perfect time-lagged movement synchrony. Therefore, in the artificial condition the *true* synchronization interval is known due to the creation of a time series pair. In contrast, in the isolated and natural condition the reference synchronization interval is rated by human raters and can be biased. With respect to the no MSI group, we did not duplicate the patient time series but included a time series of zeros as the therapist time series. If the patient did not move, this therefore resulted in two time series with no movement. The different time series are displayed in Fig 4.

As a result, we obtained 10 naturally embedded synchrony time series pairs, 10 naturally isolated synchrony time series pairs and 10 artificial synchrony time series pairs for MSI and no MSI, respectively (total $n = 3$ (naturally embedded vs. naturally isolated vs. artificial) * 2 (sync vs. no sync) * 10 (video sequences) = 60). Note that the artificial condition can be viewed as simulated time series as they can be easily replicated. We used parts of the real-interaction time series to build those time series, but this was only a matter of choosing those pulses that are realistic for human interactions, particularly realistic for the therapeutic context.

Step 3: Applying 600 configurations of TSAM algorithms

Next, multiple TSAM with different parameter settings were applied to METS of video sequences. In total, 600 different configurations were investigated by combining method, ($WCLC_F$, WCLR), type of transformation (raw data, size standardization, Box-Cox transformation with $\lambda = 0$, Anscombe transformation), smoothing (no smoothing, slight smoothing with smoothing splines ($\lambda = .900$), high smoothing with smoothing splines ($\lambda = .005$)), bandwidth (75 frames (= 3 seconds), 125 frames (= 5 seconds), 175 frames (= 7 seconds), 250 frames (= 10 seconds) and 750 frames (= 30 seconds)) and R^2 cut-off to filter spurious correlations ($R^2 > 0.0$, $R^2 > 0.1$, $R^2 > 0.2$, $R^2 > 0.25$, and $R^2 > 0.3$). In the following sections, we omit the 'F' of $WCLC_F$ for readability. Note that Box-Cox transformation with $\lambda = 0$ equals to a logarithmic transformation; thus we increased each motion energy value with 1 to avoid $\log(0)$ before applying this transformation.

Assessing the motion energy of two individuals with MATLAB [46] has the advantage that the size of the regions of interest (ROIs) can be saved. This allows the standardization of the METS by the size of the ROI. First, we determined the ratio of the larger ROI to the smaller ROI in the video sequence (range of ROI sizes ratios within a video was 1.03 to 1.99, $M = 1.37$, $SD = 0.36$). Next, we multiplied the ratio with each element of the METS of the individual with the smaller ROI size. The size standardization was used as one possible transformation aside from Box-Cox and Anscombe transformations.

With regard to the automated identification of MSI, we used WCLC and WCLR as implemented by Altmann [10], Altmann [11] and set the maximum time lag to 75 frames (= 3 seconds). Before applying both procedures but after transformation and smoothing, we added some noise ($M = 0$, $SD = .1$) to the time series to make a calculation of a correlation and regression possible. To reduce computation time, values were computed every two frames. Applying one configuration to the time series resulted in an $m \times n$ matrix (m = duration of

the video sequence in frames/2, n = number of time lags/2) of R^2 values. To determine intervals showing the highest R^2 values, peak-picking algorithm of Altmann [10], Altmann [11] was used (MATLAB script publicly available: https://github.com/10101-00001/sync_ident). The output of the algorithm is a list containing every MSI within the video sequence with its time lag, starting, end, and average R^2 value. To account for spurious correlations, we included an R^2 cut-off ($R^2 > 0.0$, $R^2 > 0.1$, $R^2 > 0.2$, $R^2 > 0.25$ and $R^2 > 0.3$) to filter MSI of the output list below the cut-off. The data of this study is publicly available at GitHub: https://github.com/DesireeSchoenherr/data_validation_study.

Step 4: Statistical analysis: Investigation of the best configuration

To determine which configuration identifies synchrony best, we were interested in high power of identification (i.e., valid identification of MSI) and low type I error (i.e., low percentage of over-identified MSI in sequences where no movement synchrony was present). Therefore, a configuration was rated as good if it showed high power in the presence of synchronization intervals and low type I error in their absence. In addition, a configuration was only rated as good if the criteria hold for *each* video sequence.

To compare the ratings with each other, human ratings of the video sequence/true synchronization intervals of the artificial time series were transposed to a time series with zeros and ones, indicating whether movement synchrony was present (= 1) or not (= 0). This time series functioned as a reference time series. In addition, a binary time series of the algorithm was built by setting the time series to one in an automatically identified MSI and to zero outside an MSI. As a measure of concordance of simulated/human and computer-rated synchronization intervals, we computed Cohen's kappa using both binary time series. A widely established scale for the interpretation of kappa suggests the following cut-offs: values ≤ 0 poor agreement, .00 – .20 slight agreement, .21 – .40 fair agreement, .41 – .60 moderate agreement, .61 – .80 substantial agreement and .81–1.00 almost perfect agreement [47]. With regard to the no MSI group, the proportion of over-identified frames (abbreviated as: *pr_out*) was calculated that equals the number of identified frames relative to the number of frames in a sequence.

The parameter configuration with the highest identification rate (IR) of synchrony was identified using a sequential procedure: First, the time series in the artificial condition were investigated to ensure that the configurations were able to detect artificial synchrony and no synchrony (high power, low type I error). Configurations with a kappa higher than .60 for *each* of the 10 synchronization intervals (high power) or a maximum value of 5% over-identified frames with respect to *each* of the 10 no synchronization intervals (low type I error) were classified as configurations with a good identification rate (IR). Configurations that yielded a minimum kappa between .41 and .60 for synchronization intervals or 5% to 10% over-identified frames for no synchronization intervals were evaluated as configurations with an acceptable IR. Configurations that did not fulfill either of these criteria were rated as configurations with a poor IR (sync: kappa < .40, no sync: more than 10% *pr_out*). Based on these classifications, the three IR levels were used to produce cross tables and conduct Fisher's exact test to check for significant overall influences of the manipulated parameters (e.g., bandwidth, smoothing amount). To examine which specific parameter configuration is best, we additionally conducted ordinal logistic regressions using the function `polr` in R [48]. Thereby, we used the IR as criterion and the parameters (levels were dummy coded) as predictors. Coefficients of the ordinal logistic regression which are higher than zero indicate that the parameter has a positive influence on the IR (higher kappa, lower *pr_out*); coefficients lower than zero indicate that the parameter has a negative influence on the IR (lower kappa, higher *pr_out*). A

significant predictor (e.g., method WCLC vs. WCLR) indicates that the IR is affected by the parameter.

Second, we investigated the data with respect to the naturally isolated condition, using *only the configurations that showed a good IR in the artificial condition*. Again, we determined the configurations with a good, moderate and bad IR. In addition, we conducted an ordinal logistic regression with respect to the naturally isolated condition, where only configurations with a good IR in the artificial condition were used.

Third, we examined the data with respect to the naturally embedded condition. Therefore, we used the *configurations with a good or acceptable IR for all video sequences in the naturally isolated condition*.

Lastly, we investigated whether the video sequence had an influence on the validity of identification using two Kruskal-Wallis-tests with the average kappa (κ_{mean}) or the average percentage of over-identified frames ($\text{pr_out}_{\text{mean}}$) as the dependent variable and the video sequence as independent variables. As our independent variables are not normally distributed ($p < .001$), we used Kruskal-Wallis-test. Note, that a sequence ID only characterizes the stimulus. Therefore, the Kruskal-Wallis-test will show us if there is a dependency between the sequence ID (categorical) and the dependent variables (κ_{mean} , $\text{pr_out}_{\text{mean}}$, both continuous). However, since the independent variable is categorical, the type of dependency cannot be characterized further.

In additional analyses, we also investigate the general influence of the parameters on the average of kappa and pr_out . Details and results are reported in the Supporting information ([S2 Appendix](#)).

Results

Artificial condition

Firstly, we considered the artificial condition and counted the number of configurations with a poor, acceptable and good IR depending on their parameter setting (e.g., smoothing, bandwidth, etc.). The corresponding cross table is reported in [Table 2](#).

To investigate the influence of different parameters on the IR, we conducted Fisher's exact tests because for some cells the assumptions of a χ^2 -test were not fulfilled (not more than 20% of the expected counts less than 5, provided that none are less than 1). We found that the choice of method, bandwidth, smoothing amount and R^2 cut-off influenced the IR in the MSI group (smoothing: $p < .001$, Cramer's $V = 0.23$; method: $p = .049$, Cramer's $V = 0.10$; bandwidth: $p < .001$, Cramer's $V = 0.66$; R^2 cut-off: $p < .001$, Cramer's $V = 0.24$). The highest influence was found for bandwidth, the lowest for method. The frequencies in [Table 2](#) (kappa) suggest that bandwidth should be rather low (175 and/or smaller), a low amount of smoothing should be applied and the R^2 cut-off should be set to 0.25. No significant effects were found with respect to transformation ($p = .77$).

Referring to the artificial no MSI condition (pr_out), we found significant effects for method, bandwidth, and R^2 cut-off (method: $p < .001$, Cramer's $V = 0.37$; bandwidth: $p < .001$, Cramer's $V = 0.50$; R^2 cut-off: $p < .001$, Cramer's $V = 0.48$). The largest influence was found for bandwidth, the lowest for the amount of smoothing. The frequencies in [Table 2](#) (pr_out) suggest that bandwidth should be rather high, the WCLR should be applied and the R^2 cut-off should be set to 0.25. No significant effects were found with respect to transformation and smoothing (transformation: $p = .28$; smoothing: $p = .41$).

To identify the best parameter also scientifically, we conducted ordinal logistic regressions. This offers the possibility to also control for other covariates. Results of the ordinal logistic

Table 2. Cross table of the variables method, bandwidth, transformation, smoothing, and R^2 cut-off value (artificial condition) and identification rate (IR) in synchronization intervals.

| | | Poor IR | | Acceptable IR | | Good IR | | total |
|----------------|------------------|---------------|--------------|--------------------|-------------------|---------------|--------------|-------------|
| | | $\kappa < .4$ | pr_out > 10% | .4 < $\kappa < .6$ | 10% > pr_out > 5% | $\kappa > .6$ | pr_out < 5% | |
| Smoothing | Raw data | 42 (21.0%) | 49 (24.5%) | 110 (55.0%) | 15 (7.5%) | 48 (24.0%) | 135 (68.0%) | 200 (33.3%) |
| | Slight smoothing | 41 (20.5%) | 47 (23.5%) | 131 (65.5%) | 17 (8.5%) | 28 (14.0%) | 136 (68.0%) | 200 (33.3%) |
| | High smoothing | 28 (14.0%) | 47 (23.5%) | 172 (86.0%) | 8 (4.0%) | 0 (0.0%) | 145 (72.5%) | 200 (33.3%) |
| Transformation | Raw data | 28 (18.7%) | 37 (24.7%) | 98 (65.3%) | 5 (3.3%) | 24 (16.0%) | 108 (72.0%) | 150 (25.0%) |
| | Size standard. | 26 (17.3%) | 35 (23.3%) | 108 (72.0%) | 7 (4.7%) | 16 (10.7%) | 108 (72.0%) | 150 (25.0%) |
| | Log-trans | 26 (17.3%) | 36 (24.0%) | 104 (69.4%) | 13 (8.7%) | 20 (13.3%) | 101 (67.3%) | 150 (25.0%) |
| Method | WCLC | 66 (22.0%) | 132 (44.0%) | 202 (67.3%) | 0 (0.0%) | 32 (10.7%) | 168 (56.0%) | 300 (50.0%) |
| | WCLR | 45 (15.0%) | 11 (3.7%) | 211 (70.3%) | 40 (13.3%) | 44 (14.7%) | 249 (83.0%) | 300 (50.0%) |
| Bandwidth | 75 | 0 (0.0%) | 48 (40.0%) | 100 (83.3%) | 8 (6.7%) | 20 (16.7%) | 64 (53.3%) | 120 (20.0%) |
| | 125 | 0 (0.0%) | 48 (40.0%) | 100 (83.3%) | 8 (6.7%) | 20 (16.7%) | 64 (53.3%) | 120 (20.0%) |
| | 175 | 1 (0.8%) | 12 (10.0%) | 83 (69.2%) | 11 (9.2%) | 36 (30.0%) | 97 (80.8%) | 120 (20.0%) |
| | 250 | 5 (4.2%) | 17 (14.2%) | 115 (95.8%) | 7 (5.8%) | 0 (0.0%) | 96 (80.0%) | 120 (20.0%) |
| | 750 | 105 (87.5%) | 18 (15.0%) | 15 (12.5%) | 6 (5.0%) | 0 (0.0%) | 96 (80.0%) | 120 (20.0%) |
| R^2 cut-off | > 0.00 | 27 (22.5%) | 71 (59.2%) | 91 (75.8%) | 28 (23.3%) | 2 (2.7%) | 21 (17.5%) | 120 (20.0%) |
| | > 0.10 | 21 (17.5%) | 24 (20.0%) | 88 (73.3%) | 4 (3.3%) | 11 (9.2%) | 92 (76.7%) | 120 (20.0%) |
| | > 0.20 | 21 (17.5%) | 24 (20.0%) | 88 (73.3%) | 4 (3.3%) | 11 (9.2%) | 92 (76.7%) | 120 (20.0%) |
| | > 0.25 | 21 (17.5%) | 0 (0.0%) | 58 (48.3%) | 0 (0.0%) | 41 (34.2%) | 120 (100.0%) | 120 (20.0%) |
| | > 0.30 | 21 (17.5%) | 24 (20.0%) | 88 (73.3%) | 4 (3.3%) | 11 (9.2%) | 92 (76.7%) | 120 (20.0%) |

Note. WCLC = windowed cross-lagged correlation, WCLR = windowed cross-lagged regression, bandwidth refers to the frame number. We used a frequency of 25 frames per second for our analysis, size standard. = size standardization, IR = identification rate, cell row proportions in brackets, κ = Cohen's kappa, pr_out = percentage of over-identified frames.

<https://doi.org/10.1371/journal.pone.0211494.t002>

regression based on kappa (log-likelihood = -196.44, $df = 16$) and pr_out (log-likelihood = -214.33, $df = 16$) are displayed in Table 3.

The coefficients for kappa indicate that transformation do not have an impact on the IR. No smoothing is best in comparison to slight and high amount of smoothing. The WCLR is superior in comparison to the WCLC (reference group). Additionally, a bandwidth of 175 and an R^2 cut-off of 0.25 (reference group) seem to be optimal for synchrony detection in MSI within the artificial condition.

The results regarding pr_out show that smoothing and transformation do not have an impact on the IR of over-identified frames. High significant coefficients indicate a positive influence on the IR (lower false positives, lower pr_out). Parameters that were best suited for the identification were the WCLR, a higher bandwidth and an R^2 cut-off of 0.25.

Within the artificial condition, 76 configurations fulfilled both criteria for a good IR ($\kappa > .6$ and pr_out < 5%) and therefore showed high power and low type I error. Using these configurations, the average and standard deviation of the rater concordance (minimum kappa, maximum kappa, mean kappa) between the artificially generated movement phenomenon (true MSI) and algorithm ratings of the 10 video sequences containing MSIs were: $M_{\kappa_{min}} = .64$, $SD_{\kappa_{min}} = .05$; $M_{\kappa_{max}} = .90$, $SD_{\kappa_{max}} = .05$, $M_{\kappa_{mean}} = .77$, $SD_{\kappa_{mean}} = .05$. The average and standard deviation of over-identified frames in the no MSI group were: $M_{pr_out_{min}} = 0.06\%$, $SD_{pr_out_{min}} = 0.17\%$, $M_{pr_out_{max}} = 1.01\%$, $SD_{pr_out_{max}} = 1.72\%$, $M_{pr_out_{mean}} = 0.32\%$, $SD_{pr_out_{mean}} = 0.70\%$.

Table 3. Results of the ordinal logistic regression with IR (kappa or pr_out) as criterion and parameters as predictors (artificial condition).

| | <i>b</i> (kappa) | <i>b</i> (pr_out) |
|------------------------|--|-------------------|
| Smoothing | reference group: smoothing = raw data | |
| Slight | -0.84* | 0.04 |
| High | -1.72* | 0.43 |
| Transformation | reference group: transformation = raw data | |
| Size-standard. | -0.39 | 0.04 |
| Log-trans | -0.13 | -0.36 |
| Anscombe | -0.75 | -0.38 |
| Method | reference group: method = WCLC | |
| WCLR | 1.11* | 4.14* |
| Bandwidth | reference group: bandwidth = 75 | |
| 125 | 0.00 | 0.00 |
| 175 | 0.97* | 3.65* |
| 250 | -4.00* | 3.39* |
| 750 | -10.61* | 3.34* |
| R ² cut-off | reference group: R ² cut-off = 0.25 | |
| 0.0 | -3.80* | -34.54* |
| 0.1 | -1.85* | -28.53* |
| 0.2 | -1.85* | -28.53* |
| 0.3 | -1.85* | -28.53* |

Note. WCLC = windowed cross-lagged correlation, WCLR = windowed cross-lagged regression, bandwidth refers to the frame number. We used a frequency of 25 frames per second for our analysis, size standard. = size standardization, IR = identification rate,

* indicates significant associations ($p < .05$), kappa = Cohen’s kappa, pr_out = percentage of over-identified frames.

<https://doi.org/10.1371/journal.pone.0211494.t003>

Naturally isolated and embedded condition

Second, we examined the data in the naturally isolated condition. We found only two configurations with identical identification of MSI and no MSI, which fulfilled our criteria of an acceptable IR. These configurations were *WCLC*, *log-transformed*, *no smoothing*, *R² cut-off = 0.25 with a bandwidth of 75 or 125 frames* (= 3 or 5 seconds). However, we also conducted an ordinal logistic regression to identify the best parameters scientifically. Note that we only used the configurations that were rated as good in the artificial condition; thus some parameter levels are missing. Results are displayed in [Table 4](#) (kappa: log-likelihood = 12.56, $df = 13$; pr_out: log-likelihood = -5.57, $df = 13$). As reference values, we used the best configuration that was identified (*WCLC*, *log-transformed*, *no smoothing*, *R² cut-off = 0.25*, *bandwidth = 75 frames*).

Result of the IR based on kappa showed that the reference configuration is superior to all other parameter settings. However, with respect to the IR based on pr_out there are no differences between different parameter configurations.

In the naturally embedded condition, no configuration was rated as having a good or acceptable IR. In [Table 5](#), means of kappa (MSI) and the average pr_out (no MSI) for the configurations with the best IR of the artificial and naturally isolated conditions as well as the artificial condition only are shown for each of the three time series types (artificial, naturally isolated, naturally embedded). In the naturally embedded time series, the IR is rather low; approximately 50% of frames are over-identified in the no MSI group.

Table 4. Results of the ordinal logistic regression with IR (kappa) as criterion and parameters as predictors (naturally isolated condition).

| | <i>b</i> (kappa) | <i>b</i> (pr_out) |
|------------------------|---|-------------------|
| Smoothing | reference group: smoothing = raw data | |
| Slight | -0.95 | 0.00 |
| Transformation | reference group: transformation = log-transformed | |
| Raw data | -28.70* | 0.01 |
| Size-standard. | -63.08* | -6.50 |
| Anscombe | -26.29* | -6.50 |
| Method | reference group: method = WCLC | |
| WCLR | 0.95 | 9.40 |
| Bandwidth | reference group: bandwidth = 75 | |
| 125 | 0.00 | 0.01 |
| 175 | -61.34* | -7.38 |
| R ² cut-off | reference group: R ² cut-off = 0.25 | |
| 0.0 | -23.57* | -8.97 |
| 0.1 | -21.06* | -8.06 |
| 0.2 | -21.06* | -8.06 |
| 0.3 | -21.06* | -8.06 |

Note. WCLC = windowed cross-lagged correlation, WCLR = windowed cross-lagged regression, bandwidth refers to the frame number. We used a frequency of 25 frames per second for our analysis, size standard. = size standardization, IR = identification rate,

* indicates significant associations ($p < .05$), kappa = Cohen’s kappa, pr_out = percentage of over-identified frames.

<https://doi.org/10.1371/journal.pone.0211494.t004>

Influence of the video sequence

Lastly, we examined the influence of the video sequence on kappa_{mean} and the amount of over-identified frames (pr_out_{mean}). We found a significant effect with respect to all three time series types when comparing kappa_{mean} between video sequences (artificial: $\chi^2 = 520.47$, $df = 9$, $p < .001$; naturally isolated: $\chi^2 = 552.81$, $df = 9$, $p < .001$; naturally embedded: $\chi^2 = 717.54$, $df = 9$, $p < .001$). Ranks of the sequences are displayed in [S1 Table](#). For example, video

Table 5. Cohen’s kappa (MSI) and over-identified frames pr_out (no MSI) of the two best configurations of the artificial and naturally isolated conditions and the best configurations of the artificial condition for all three time series types.

| Configuration | Sync sequences | | | No sync sequences | | |
|---|----------------|---------|---------|-------------------|---------|---------|
| | Simu | Nat iso | Nat emb | Simu | Nat iso | Nat emb |
| Best configurations artificial and naturally isolated synchrony | | | | | | |
| WCLC_75_log_nosmooth_0.25 | .83 | .64 | .02 | 1.20 | 2.13 | 57.66 |
| WCLC_125_log_nosmooth_0.25 | .83 | .64 | .02 | 1.20 | 2.13 | 57.66 |
| Best configurations artificial synchrony | | | | | | |
| WCLC_75_raw_slight_0.25 | .84 | .57 | -.02 | 0.50 | 2.82 | 44.63 |
| WCLC_75_size_slight_0.25 | .84 | .57 | -.02 | 0.00 | 2.82 | 43.27 |
| WCLC_125_raw_slight_0.25 | .84 | .57 | -.02 | 0.50 | 2.82 | 44.63 |

Note. M_{kappa} = mean of all stimuli (sequences) with respect to kappa_{mean}; pr_out = proportion of over-identified frames to all no sync frames; simu = artificial condition/simulated synchrony; nat iso = naturally isolated condition/synchrony; nat emb = naturally embedded condition/synchrony, WCLC = windowed cross-lagged correlation; 75 or 125 indicates bandwidth; raw = raw data; log = logarithmic transformed, size = size-standardized; slight = slight smoothing; 0.25 indicates the R² cut-off, WCLC = WCLC_F.

<https://doi.org/10.1371/journal.pone.0211494.t005>

sequence number 7 was the sequence with a very low kappa with respect to the naturally isolated condition. However, it was the best sequences referring to the artificial and naturally embedded condition. We also found a significant effect with respect to two time series types when comparing pr_out_mean between video sequences (naturally isolated: $\chi^2 = 1631.10$, $df = 9$, $p < .001$; naturally embedded: $\chi^2 = 817.70$, $df = 9$, $p < .001$). There were no differences for the artificial condition ($\chi^2 = 6.13$, $df = 9$, $p = .73$). For example, video sequence 1 was best with respect to the naturally isolated condition, whereas for the naturally embedded condition video sequence number 15 was best. Since we do not have any variables describing these sequences further, it remains an open question which characteristics of a video sequence are important for being a good sequence.

Discussion

The aim of this study was the validation of TSAM regarding their capability to correctly identify synchronization intervals given an external criterion. The peak-picking algorithm in combination with the WCLC or WCLR of Altmann [10], Altmann [11] allows the identification of MSI so that predefined synchronization intervals (e.g., rated by a human rater or artificially generated by a time series simulation) can be compared with MSI identified by TSAM. We tested the validity of varying configurations with MSI and no MSI to find a configuration with high power (correct positives) and low type I error (false positives) with respect to the identification of MSI. Former studies having tested TSAM or varied parameter settings are rare. Ramseier and Tschacher [49] showed that synchrony is dependent on the time lag and segment size by using different parameter settings in a case study. Altmann [10], Altmann [11] tested the WCLC and WCLR by using simulated time series in a case study. Therefore, the current study contributes to the examination of the impact of different parameter settings on the identification of movement synchrony by using $n = 60$ time series pairs.

Artificial condition

Our results show a clear influence of method, bandwidth, smoothing amount and R^2 cut-off on the identification rate (IR) in MSI. These parameters, except for the smoothing amount, also affect the frequency of over-identified frames in the no MSI condition. The WCLR is significantly superior to the WCLC. A bandwidth of 175 frames (= 7 seconds) leads to the best results, followed by smaller bandwidths such as 75 or 125 frames (= 3 and 5 seconds) regarding to the MSIs. With regard to smoothing, it is advisable to use raw data or to smooth the data slightly at most. An R^2 cut-off of 0.25 provides the best IR for MSI and no MSI.

For the subsequent analyses, we excluded configurations showing an acceptable or poor IR in the artificial condition. Of all $N = 600$ configurations within the artificial condition, there were 76 that displayed a good IR ($\kappa > .6$, $pr_out < .05$) for each of the video sequences. On average, these configurations had a very high inter-rater level of agreement between artificially generated MSI and computer-based ratings and a small proportion of over-identified frames ($M_{kappa_mean} = .77$, $SD_{kappa_mean} = .05$, $M_{pr_out_mean} = 0.32\%$, $SD_{pr_out_mean} = 0.70\%$). This indicates that the presented methods are able to detect perfect synchrony with a constant time lag (also known as echoing). The influences of specific parameters are discussed in detail in the following.

A parameter that affected the IR of synchrony was the amount of smoothing. So far, different smoothing amount have not yet been compared. The graphics in the introduction illustrate that if the time series does not contain any errors due to video defects, it is quite plausible to not smooth the data too much to avoid shifting or erasing peaks. Accordingly, our results

suggest that no smoothing or slight smoothing is advisable. If the time series contains such errors, filtering is important to avoid biases in MSI identification.

Transformation appears to have no significant influence on the identification of MSI. Transformations are performed to achieve stationarity and to transform the distribution into an approximately normal distribution [10]. The use of windows in the calculation of the MSI was also implemented to ensure local stationarity. It appears that the use of the windows is sufficient to deal with distribution-related challenges and there is no need for a stabilizing transformation.

Regarding the method used (WCLC vs. WCLR) we found a significant effect in the artificial condition. The WCLR showed higher kappas and lower false positives than the WCLC. The effect size was small to medium (Cramer's $V = 0.10$ – 0.37) in favor of WCLR. The difference between WCLC and WCLR may be explained by the fact that the WCLR was developed for cyclic or auto-correlated time series.

With regard to bandwidth, we found that rather small bandwidths lead to good results. Additionally, bandwidth seems to have the highest impact on results. Ramseyer and Tschacher [49] reported high synchrony values with bandwidths of 60 to 90 seconds (i.e., 600 to 900 frames). These contradictory results can be explained by the different synchrony measure and research question. While Ramseyer and Tschacher [49] calculated the strength of synchrony of the total interaction [17], our algorithms measured the frequency of synchrony [10, 17]. Moreover, while Ramseyer and Tschacher [49] classified the highest cross-correlation in comparison to surrogate-based pseudosynchrony, they did not compare the localization of a synchronization interval by a human rater/simulated MSIs to the identification by TSAM. Therefore, the criterion against which synchrony is compared differs: for Ramseyer and Tschacher [5] it is synchrony measured in surrogate datasets, in the current study it is either simulated perfect/imperfect synchrony, or synchrony rated by human raters. In addition, the different stimulus material has an influence on the results, as shown in our study (Ramseyer and Tschacher [49]: split-screen, two cameras; our scenario: naturalistic therapy sessions, one camera).

We introduced an R^2 cut-off to further control for spurious correlations. Our results suggest that an R^2 cut-off improves the IR. The multiple executions of the correlations or regressions to form the $m \times n$ R^2 matrix (m denotes the number of time points in frames, n indicates the number of time lags in frames) seem to have the effect that random intervals are identified as synchronous intervals. An increase of the R^2 cut-off can be understood as controlling for such spurious correlations. Note that this procedure is only valid with high true correlations and may be problematic if synchrony is weak.

Best configuration in the naturally isolated and artificial conditions

For further analyses, we only included configurations that had been classified as good in the artificial condition. Investigating the naturally isolated condition, we found only two configurations with an acceptable IR (WCLC, logarithmically transformed, not smoothed, $R^2 > 0.25$, bandwidth of 75 or 125 frames). It is noticeable that both configurations use logarithmically transformed data, seemingly contradictory to the results in the artificial condition. This inconsistency, however, can be explained as follows: The plots of the time series with identified synchronization intervals (Fig 5) display that the peaks have different shapes in the naturally isolated condition in contrast to the artificial condition. Therefore, a large movement can occur synchronously with a rather weak movement. This difference in the peak heights is matched by the logarithmic transformation and thus allows a better identification of the synchronization intervals compared to a time series that has not been logarithmically transformed

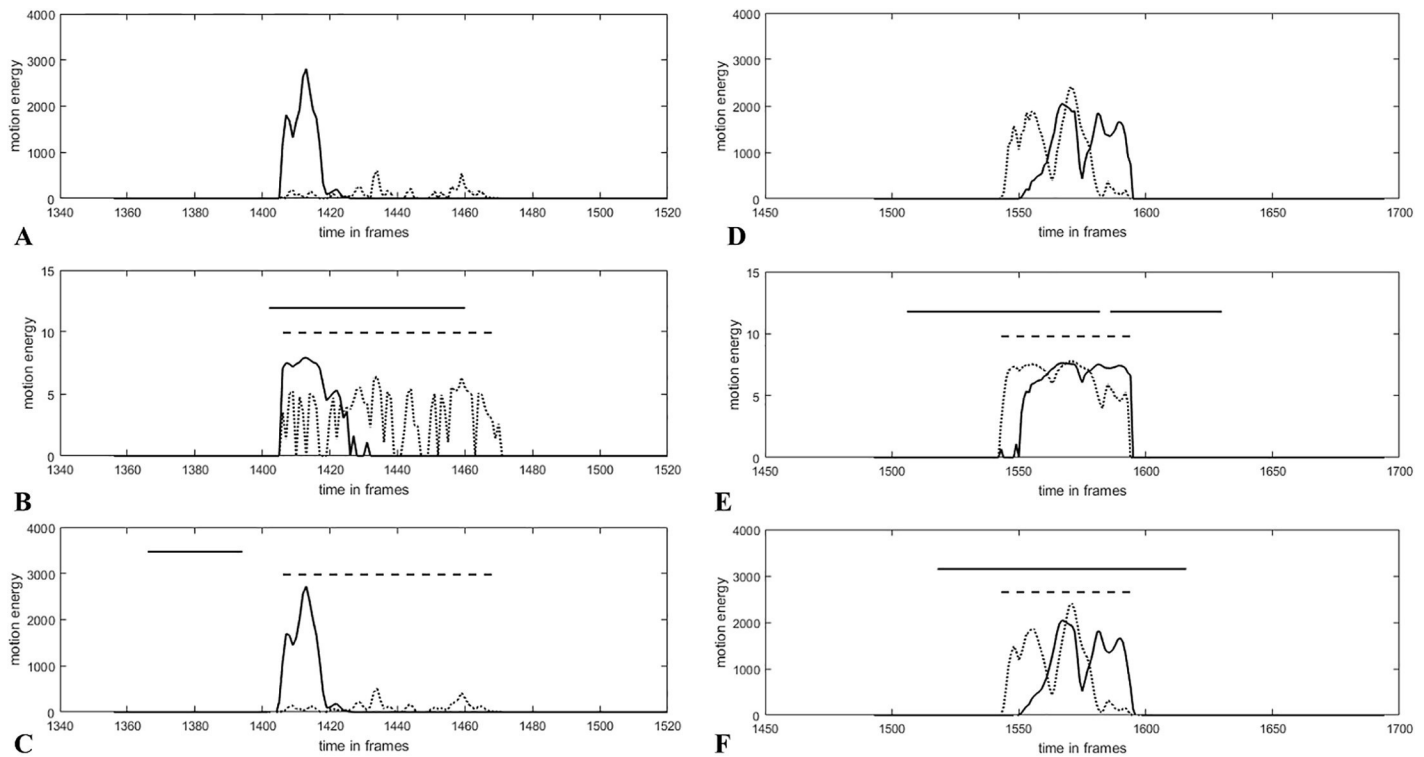


Fig 5. Time series (patient and therapist movements) of the naturally isolated condition. Solid horizontal lines indicate the computer-based synchrony interval (WCLC, bandwidth 125, $R^2 > 0.25$); dashed horizontal lines indicate the human-rated synchrony interval of the video sequence. (A) Time series of human-rated synchronization interval with different peak heights, (B) logarithmically transformed time series shown in A, (C) slightly smoothed and not transformed time series shown in A, (D) time series of human-rated synchronization intervals with similar peak heights, (E) logarithmically transformed time series shown in D, (F) slightly smoothed and not transformed time series shown in D.

<https://doi.org/10.1371/journal.pone.0211494.g005>

(Fig 5, Plot A, B, C). If both time series have approximately the same peak heights, the results differ only slightly (Fig 5, Plot D, E, F). In line with this, the results of the ordinal logistic regression show that the log-transformation is superior to other transformations in MSI detection, provided that the IR is acceptable for all videos. If the general influence of the parameters is investigated (see Supporting information, [S2 Appendix](#)), also the Anscombe transformation has a significant influence on the results. Therefore, we recommend using logarithmically transformed time series (or another variance-stabilizing transformation) for the identification of synchrony (based on human-rated synchronization intervals).

For bandwidth, the results point towards the preference of lower bandwidths (75 or 125) as indicated by both best configurations.

Another explanation for the differences between artificial and naturally isolated time series is that the characteristics of the reference material differ substantially. In the artificial condition, a synchronization interval is created as echoing of a given time series part so that the time series of both persons are perfectly equal with a predefined time lag. In contrast, in the naturally embedded configuration the “true” synchronization interval (respectively its beginning and end) is rated by human raters. Our study revealed that a lot of algorithms with very high identification quality in the artificial configuration failed in the naturally embedded configuration. This could mean that the algorithms had another synchrony concept than the human raters in our study. Furthermore, it should be noted that the algorithms and human raters had different levels of information about the nonverbal interpersonal interaction under study. Whereas the algorithms “decided” about the similarity of time series parts of an aggregated

movement measure (the motion energy), the decision of a human rater refers to the information in the video including gaze behavior or direction of movement (which were not available for the algorithms). On that reason, future studies should examine simulated stimulus material whereby the time-lag and the shape of the echoed time series are modified. The big advantage of simulated time series is that the synchronization intervals can be defined.

Naturally embedded condition

With respect to the naturally embedded time series, we did not find a configuration that reached a kappa of $> .4$ for all sequences. A reason for this finding may be that the rating of the synchronization intervals was carried out by three human raters who were really similar to each other. Therefore, the construct to which the algorithm-construct of synchrony was compared to was synchrony as rated by three young female psychologists. Movement synchrony was identified when there was an obvious relationship between the movements of both persons interacting. However, the construct measured by the computer-based methods disregards the context of the displayed movements. Only simultaneous or slightly time-delayed movements of the interaction partners are identified as synchrony. That is, an interval in which the therapist is looking for something in a folder and the patient scratches his head would be identified as synchrony by methods, but not by human raters. Fig 6 clearly displays the difference in ratings. The phenomenon between approx. 1125 frames and 1160 frames as well as the phenomenon between 1340 frames and 1360 frames are evaluated as not synchronous by the raters, but are nevertheless identified by the methods, leading to low rater agreement between human rating and computer rating. A possible solution to reduce the over-identification of frames could be to smooth small deflections (Fig 6, $t = 1375$ – 1400 frames) by applying a moving median (bandwidth of 5 values) [46], or to modify the peak-picking algorithm. Studies using a more representative sample of raters may also be requested. However, Bernieri, Reznick [39] showed that even with a large number of raters ($n = 20$), ratings are not very reliable. We used similar raters to not additionally include a rating bias based on gender or age because studies show an influence in synchrony (respectively perceiving synchrony) for these variables [50, 51]. Moreover, calculating correlations makes the assumption of a linear relationship between both time series. Allowing for other dependencies may also result in an improved detection of synchrony [52].

Another important aspect to consider is the output of the MEA. MEA aggregates all movements displayed in a video. Therefore, movements of arms, torso and head are aggregated to one movement index. A human rater can differentiate between these different body parts and also considers gaze direction and may therefore understand movement synchrony differently. Moreover, our motion analysis was based on 2-D video recordings and can only partially represent 3-dimensional movements. Another possibility for future research projects is the usage of 3-D cameras such as the Microsoft Kinect.

To summarize, we identified 76 of 600 configurations that allow a good identification of artificially generated MSI and a low false identification rate with respect to artificially generated no MSI. These configurations were found in the artificial condition. In the artificial condition, the time series pair was generated with perfect (time-delayed) synchrony. Therefore, 76 were able to identify perfect echoing with high concordance. These configurations used a low amount of smoothing, the WCLC or WCLR, an R^2 cut-off of 0.25 and rather low bandwidths. Furthermore, we were able to show that computer algorithms identify synchrony in a different way than humans, that is they detect similarity of time series with a good identification rate (artificial condition), but not multimodal synchrony within a context as rated by humans in the present study.

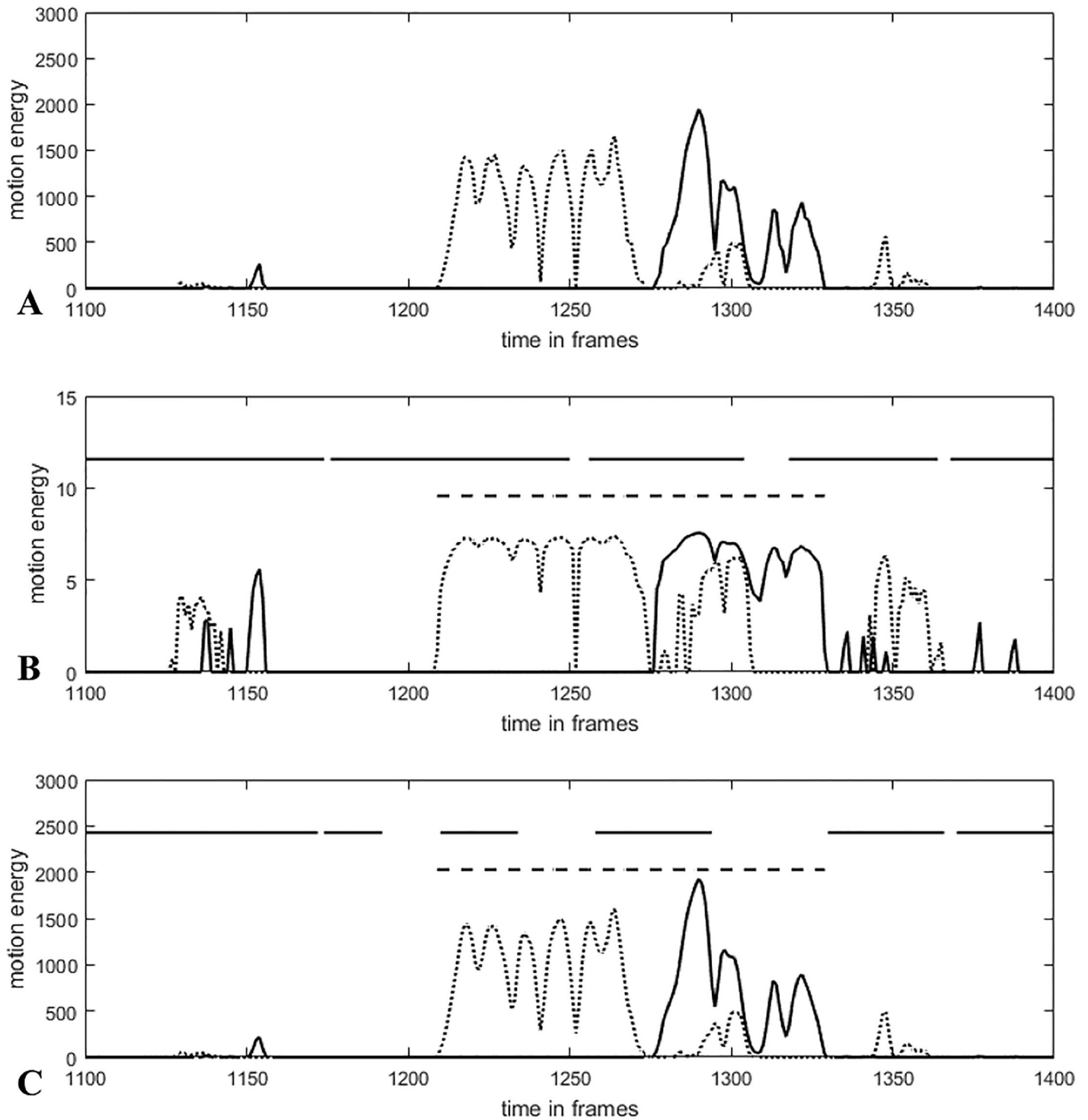


Fig 6. Time series (patient and therapist movements) of the naturally embedded condition. Solid horizontal lines indicate the computer-based synchrony interval (WCLC, bandwidth 125, $R^2 > 0.25$), dashed horizontal lines indicate the human-rated synchrony interval of the video sequence. (A) Time series of the human-rated synchrony phenomenon, (B) logarithmically transformed time series shown in A, (C) slightly smoothed and not transformed time series shown in A.

<https://doi.org/10.1371/journal.pone.0211494.g006>

Limitations & future research

To our knowledge, this study is the first testing parameter settings with respect to time series analysis methods in a validation study using $n = 60$ time series pairs and 600 different parameter configurations. Therefore, the current study contributes to the investigation of the validity of automated algorithms for the identification of nonverbal synchrony.

One limitation of the present study lies in the fact that we defined the *true* MSI by ratings of similar human raters. It is possible that different instructions and an extended training are required to better capture the construct. Although raters were instructed and familiar with the concept of movement synchrony, human raters have an internal model of synchrony that likely includes keeping eye contact and signals of other modalities that are not assessed by the computer-assisted methods for automated synchrony identification. The resulting difference could be reduced by further instructions or an extended rater training with regard to the concept of movement synchrony. However, it should be kept in mind that all algorithms (parameter configurations) studied in the naturally isolated and naturally embedded condition (in which human ratings were used as reference) were tested and approved by the good IR in the artificial condition. In the artificial condition, the true synchronization intervals were created by programming / simulation and are thus largely unbiased.

Furthermore, the time lag of synchrony was not systematically varied in the stimulus material. However, it is possible that time-lag has an impact on the IR. In a further study, it would also be important to systematically vary the time lag and more finely grade the parameters.

We were able to show that the stimulus material had a significant impact on the IR. This means that generalizability to other study designs is restricted. Further, in future studies, the amount of stimulus material should be increased. Our video sequences had specific characteristics (e. g., therapy videos of patients with social anxiety disorder, one camera setting) and a specific format (size: 640 x 480, etc.). The creation of a publicly accessible video database with annotations of synchronous movements could make a great contribution to the further development of TSAM. With the help of such a database, methods could be validated and the construct of nonverbal synchrony could be defined more explicitly. Such a database would also provide the opportunity to further investigate video characteristics and parameter settings.

Conclusion

Numerous studies have demonstrated that several problems occur with human ratings (e.g., reduction number of values of the assessed behavior, because of the rater's capacity, problems defining an equal construct between raters) and that the reliability of human raters is weak [53]. Additionally, human ratings require more time and personal resources and training seems to be important to ensure that all raters assess the same construct. In contrast, automated synchrony identification is economic and reliable. With our study, we demonstrated that the choice of parameter settings is essential for the generation of meaningful results using automated methods for synchrony detection. The tests of the artificial condition suggested that WCLR without smoothing, small bandwidths (75 or 125 frames) and an R^2 cut-off of 0.25 is best suited for the identification of perfect echoing. With respect to the naturally isolated condition, we identified WCLC with a bandwidth of 75 or 125 frames (3 or 5 seconds, respectively), no smoothing/slight smoothing, but a logarithmically transformed time series and an additional R^2 cut-off of 0.25 as the best configuration to detect the interpersonal synchrony of body movements (given acyclic time series and MSI showing two different complex peaks). This configuration also showed a good IR regarding high correct positives and low false positives in the artificial condition. In addition, we found low concordance in the naturally embedded condition due to the fact that human raters assessed movement synchrony differently than

the applied configurations. Human raters rate synchronous movements by including variables such as eye contact and context variables, resulting in another concept of synchrony. Movement synchrony of our best algorithm goes beyond this definition by assessing all simultaneous or slightly time-delayed highly correlated time series trajectories. Thus, unconscious processes resulting in movement synchrony, which may also have an influence on the therapeutic process, are additionally assessed.

Supporting information

S1 Appendix. Influence of the video sequence.

(DOCX)

S2 Appendix. General influence of the parameters in all conditions.

(DOCX)

S1 Table. Ranking of the different video sequences based on Kruskal-Wallis-tests.

(DOCX)

S2 Table. Coefficients and significance of the ordinal logistic mixed effects regression (criterion IR by kappa, MSI) for all conditions.

(DOCX)

S3 Table. Coefficients and significance of the ordinal logistic mixed effects regression (criterion IR by pr_out, noMSI) for all conditions.

(DOCX)

Acknowledgments

We cordially thank U. Willutzki, St. Herpertz, J. Hoyer, P. Joraschky, W. Hiller, E. Leibing, M. Beutel and F. Leichsenring for providing video recordings of the project. We thank K. Boyle, native speaker, for English proof reading. We thank Lukas A. Knitter and Alisa Shugaley for student assistance.

Author Contributions

Conceptualization: Désirée Schoenherr, Uwe Altmann.

Data curation: Désirée Schoenherr.

Formal analysis: Désirée Schoenherr.

Funding acquisition: Bernhard M. Strauss, Wolfgang Lutz.

Investigation: Désirée Schoenherr.

Methodology: Désirée Schoenherr, Uwe Altmann.

Project administration: Bernhard M. Strauss, Wolfgang Lutz, Uwe Altmann.

Resources: Bernhard M. Strauss, Ulrich Stangier.

Software: Uwe Altmann.

Supervision: Uwe Altmann.

Writing – original draft: Désirée Schoenherr.

Writing – review & editing: Désirée Schoenherr, Jane Paulick, Bernhard M. Strauss, Anne-Katharina Deisenhofer, Brian Schwartz, Julian A. Rubel, Wolfgang Lutz, Uwe Altmann.

References

1. Tonti M, Gelo OCG. Rate of speech and emotional-cognitive regulation in the psychotherapeutic process: a pilot study. *Research in Psychotherapy: Psychopathology, Process and Outcome*. 2016; 19(2). <https://doi.org/10.4081/ripppo.2016.232>
2. Greene MG, Adelman RD, Friedmann E, Charon R. Older patient satisfaction with communication during an initial medical encounter. *Social Science & Medicine*. 1994; 38(9):1279–88.
3. Paulick J, Deisenhofer A-K, Ramseyer F, Tschacher W, Boyle K, Rubel J, et al. Nonverbal Synchrony: A new approach to better understand psychotherapeutic processes and drop-out. *Journal of Psychotherapy Integration*. 2017.
4. Galbusera L, Finn MT, Fuchs T. Interactional synchrony and negative symptoms: An outcome study of body-oriented psychotherapy for schizophrenia. *Psychother Res*. 2016:1–13.
5. Ramseyer F, Tschacher W. Nonverbal Synchrony in Psychotherapy: Coordinated Body Movement Reflects Relationship Quality and Outcome. *Journal of Consulting and Clinical Psychology*. 2011; 79(3):284–95. PMID: [21639608](https://pubmed.ncbi.nlm.nih.gov/21639608/)
6. Watanabe T. A study of motion-voice synchronization. *Bulletin of JSME*. 1983; 26(222):2244–50. <https://doi.org/10.1299/jsme1958.26.2244>
7. Messinger DS, Mahoor MH, Chow S-M, Cohn JF. Automated measurement of facial expression in infant–mother interaction: A pilot study. *Infancy*. 2009; 14(3):285–305. <https://doi.org/10.1080/15250000902839963> PMID: [19885384](https://pubmed.ncbi.nlm.nih.gov/19885384/)
8. Feldman R. Parent–infant synchrony and the construction of shared timing; physiological precursors, developmental outcomes, and risk conditions. *Journal of Child psychology and Psychiatry*. 2007; 48(3–4):329–54. <https://doi.org/10.1111/j.1469-7610.2006.01701.x> PMID: [17355401](https://pubmed.ncbi.nlm.nih.gov/17355401/)
9. Bailenson JN, Yee N. Digital chameleons: Automatic assimilation of nonverbal gestures in immersive virtual environments. *Psychological science*. 2005; 16(10):814–9. <https://doi.org/10.1111/j.1467-9280.2005.01619.x> PMID: [16181445](https://pubmed.ncbi.nlm.nih.gov/16181445/)
10. Altmann U. Investigation of movement synchrony using windowed crosslagged regression. *Analysis of Verbal and Nonverbal Communication and Enactment: The Processing Issue*. 2011:344–54.
11. Altmann U. Synchronisation nonverbalen Verhaltens: Weiterentwicklung und Anwendung zeitreihenanalytischer Identifikationsverfahren. [Synchronization of nonverbal behavior: Development and Application of time series analysis methods]: Springer-Verlag; 2013.
12. Grammer K, Kruck K, Magnusson M. The courtship dance: Patterns of nonverbal synchronization in opposite-sex encounters. *Journal of Nonverbal Behavior*. 1998; 22(1):3–29. <https://doi.org/10.1023/A:1022986608835>
13. Bernieri FJ, Rosenthal R. Interpersonal coordination: Behavior matching and interactional synchrony. In: Feldman R, Rimé B, editors. *Fundamentals of nonverbal behavior*. xiv. New York, NY, US: Cambridge University Press; Paris, France: Editions de la Maison des Sciences de l'Homme; 1991. p. 511 pp.
14. Chetouani M, Delaherche E, Dumas G, Cohen D. 15 Interpersonal Synchrony: From Social Perception to Social Interaction. *Social Signal Processing*. 2017:202.
15. Delaherche E, Chetouani M, editors. *Multimodal coordination: exploring relevant features and measures*. Proceedings of the 2nd international workshop on Social signal processing; 2010: ACM.
16. Boker SM, Rotondo JL, Xu M, King K. Windowed cross-correlation and peak picking for the analysis of variability in the association between behavioral time series. *Psychological Methods*. 2002; 7(3):338. <https://doi.org/10.1037/1082-989X.7.3.338> PMID: [12243305](https://pubmed.ncbi.nlm.nih.gov/12243305/)
17. Schoenherr D, Paulick J, Worrack S, Strauss B, Rubel JA, Schwartz B, et al. Quantification of nonverbal synchrony using linear time series analysis methods: Lack of convergent validity and evidence for facets of synchrony. *Behavior Research Methods*. 2018:1–23.
18. Moulder RG, Boker SM, Ramseyer F, Tschacher W. Determining synchrony between behavioral time series: An application of surrogate data generation for establishing falsifiable null-hypotheses. *Psychol Methods*. 2018. <https://doi.org/10.1037/met0000172> PMID: [29595296](https://pubmed.ncbi.nlm.nih.gov/29595296/).
19. Ashenfelter KT, Boker SM, Waddell JR, Vitanov N. Spatiotemporal symmetry and multifractal structure of head movements during dyadic conversation. *Journal of Experimental Psychology: Human Perception and Performance*. 2009; 35(4):1072. <https://doi.org/10.1037/a0015017> PMID: [19653750](https://pubmed.ncbi.nlm.nih.gov/19653750/)
20. Bilakhia S, Petridis S, Nijholt A, Pantic M. The MAHNOB Mimicry Database: A database of naturalistic human interactions. *Pattern recognition letters*. 2015; 66:52–61. <https://doi.org/10.1016/j.patrec.2015.03.005>
21. Boker SM, Rotondo JL. Symmetry building and symmetry breaking in synchronized movement. *Mirror neurons and the evolution of brain and language*. 2002; 42:163.

22. Bozkurt E, Yemez Y, Erzin E. Multimodal analysis of speech and arm motion for prosody-driven synthesis of beat gestures. *Speech Communication*. 2016; 85:29–42. <https://doi.org/10.1016/j.specom.2016.10.004>
23. Campbell N. An audio-visual approach to measuring discourse synchrony in multimodal conversation data: September; 2009.
24. Michelet S, Karp K, Delaherche E, Achard C, Chetouani M. Automatic imitation assessment in interaction. *Human Behavior Understanding*: Springer; 2012. p. 161–73.
25. Nagaoka C, Komori M. Body movement synchrony in psychotherapeutic counseling: A study using the video-based quantification method. *IEICE transactions on information and systems*. 2008; 91(6):1634–40. <https://doi.org/10.1093/ietisy/e91-d.6.1634>
26. Paxton A, Dale R. Frame-differencing methods for measuring bodily synchrony in conversation. *Behavior Research Methods*. 2013; 45(2):329–43. <https://doi.org/10.3758/s13428-012-0249-2> PMID: [23055158](https://pubmed.ncbi.nlm.nih.gov/23055158/)
27. Ramseyer F, Tschacher W. Nonverbal synchrony of head- and body-movement in psychotherapy: different signals have different associations with outcome. *Frontiers in Psychology*. 2014; 5(979). <https://doi.org/10.3389/fpsyg.2014.00979> PMID: [25249994](https://pubmed.ncbi.nlm.nih.gov/25249994/)
28. Ramseyer F, Tschacher W. Synchrony in dyadic psychotherapy sessions. *Simultaneity: Temporal structures and observer perspectives*. 2008:329–47.
29. Sun X, Nijholt A, Truong KP, Pantic M, editors. Automatic visual mimicry expression analysis in interpersonal interaction 2011: IEEE.
30. Tschacher W, Rees G, Ramseyer F. Nonverbal synchrony and affect in dyadic interactions. *Frontiers in Psychology*. 2014; 5(1323). <https://doi.org/10.3389/fpsyg.2014.01323> PMID: [25505435](https://pubmed.ncbi.nlm.nih.gov/25505435/)
31. Tronick ED, Als H, Brazelton TB. Mutuality in mother-infant interaction. *Journal of communication*. 1977; 27(2):74–9. <https://doi.org/10.1111/j.1460-2466.1977.tb01829.x> PMID: [853138](https://pubmed.ncbi.nlm.nih.gov/853138/)
32. Yang N, Wang Z, Hu W. Synchrony expression analysis of human-human interaction using wearable sensors. *Control and Automation (ICCA)*. 2016:611–5.
33. Grammer K, Honda M, Juette A, Schmitt A. Fuzziness of nonverbal courtship communication unblurred by motion energy detection. *Journal of personality and social psychology*. 1999; 77(3):487. <https://doi.org/10.1037/0022-3514.77.3.487> PMID: [10510505](https://pubmed.ncbi.nlm.nih.gov/10510505/)
34. Anscombe FJ. The transformation of Poisson, binomial and negative-binomial data. *Biometrika*. 1948; 35(3/4):246–54. <https://doi.org/10.2307/2332343>
35. Box GE, Cox DR. An analysis of transformations. *Journal of the Royal Statistical Society Series B (Methodological)*. 1964:211–52.
36. Safran JD, Muran JC, Eubanks-Carter C. Repairing alliance ruptures. *Psychotherapy*. 2011; 48(1):80. <https://doi.org/10.1037/a0022140> PMID: [21401278](https://pubmed.ncbi.nlm.nih.gov/21401278/)
37. Schönbrodt FD, Perugini M. At what sample size do correlations stabilize? *Journal of Research in Personality*. 2013; 47(5):609–12. <https://doi.org/10.1016/j.jrp.2013.05.009>
38. Ramseyer F, Tschacher W. Nonverbal synchrony or random coincidence? How to tell the difference. *Development of multimodal interfaces: active listening and synchrony*: Springer; 2010. p. 182–96.
39. Bernieri FJ, Reznick JS, Rosenthal R. Synchrony, pseudosynchrony, and dissynchrony: Measuring the entrainment process in mother-infant interactions. *Journal of personality and social psychology*. 1988; 54(2):243. <https://doi.org/10.1037/0022-3514.54.2.243>
40. Louwse MM, Dale R, Bard EG, Jeuniaux P. Behavior matching in multimodal communication is synchronized. *Cognitive science*. 2012; 36(8):1404–26. <https://doi.org/10.1111/j.1551-6709.2012.01269.x> PMID: [22984793](https://pubmed.ncbi.nlm.nih.gov/22984793/)
41. Gottman JM, Ringland JT. The analysis of dominance and bidirectionality in social development. *Child Development*. 1981:393–412. <https://doi.org/10.2307/1129157>
42. Dean RT, Dunsmuir WT. Dangers and uses of cross-correlation in analyzing time series in perception, performance, movement, and neuroscience: The importance of constructing transfer function autoregressive models. *Behavior research methods*. 2016:1–20.
43. Leichsenring F, Salzer S, Beutel ME, Herpertz S, Hiller W, Hoyer J, et al. Long-term outcome of psychodynamic therapy and cognitive-behavioral therapy in social anxiety disorder. *American Journal of Psychiatry*. 2014.
44. Leichsenring F, Salzer S, Beutel ME, Herpertz S, Hiller W, Hoyer J, et al. Psychodynamic therapy and cognitive-behavioral therapy in social anxiety disorder: a multicenter randomized controlled trial. *American Journal of Psychiatry*. 2013. <https://doi.org/10.1176/appi.ajp.2013.12081125> PMID: [23680854](https://pubmed.ncbi.nlm.nih.gov/23680854/)
45. Anvsoft Inc. Any Video Converter. 2009.

46. Altmann U, Schoenherr D, Paulick J, Knitter L, Worrack S, Schiefele A-K, et al. Introduction, Practical Guide, and Validation Study for Measuring Body Movements Using Motion Energy Analysis. submitted. 2019.
47. Landis JR., Koch GG. The Measurement of Observer Agreement for Categorical Data. *Biometrics*. 1977; 33(1):159–74. <https://doi.org/10.2307/2529310> PMID: 843571
48. R Core Team. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2017.
49. Ramseyer F, Tschacher W. Movement Coordination in Psychotherapy: Synchrony of Hand Movements is Associated with Session Outcome. A Single-Case Study. *Nonlinear dynamics, psychology, and life sciences*. 2016; 20(2):145–66. PMID: 27033131
50. Stel M, van Knippenberg A. The role of facial mimicry in the recognition of affect. *Psychological Science*. 2008; 19(10):984–5. <https://doi.org/10.1111/j.1467-9280.2008.02188.x> PMID: 19000207
51. Namy LL, Nygaard LC, Sauerteig D. Gender differences in vocal accommodation: The role of perception. *Journal of Language and Social Psychology*. 2002; 21(4):422–32. <https://doi.org/10.1177/026192702237958>
52. Dionisio A, Menezes R, Mendes DA. Mutual information: a measure of dependency for nonlinear time series. *Physica A: Statistical Mechanics and its Applications*. 2004; 344(1–2):326–9. <https://doi.org/10.1016/j.physa.2004.06.144>
53. Altorfer A, Jossen S, Würmle O. Eine Methode zur zeitgenauen Aufnahme und Analyse des Bewegungsverhaltens. *Zeitschrift für Psychologie*. 1997; 205(1):83–117.

Publikation 3

Nonverbal synchrony predicts premature termination of psychotherapy for social anxiety disorder

Désirée Schoenherr¹, Jane Paulick², Bernhard M. Strauss¹, Anne-Katharina Deisenhofer², Brian Schwartz², Julian Rubel², Wolfgang Lutz², Ulrich Stangier³, Uwe Altmann¹

¹University Hospital Jena, Institute of Psychosocial Medicine and Psychotherapy, Jena, Germany

²Trier University, Department of Clinical Psychology and Psychotherapy, Trier, Germany

³Goethe University Frankfurt/Main, Department of Clinical Psychology and Psychotherapy, Frankfurt/Main, Germany

Acknowledgement

This study was funded by the German Research Foundation (DFG project “Timing of nonverbal patient-therapist-interaction and therapeutic success of social phobic patients (TIMPATHIN)”, funding number GZ: STR 306/28-1 and LU 660/8-1) and uses video recordings from the SOPHONET treatment study funded by the German Federal Ministry of Education and Research (BMBF, FKZ 01GV0607). We cordially thank U. Willutzki, St. Herpertz, J. Hoyer, P. Joraschky, W. Hiller, E. Leibing, M. Beutel and F. Leichsenring for providing video recordings of the project. We thank K. Boyle and San Francisco Edit for English proof reading. We also thank L. Knitter and A. Shugaley (student assistants) for assistance in pre-processing the data.

Abstract

Premature termination is a problem in psychotherapy. In addition to the examination of demographic and clinical variables as predictors of dropout, research indicates the importance of dyadic variables. Nonverbal synchrony (e.g., movement synchrony) operationalizes the coordination of patient and therapist and is a promising candidate for predicting premature termination. This secondary data analysis included data on patients with social anxiety disorder (N=267) that were treated with >20 sessions of cognitive-behavioral therapy or psychodynamic therapy. Therapy outcome was measured by the Inventory of Interpersonal Problems and the Beck Depression Inventory. Individual movements in the third session were assessed by motion energy analysis. Movement synchrony was identified with a windowed cross-lagged correlation and peak-picking algorithm. We performed logistic regressions and mixed effects Cox regressions to investigate synchrony as a predictor of premature termination. Therapist-patient dyads that included a patient who terminated psychotherapy prematurely had significantly lower movement synchrony at the beginning of therapy than patients who completed therapy. Gender-matching and therapeutic approach had a (marginally) significant effect in moderating the relationship. Therefore, low movement synchrony in early therapy sessions may contain clues to premature termination and reflect a mismatch between the patient and therapist or problems in the therapeutic alliance.

Keywords: social anxiety disorder, premature termination, movement synchrony, motion energy analysis, windowed cross-lagged correlation

Clinical Impact Statement

Question: Movement synchrony at the beginning of therapy, measured by frame-by-frame coding, was investigated as a predictor of premature termination of psychotherapy. *Findings:* High nonverbal synchrony was associated with a low rate of premature therapy termination in patients with social anxiety disorder. *Meaning:* Because premature termination is generally

viewed as a failure in psychotherapy, information on nonverbal aspects may help therapists prevent premature termination. *Next steps*: Nonverbal elements should be explicitly integrated into psychotherapy practice.

Social anxiety disorder (SAD) is a common psychological disorder, with a 12.1% lifetime prevalence and 7.1% 1 year prevalence (Ruscio et al., 2008). SAD is often concurrent with other disorders, such as depression or other anxiety disorders (Kessler, 2003). SAD is characterized by discomfort in social interactions and concerns about being judged or humiliated by others, with patients experiencing fear in specific social situations that they then avoid (American Psychiatric Association, 2013). At the cognitive level, SAD is characterized by increased self-focus (Clark & Wells, 1995; Heerey & Kring, 2007). Patients with SAD monitor their own thoughts and behavior; thus, their attention shifts towards the self, which manifests in verbal (e.g., increased self-talk) and nonverbal (e.g., avoidance of eye contact, increased self-touch) behaviors (Farabee, Ramsey, & Cole, 1993; Heerey & Kring, 2007). Furthermore, studies have shown that patients with SAD lack interpersonal skills in conversations (Voncken & Bögels, 2008). Therefore, people that interact with individuals with social anxiety are more likely to experience a communicative imbalance and evaluate the interaction quality as low (Heerey & Kring, 2007). This feeling may arise from a disturbance in the nonverbal responsiveness (i.e., capability to synchronize with the interaction partner on a nonverbal level).

Nonverbal synchrony is a new concept that operationalizes how much contact individuals have with each other and refers to a phenomenon that evokes in an observer the impression that the behaviors of interacting partners are interrelated (Altmann, 2013; Bernieri & Rosenthal, 1991). Because the social context and being in contact with other individuals is impaired in SAD, the investigation of nonverbal synchrony may lead to new insights into the symptomatology and psychotherapy process. Psychotherapy research has increasingly examined interactions between nonverbal patient and therapist behaviors by evaluating nonverbal synchrony in different (mostly disorder-heterogeneous) samples. The findings

suggest a relationship between alliance, outcome measures, and nonverbal synchrony (Galbusera, Finn, & Fuchs, 2016; Paulick et al., 2017; Ramseyer & Tschacher, 2011).

The In-Sync model of psychotherapy assumes that the perceptual-motor processes of nonverbal synchrony are the basis for forming higher cognition and a therapeutic relationship through, for example, sharing a subjective experience with another person (Koole & Tschacher, 2016). Therefore, nonverbal synchrony can influence one's evaluation of an interaction, of the interaction partner, and one's prosocial behavior. For example, individuals who move in parallel may perceive their pair as a unit and rate their interaction partner as sympathetic (Vicaria & Dickens, 2016). The neuronal basis of nonverbal synchrony may be the mirror neuron system. Observing a movement activates the neurons in the observer's brain, typically the same neurons that are involved in executing the action (mirror neurons). Thus, perception and action are automatically linked, which facilitates the initiation of a movement or posture similar to what is observed (Iacoboni, 2009). One possible mechanism of synchrony is simultaneous actions blurring the boundaries between self and the other person (Mogan, Fischer, & Bulbulia, 2017). Thus, integration of the other person and self occurs, resulting in the application of positive, self-related attributions to the dyad. Furthermore, attention shifts towards the dyad and interpersonal variables gain importance (Mogan et al., 2017). In addition, synchrony facilitates a feeling of social connection/belonging (Marsh, Richardson, & Schmidt, 2009).

However, the connection postulated by the In-Sync model between nonverbal synchrony and a therapeutic alliance or success appears to hold only under specific conditions. For example, studies of movement synchrony have shown significant correlations with the therapeutic outcome (Ramseyer & Tschacher, 2011) and premature termination (Paulick et al., 2017) in patients treated with cognitive-behavioral therapy (CBT). However, the effect on outcome found by Ramseyer and Tschacher (2011) was not found by Paulick et al. (2017),

despite using the same methodology. Differences between these studies were found in post-hoc comparisons of responders and non-responders. Particularly, Paulick et al. (2017) found that patients with an improvement and regular termination had a modest level of synchrony, whereas patients who dropped out had low levels of synchrony and patients with regularly ended therapy without improvement had high levels of synchrony. In addition, among patients with schizophrenia, a marginally significant positive effect indicated that high synchrony in body-oriented therapy was present in dyads that included patients with improved symptomatology at the end of therapy (Galbusera et al., 2016). Findings on the synchrony of facial expressions have contrasted with others showing a largely positive connection between movement synchrony and therapeutic success. One study with a heterogeneous sample found a significant correlation between high facial expression synchrony and therapy failure (Rasting & Beutel, 2005). This raised the question of whether the connection between nonverbal synchrony and therapy success depends on the psychological disorder studied (e.g., affective vs. schizophrenic) or the behavior modality measured (e.g., body movements vs. facial expressions). In addition, studies have suggested that other variables, such as gender-matching (Grammer, Kruck, & Magnusson, 1998) or the context (Lakin, Jefferis, Cheng, & Chartrand, 2003; Paxton & Dale, 2013a), may influence the association between nonverbal synchrony and therapeutic success.

In the present study, we examined patients with SAD because the disorder is inherently connected to social interactions and influences the behavior of the interaction partner. Studies in experimental settings have shown that individuals with social anxiety do not experience the positive consequences of nonverbal synchrony (Vrijnsen, Lange, Dotsch, Wigboldus, & Rinck, 2010). Others have found that individuals with social anxiety mimic expressions to a lesser extent than normal, which results in low interaction quality (Vrijnsen, Lange, Becker, & Rinck, 2010). However, some studies have suggested that individuals with social anxiety do not lack

mimicry of all expressions, as they mimic positive expressions, such as smiling, to a greater extent than normal (Dijk, Fischer, Morina, van Eeuwijk, & van Kleef, 2018).

Premature termination is a general problem in psychotherapy. Overall, approximately every fifth patient drops out of therapy prematurely, and approximately 15% of patients with anxiety disorders terminate therapy prematurely (Hans & Hiller, 2013; Swift & Greenberg, 2012). General predictors can be categorized according to the treatment (e.g., manual-guided), setting (e.g., e-therapy), therapist characteristics (e.g., experience), and patient characteristics (e.g., demographic, clinical variables) (Swift & Greenberg, 2012); the latter category is the most commonly investigated. According to some studies (Swift & Greenberg, 2012; Zimmermann, Rubel, Page, & Lutz, 2017), gender and employment do not predict therapy dropout, but other studies have demonstrated that these variables could influence dropout rates (Altmann et al., 2014; Saxon, Barkham, Foster, & Parry, 2017). No valid predictors have been identified for preventing patients with SAD from prematurely terminating therapy. Some studies have shown that premature therapy termination is affected by clinical variables, demographic variables, or other comorbid disorders (e.g., depression), but most studies have not supported these findings (Eskildsen, Hougaard, & Rosenberg, 2010; Hoyer et al., 2014). Therefore, it may be particularly interesting to consider other variables as predictors of premature termination. Accordingly, Swift and Greenberg (2012) highlighted the need for investigating process variables. Nonverbal synchrony between the patient and therapist represents a new approach for examining the process of therapy.

The present study aimed to investigate the frequency of movement synchrony as a predictor of premature termination in patients with SAD. Movement synchrony was defined as the frequency of synchronization intervals in which a linear, video-based covariation could be detected in movements. Paulick et al. (2017) demonstrated that movement synchrony is lowest for patients who do not improve from therapy and end therapy prematurely. With respect to

SAD patients, studies have reported generally low levels of nonverbal synchrony. We hypothesize that patients with a low synchronization frequency terminate therapy prematurely. Here, we controlled for therapist effects on premature termination and conducted subgroup analyses to explore the influence of other variables (e.g., the treatment condition).

Methods

Data Source

This retrospective study is a secondary data analysis. Data for the study were retrieved from 1) a previous multicenter study conducted by the Social Phobia Psychotherapy Network (SOPHO-NET; Leichsenring et al., 2009, 2013) and from 2) a database in an outpatient clinic in southwest Germany that routinely performed video-recordings and data collection between 2008 and 2016. Therapeutic approaches in the multicenter SOPHO-NET study were short-term manual-guided CBT (CBT_{man}) or short-term manual-guided psychodynamic-oriented psychotherapy (PDT_{man}). Participants were recruited at five different outpatient centers in Germany. Each center provided both CBT_{man} and PDT_{man} psychotherapies. Before inclusion, therapists were trained (CBT_{man} vs. PDT_{man}) and had to treat two pilot cases in accordance with the manual under supervision. Continuous site and cross-site supervision methods were applied to ensure treatment quality and manual adherence. The other study center was an outpatient university clinic in south-west Germany that performed naturalistic cognitive-behavioral psychotherapy (CBT_{nat}). This video study was approved by the local ethics committees at all involved centers and was conducted in accordance with the Guidelines for Good Clinical Practice.

Treatment

Participating therapists were either advanced postgraduates or licensed therapists. Therapists in training were supervised every fourth session. CBT_{man} and PDT_{man} included five preparatory sessions, which are compulsory in the German health care system and cover mostly diagnostic and administrative issues. These sessions were followed by 25 50-min weekly therapy sessions. CBT_{man} was based on Clark and Well's model and followed the Stangier, Clark, and Ehlers (2006) manual (Leichsenring et al., 2013; Stangier et al., 2006). The average treatment duration was 26.0 (SD = 4.9) sessions for patients in the CBT_{man} arm. PDT_{man} was based on Luborsky's model of psychodynamic therapy. During the initial sessions, interviews were conducted to collect information on important episodes in relationships between the patient and other persons (Leichsenring, Beutel, & Leibing, 2007). The average treatment duration was 21.7 (SD = 7.3) sessions in the PDT_{man} arm.

CBT_{nat} did not follow a manual or have a pre-defined treatment duration. Typically, treatment was provided in 50-min sessions on a weekly basis. Treatment was based on an integrative CBT approach (Lutz, Schiefele, Wucherpennig, Rubel, & Stulz, 2016). The therapy was planned based on the patient's symptoms and terminated consensually. On average, patients in this group received 42.7 (SD = 18.9) treatment sessions before termination. In this arm, long-term therapies were pursued if the patient required more psychotherapy.

In all arms, session 1 mostly covered administrative issues and an assessment of the medical history and questionnaire data. Therapists treated an average of 2.25 patients (range: 1 to 10).

Study Design

All sessions were video-recorded after obtaining patient and therapist consent. The patient and therapist sat obliquely to one another. Videos of the manual-guided therapy were

recorded by a one-hand camera; videos of CBT_{nat} were recorded using two cameras with split-screen technology.

Participants

In the CBT_{man} and PDT_{man} arms, patients were included when they had a primary diagnosis of SAD according to the Structural Clinical Interview for DSM-IV (SCID; Wittchen, Wunderlich, Gruschwitz, & Zaudig, 1997) and Anxiety Disorders Interview Schedule (Brown, Barlow, & DiNardo, 1994) and a score >30 on the Liebowitz Social Anxiety Scale (LSAS; Mennin et al., 2002). Patients were excluded when they had cluster A or B personality disorders, substance abuse (acute), psychotic symptoms, risk of self-harm, mental disorders due to a known physiological condition, severe medical conditions, or other psychotherapeutic or psychopharmacological interventions. After applying these clinical inclusion and exclusion criteria, 495 patients were included in the study.

The naturalistic group included patients diagnosed with SAD as a primary or secondary diagnosis based on the SCID. A clinical interview was conducted by intensively trained, independent clinicians before actual therapy began. Patients with substance dependencies and psychosis (acute) were excluded. After applying the clinical inclusion and exclusion criteria, 134 patients were included in this arm, including 20 patients who also participated in the Paulick et al. (2017) study. However, independent nonverbal synchrony measures were used because the two studies analyzed different therapy sessions (i.e., different video files). In all arms, we excluded patients who terminated therapy before the third session.

Patients were also screened using technical inclusion criteria. We required at least one video from the 3rd, 8th, 20th, or 29th therapy session (± 1 session) because they are close to the SOPHO-NET measurement time points. This criterion excluded 274 patients. Furthermore, the patient and therapist had to be visible in the video for 15 min during the first half of the therapy session, after questionnaire completion and clarification of administrative issues, and the

selected session had to be based on verbal interactions. Therefore, sessions that applied other specific techniques (e.g., progressive muscle relaxation) were excluded. Videos were also excluded when one of the interacting individuals stood up during the 15 min of relevant recording, when a third person was present, when less than 15 min were available in the first half of the session, when video quality was low, and when any overlap occurred between the patient's and therapist's bodies in the video. These criteria excluded 88 patients. Patients in all three arms provided informed consent for study participation. The final analyses included 267 dyads with 267 patients and 119 therapists.

Measures

Premature termination. Premature termination was operationalized as termination by the patient without planning or reaching a consensus with the therapist about ending therapy. In the CBT_{man} and PDT_{man} arms, patients also had to drop out before the 30th session. Therapy records were reviewed to obtain this information. In addition, we considered the last session attended by the patient; therefore, it was possible to relate the categorical event of dropout to the time when the patient terminated therapy. According to discussions about the definition of dropout, we also included a clinically significant change (CSC) in interpersonal problems as a criterion. We differentiated between early and late dropout using the 10th session as a cut-off in accordance with Hynan (1990), who used the fifth session as a cut-off. The reason for the difference is that, in the German health care system, the first five sessions are probatoric to examine the indication for therapy before requesting it from health insurance. Thus, we considered various dropout variables: dichotomous; categorical: early dropout (before 10th session), late dropout (after 10th session), regular terminated; CSC, dichotomous; and time until dropout (survival analysis).

Initial symptom severity was measured in terms of interpersonal problems and depression, which were common measures in this multicenter study.

Inventory of interpersonal problems. Interpersonal problems were assessed in the first session using the 32-item version of the German Inventory of Interpersonal Problems (IIP; Thomas, Brähler, & Strauß, 2011). This questionnaire assessed self-reported difficulties in interactions with other people and contained eight scales (domineering, vindictive, cold, socially inhibited, nonassertive, overly accommodating, self-sacrificing, and intrusive) with four items each. Responses to each item were selected from a 5-point Likert scale ranging from 0 (no problems) to 4 (severe problems). The retest reliability and validity of the questionnaire were considered good (Horowitz, Strauss, Thomas, & Kordy, 2016). Cronbach's alpha ($\alpha = .83$, measured at the start of therapy; $N = 267$) showed good internal consistency in the total IIP score.

Beck Depression Inventory. The degree of depressive symptoms was assessed in the first session using the German version of the Beck Depression Inventory (Beck, Steer, & Brown, 1996; Beck, Ward, Mendelson, Mock, & Erbaugh, 1961). Two versions of the inventory were applied, BDI-I for CBT_{man} and PDT_{man}, and BDI-II for CBT_{nat}. Consequently, only items present in both inventories were analyzed. For pragmatic issues, a new 16-item scale was created that included items A-M, P, R, and U of the BDI-I, and items 1-10, 12, 13, 16, 18, 19, and 23 of the BDI-II. Summed scores were calculated when at least 80% of the items were answered. Cronbach's alpha values were .84 for the new BDI-I scale .85 for the new BDI-II scale, and .87 for the new combined scale. Correlations between each of the original scales (BDI-I and BDI-II) and the new scale were $r = .982$ and $r = .984$, respectively. Both the retest reliability ($>.60$) and criterion-based validity were good (Beck, Steer, & Carbin, 1988; Wang & Gorenstein, 2013).

Assessment of body movements. In the present study, we analyzed movements observed in the third session (probatoric session); this session was considered to represent the beginning of therapy, when a relationship between the patient and therapist is already

established. The same analyses were performed for each camera technique. First, all video recordings were converted by Any Video Converter 3.0 (Anvsoft Inc., 2009) into a coherent AVI format with a size of 640×480, a frame rate of 25 fps, and a bit rate of 2000 Kb/s. We used motion energy analysis (MEA) to assess body movements during the first 15 min of each session (starting after completing the questionnaires and clarifying administrative issues; the latest start time was 10 min after the session began). The MEA is an automated, objective frame-differencing technique. Body movement assessments with the MEA have been shown previously to be reliable (Altmann et al., 2017) and valid (Paxton & Dale, 2013b). We used the first 15 min because previous studies demonstrated that this interval is representative of the entire 50-min session (Paulick et al., 2017; Ramseyer & Tschacher, 2011). Assuming a stable background, motion energy values are calculated as the sum of the pixel changes between frames. These values can be interpreted to reflect movement intensity. The cut-off for a meaningful pixel change was determined empirically based on the analysis of 40 videos and set to the 99% quantile of the pixel intensity change (Altmann, 2013). We used a pixel change of 12 as the cut-off value. The value is in accordance with the range of 10 to 15 in the software from Ramseyer and Tschacher (2011). Movements can be evaluated for each region of interest (ROI), e.g., one person, separately (Altmann et al., 2018; Ramseyer & Tschacher, 2011). We used Altmann et al.'s (2013) algorithm (script publicly available at [github: https://github.com/10101-00001/MEA](https://github.com/10101-00001/MEA)). MEA scores obtained using the tools from Altmann (2013) and Ramseyer and Tschacher (2011) showed high correspondence ($r = .996$). We used the Matlab version of Altmann (2013) because the MEA could be performed within the Windows operating system and ROI sizes could be saved to process the time series further. We defined the ROI as the region that included the entire upper body and all potential upper-body movements. In addition, we introduced two ROIs (10×10 pixels) into the background to filter out video errors. We then divided the time series values by the corresponding ROI size to

convert the motion energy time series to intensity values from 0 to 100%. Subsequently, we labeled all frames as “missing data” when the background ROIs had pixel changes >5%. Based on all available time series (n = 891), we found that video errors could be detected by an increase in motion energy of approximately 15%, followed by a decrease of approximately 15%, within three frames. Therefore, these values were also labeled as missing data. Afterwards, we interpolated the missing values for sections with fewer than eight consecutive missing values using linear interpolation based on the neighboring values. Finally, a moving median of bandwidth 5 was applied.

In a validation study comparing human-rated synchronization intervals to intervals detected by the algorithm (Schoenherr, Paulick, Deisenhofer, et al., 2018), transforming the time series to logarithmic values resulted in more comparable peak heights for the therapists’ and patients’ time series, and synchrony could be detected more accurately. Therefore, we transformed motion energy values to logarithmic values.

Quantification of movement synchrony. First, we applied a windowed cross-lagged correlation (WCLC) with a window width of 5 s (125 frames) and maximum time lag of ± 5 s (125 frames) in steps of 0.04 s (1 frame) with overlapping windows. These settings were based on a previous validation study performed by Schoenherr, Paulick, Deisenhofer, et al. (2018). We controlled for pseudo-synchrony by applying a parametric test (Altmann, 2011, 2013); each cross-correlation was tested against zero before using it for the detection of synchronization intervals. Second, we used Altmann’s (2011, 2013) peak-picking algorithm to identify intervals exhibiting linear interdependencies. The algorithm is inspired by the peak-picking algorithm of Boker, Rotondo, Xu, and King (2002). In contrast to their algorithm, synchronization intervals were identified by connecting neighboring peaks with an equal time lag (1 frame tolerance) and reducing overlapping intervals based on the size of the cross-correlation function (Altmann, 2011). Moreover, all synchronization intervals that lasted less than 0.4 s were

excluded because the intervals are too short to comprise meaningful synchrony. The combination of WCLC and peak-picking algorithm has been demonstrated to be superior to Boker et al.'s (2002) algorithm (Altmann, 2011). By comparing synchronization intervals and non-synchronization intervals rated by humans to intervals rated by this algorithm, intervals with $R^2 < .25$ were excluded because this procedure has been shown to exclude randomly occurring synchronization intervals (Schoenherr, Paulick, Deisenhofer, et al., 2018). In contrast to the synchrony index of Ramseyer and Tschacher (2011), this method results in a list of intervals with meaningful synchrony, including the beginning, end, duration, time lag, and average R^2 ; the R^2 value reflects the strength of the association in each synchronization interval. Third, the ratio of synchronous time to total time, multiplied by 100, was used as a global value; in contrast to the averaged cross-correlation as a measure of average interrelatedness in Ramseyer and Tschacher (2011). Therefore, the global score (synchrony total) denoted the percentage of time that the patient and therapist exhibited synchronous behavior (either simultaneous or time-lagged synchrony). The literature suggests a negative association between synchrony and interpersonal problems (Ramseyer and Tschacher, 2011). Our global score was validated with the inventory of interpersonal problems (negative association) (Schoenherr, Paulick, Worrack, et al., 2018). The MATLAB scripts are publicly available (github: https://github.com/10101-00001/sync_ident). Applying the WCLC also offers the opportunity to differentiate between patient-induced synchrony and synchrony led by the therapist. Therefore, we also conducted analyses with these two variables.

Statistical Methods

Imputation of missing values. The rate of missing values in our data varied from 0% to 34.8% (mean: 18.7%, median: 25.1%). Demographic variables had little missing data (0% - 8%), questionnaire data had a modest amount of missing data (before treatment: ~4.6%, after treatment: 23.8%), and video data had the most missing data (~34.5%). Missing values were

due to incomplete questionnaires and non-existent or poor-quality video recordings (e.g., varying light conditions or a third person in the room). Little's MCAR test suggested that values were missing completely at random within and between variables ($\chi^2 = 539.210$, $df = 511$, $p = .187$). We used the R package (version 3.4.0; R Core Team, 2017) *missForest* (Stekhoven, 2011) to impute missing values with a maximum number of 20 iterations and 1000 random trees (seed = 47). The *missForest* package has been demonstrated to be superior to other imputation techniques (Stekhoven & Buhlmann, 2012). Missing values for movement synchrony were imputed with respect to sessions 3 and 8 only. Indices for fitting the imputations showed a good fit to the data (normalized root-mean-square error = .196; proportion of falsely classified = .048).⁴

Matching. We included different therapeutic approaches to investigate whether nonverbal synchrony is a common mechanism. Patients were not randomized to the naturalistic or manual-guided settings. Therefore, we matched dyads from the manual-guided therapeutic approaches (CBT_{man} and PDT_{man} conditions combined) to dyads from the naturalistic approach by propensity score matching. We used all available baseline variables for matching: patient age, therapist age, patient gender, therapist gender, high school diploma (dichotomous: yes=1, no=0), current relationship (dichotomous: yes=1, no=0), BDI sum score determined in session 1, and IIP sum score at session 1. We used the R package *MatchIt* to conduct the full matching (Ho, Imai, King, & Stuart, 2011). Manual-based therapeutic approaches were defined as the treatment group (weight = 1). The groups were well matched as demonstrated by all standard mean differences being <.25.

Data analysis. First, we described the subsamples using descriptive statistics. To select predictors, we conducted point-biserial correlations (for metric predictor variables) and Cramers' V (for nominal predictor variables). Associations between dropout (criterion) and

⁴ Results were consistent using the dataset with missing values.

demographic or clinical variables (predictor) were rather small and, for most of the variables, not significantly different from zero (gender: $V = .05$, age: $r = -.05$, current relationship: $V = .09$, high school degree or higher: $V = -.09$, BDI before treatment: $r = .08$, IIP before treatment: $r = -.04$, movement synchrony s8: $r = -.06$). Two variables had marginally significant associations with dropout (number of comorbid diagnoses: $V = .20$, $p = .03$, movement synchrony session 3: $r = -.18$, $p = .06$). To not overfit the models, we used these variables (number of comorbid diagnoses, movement synchrony in session 3) as predictors.

Next, we examined the prediction of premature termination using different outcomes (binary dropout variable versus differentiation in early and late dropout versus CSC). First, we fit a logistic regression with dropout as the criterion and the total amount of movement synchrony in session 3 and number of comorbid diagnoses as predictors. Second, we conducted two logistic regressions using the amount of synchrony led by the patient (therapist follower) and the amount of synchrony led by the therapist (patient follower). As the time point at which dropout occurred is important, we also conducted a multinomial regression with a variable containing information on whether the patient ended therapy regularly, prematurely early (before the 10th session), or prematurely late (after the 10th session, see Hynan, 1990) as a criterion. Third, in accordance with Swift, Callahan, and Levine (2009), we also investigated the CSC in interpersonal problems measured by the IIP as a criterion in logistic regression. We also investigated the combination of dropout (dichotomously) and CSC (dropout with improvement, dropout without improvement, regular end with improvement, and regular end without improvement) as a criterion. In all regressions, we included the propensity scores as regression weights and therapist effect as a random intercept.

By using logistic regression, only raw frequencies of completion and premature termination are considered. This method does not take into account the time until a patient drops out of therapy. In addition, data from psychotherapy research is often right-censored,

meaning we only observe a certain interval of a person's development. We know whether somebody dropped out of therapy before the end of our observation interval, but we do not know whether another patient would have also dropped out of therapy if we had observed him or her longer (Corning & Malofeeva, 2004). Both features are addressed by Cox regression (survival analysis). In our survival analysis, we evaluated therapist and site effects by specifying a propensity score-weighted, mixed effects Cox regression model without predictors, and with the "therapist ID/site ID" as a cluster variable (zero model).⁵ Next, we conducted a sophisticated model including all important possible influences: a time-dependent criterion (time until premature termination) and the predictors. For the survival analysis, we used the *survival* package (Therneau, 2015) in R. Lastly, we explored the association between dropout and movement synchrony in session 3 within subgroups. In these analyses, we tested subgroup associations for the three experimental conditions for same-sex versus mixed-sex dyads because previous studies suggested that the dyad type has an important influence (Grammer et al., 1998). Therefore, we used a mixed effects PS-weighted Cox regression with the predictor movement synchrony in session 3.

⁵ Note: in a Cox regression model, no residual variable exists. Thus, a calculation of the intra-class correlation (ICC) was not possible.

Results

Descriptive Statistics

Patient demographics, clinical variables, and pre-treatment severity are shown in Table 1. Most patients had a primary diagnosis of SAD. Patients with no primary diagnosis of SAD had SAD as a secondary diagnosis. Half of the sample was female and in a current relationship. The movement synchrony was 24% over the entire sample (completer group: 25%, dropout group: 22%) for session 3 (i.e., 24% of the time spent in synchrony). Average dropout times were 13.9 session for CBT_{man} and PDT_{man} overall, 31.6 for CBT_{nat} overall, 6.4 for early dropout from CBT_{man} and PDT_{man}, 19.9 for late dropout from CBT_{man} and PDT_{man}, and 31.6 for late dropout from CBT_{nat}. CBT_{nat} had no early dropouts.

Logistic Regression

The results of the logistic regression are presented in Table 2 and indicate that low movement synchrony at session 3 is associated with dropout. Furthermore, lower movement synchrony led by the patient (therapist follower) is a significant predictor of dropout. Results referring to the time point at which the patient dropped out of therapy show a significant relationship between movement synchrony at session 3 and early dropout (before 10th session) for patient-leading synchrony and therapist-leading synchrony. High movement synchrony at session 3 also has a positive impact on the probability of a CSC, especially when the therapist is leading. Analyses with respect to the combined measure of CSC and dropout showed that synchrony was especially relevant when the patient ended therapy prematurely without improvement (both therapist- and patient-leading synchrony). Note that the results of the group dropout with improvement are not trustworthy because the sample size was very low (N = 2). The number of comorbid diagnoses was a significant predictor of dropout only for patients ending therapy prematurely without improvement.

Survival analysis

In considering the nested structure of the data, we found no significant site effect ($\chi^2 = 0.06$, $df = 1$, $p = .81$), but a marginally significant therapist effect on dropout rate ($\chi^2 = 2.93$, $df = 1$, $p = .09$). Therefore, we included a random intercept in all analyses (random effect for the individuals within therapists). We used the same predictors as in the logistic regression. Results are given in Table 2. Movement synchrony at session 3 was a significant predictor of premature termination (hazard ratio: 0.95). Furthermore, movement synchrony at session 3 when led by the patient was a significant predictor (hazard ratio: 0.90). A patient in a dyad with 22% movement synchrony led by the patient is 0.90 times as likely to drop out as a patient in a dyad with 21% movement synchrony led by the patient. Therefore, a 1% increase in synchrony (i.e., being in synchrony for another 9 seconds) results in a 10% decrease in the dropout rate.

Subgroup analyses. We found that the variables experimental condition and dyad type had a marginally significant moderating effect on the association between movement synchrony during session 3 and the dropout rate (hazard ratio experimental condition: CBT_{nat}: HR = 0.97^{ns} [CI: 0.90 – 1.04], CBT_{man}: HR = 1.02^{ns} [CI: 0.92 – 1.12], PDT_{man}: HR = 0.91^{†6} [CI: 0.81 – 1.01], dyad type: mixed-sex dyad HR = 0.94^{*} [CI: 0.84 – 0.98], same-sex dyad: HR = 0.97^{ns} [CI: 0.91 – 1.03]). The association between movement synchrony at session 3 and premature termination was strongest among patients who were in therapy in a mixed-sex dyad (no gender matching) and psychodynamic therapy.

Discussion

Our main finding that a 1% increase in the total frequency of synchrony during the third probatoric session was associated with a 5% reduction in the dropout rate is consistent with the existing literature. In patients with heterogeneous disorders, Paulick et al. (2017) showed that

⁶ † indicates marginally significant with a 10% alpha level, ^{ns} indicates non-significance.

those who dropped out of therapy had a lower level of synchrony than those who did not improve but terminated therapy consensually. Referring to premature termination, phases in which the patient leads the synchrony and the therapist is following seem to be very important. This may reflect the therapists' responsiveness to the patient. As movements are often shown in times when one is speaking (e.g., emphasizing speech with gestures), this may also indicate that it is important to let the patient speak (and move) during this early phase of therapy; thus, therapists should only react to the behaviors shown and not dominate the interaction. This is in line with the content of these sessions because the patient speaks about his/her reasons for searching therapy and the indication for therapy is evaluated. However, the current data cannot address the reason for a link between synchrony and premature termination. Referring to the clinical outcome, therapist-leading synchrony seems to also have an important influence. Intervals when the therapist leads synchrony are valid predictors of the probability of a CSC in interpersonal problems. Therefore, the different leading/pacing sequences seem to have different effects on various outcomes, which needs further attention in clinical research.

In line with our results, other studies have shown that patients who drop out commonly report a low therapeutic alliance; specifically, they disliked the therapist and did not feel respected (Mohl, Martinez, Ticknor, Huang, & Cordell, 1991). Low synchrony may reflect a mismatch between patient and therapist, a lack of empathy, or low alliance. For patients with SAD, less movement synchrony may confirm the anxious fears of being negatively judged in the interaction; thus, premature termination may be viewed as avoidance behavior by patients with severe impairments in nonverbal synchrony (Heerey & Kring, 2007). The therapist him/herself seems to be related to dropout; thus, regarding movement synchrony, some therapists may be more rigid or responsive to bodily synchrony. This is in line with many studies showing therapist effects within psychotherapy (Zimmermann et al., 2017).

Our results suggest that movement synchrony may be especially relevant in very early phases of therapy. If movement synchrony is low, early dropout from therapy (before the 10th session) is likely. This is in line with previous studies showing that low therapeutic alliance during early phases of therapy leads to a higher dropout rate in psychotherapy and psychiatric outpatient interventions (Johansson & Eklund, 2006; Mohl et al., 1991). In addition, our results show that, with another definition of dropout, CSC in interpersonal problems, movement synchrony in the third session of therapy may influence the success of therapy.

Our further analyses identified potential moderating variables, including dyad type and therapeutic approach. We found that dyad type may moderate the association between nonverbal synchrony and dropout. When movement synchrony was low, the risk of premature termination was not as high in same-gender dyads as in mixed-gender dyads. This result suggests that matching the therapist and patient may facilitate a therapeutic relationship and prevent dropout. Consistent with this finding, Wintersteen, Mensinger, and Diamond (2005) showed that, in adolescents undergoing therapy due to substance abuse, gender-matched dyads had higher alliance ratings and lower therapy dropout rates than gender-mismatched dyads. However, it also may indicate that movement synchrony between therapist and patient may be an opportunity to compensate for different demographic characteristics. A potential moderation of premature therapy termination by the therapeutic approach may be explained by the differences in approaches. For example, CBT comprises goal-oriented and active intervention strategies, whereas PDT focuses mainly on the re-enactment and treatment of conflictual relationships (Luborsky, 1995). Thus, CBT is more structured than PDT. Moreover, PDT highlights the relational dimension; thus, apart from verbal content, it is particularly important *how* the therapist says something (Leichsenring, Hiller, Weissberg, & Leibing, 2006). Accordingly, movement synchrony between interaction partners is likely to be more important in PDT than in CBT. Therefore, it may be important to investigate the therapeutic approach

(treating conflictual relationships in PDT vs. cognitive restructuring in CBT) and which specific intervention was carried out during phases of synchrony. We think that the effect was not significant because of the low sample size in the smallest group (N=12).

In addition, the context of what is occurring in the session at the time of shifts in leading/following of synchrony may be important. However, the potential moderators require further investigation because the confidence intervals from the hazard ratios did not detect significant differences between the subgroups.

Strengths and Limitations

The main strength of this study was that it was the first study to investigate movement synchrony in a large sample (>250 dyads) of patients with a single disorder. We could rule out bias due to familiarity with the research subject because neither the therapists nor the patients were informed about the research topic. We also accounted for therapist effects by applying HLM and non-randomization of manual-based and naturalistic approaches by applying propensity scores. Due to our limited sample size (48 individuals in the smallest group), we avoided overfitting the model by not incorporating interaction terms. Furthermore, we used the MEA and WCLC for a highly objective, reliable assessment of movement synchrony. To define dropout, we used therapist ratings and CSC to ensure the meaningfulness of movement synchrony for the therapeutic process.

One limitation of our study was that the sample was not representative of the population of patients with SAD because not all patients had a primary diagnosis of SAD. Thus, the generalizability of this study is limited. The CBT_{nat} arm included patients with secondary SAD diagnoses to ensure comparable sample sizes among groups.

A potential limitation was the MEA, which evaluated aggregated motions of different body parts. Therefore, we could not differentiate between head nods and gestures. In addition, opposite movements may also lead to high movement synchrony. Future studies should

investigate the impact of motions by different body parts. Other modalities (prosody, verbal content, etc.) should also be evaluated to obtain an integrated view of the therapeutic process. From a methodological point of view, our chosen time lag of ± 5 s may have introduced bias (recommended time lags range from 1.5 s to 7 s) (Altmann, 2011; Louwerse, Dale, Bard, & Jeuniaux, 2012; Ramseyer & Tschacher, 2011; Robinson, Herman, & Kaplan, 1982).

Implications for Future Research

Future studies should investigate empirically whether and how synchrony can be manipulated in a psychotherapeutic context. It would also be interesting to determine whether the ability to synchronize is a personality trait or a dyadic interaction pattern. Moreover, for a comprehensive view of the therapy process, other nonverbal modalities (e.g., eye contact, prosody) should be examined. Future studies should focus on synchrony in different therapeutic contexts. To date, studies have focused on synchrony in disorder-heterogeneous samples that received CBT. However, different disorders have shown differences in synchrony (Paulick et al., 2018); therefore, it may be useful to investigate samples with homogeneous disorders. The In-Sync model of psychotherapy should be extended to include the differential effects of movement synchrony in various modalities, disorders, therapeutic approaches, and stages of therapy.

Implications for Clinical Practice

Nonverbal elements appear to have an influence on the therapeutic process. Therefore, they should be explicitly integrated into psychotherapy. A “common ground” between the therapist and patient (e.g., matching gender) may positively influence the dropout rate. However, some studies have shown that matching the patient and therapist by demographic variables is inappropriate (Sterling, Gottheil, Weinstein, & Serota, 2001). Nonverbal communication in therapy could compensate for mismatched demographic or cognitive traits. However, other studies have demonstrated that a conscious imitation of patient actions could

be counterproductive because it may elicit negative feelings (White & Argo, 2011); thus, the positive effects of imitation only occur when a behavior is perceived as authentic (Stel & Vonk, 2010). Preferably, therapists should be trained in how their own and their patients' nonverbal communications are perceived. This training could create an awareness of unconscious nonverbal behavior, which could guide them in addressing non-synchronous nonverbal behavior explicitly.

References

- Altmann. (2011). Investigation of movement synchrony using windowed crosslagged regression. *Analysis of Verbal and Nonverbal Communication and Enactment: The Processing Issue*, 344-354. doi:10.1007/978-3-642-25775-9_31
- Altmann. (2013). *Synchronisation nonverbaler Verhaltens: Weiterentwicklung und Anwendung zeitreihenanalytischer Identifikationsverfahren. [Synchronization of nonverbal behavior: Development and Application of time series analysis methods]*: Springer-Verlag.
- Altmann, Schoenherr, Paulick, Knitter, Worrack, Schiefele, . . . Strauss. (2018). Introduction, Practical Guide, and Validation Study for Measuring Body Movements Using Motion Energy Analysis. *submitted*.
- Altmann, Wittmann, Steffanowski, von Heymann, Bruckmayer, Pfaffinger, . . . Fembacher. (2014). Abbrüche antragspflichtiger ambulanter Psychotherapien: Häufigkeit, Risikofaktoren, Outcome [Non-Completion of Insurance-Funded Outpatient Psychotherapy: Frequency, Risk Factors, Outcome]. *Psychiatrische Praxis*, 41(06), 305-312. doi:10.1055/s-0034-1370172
- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders (5th ed.)*. Washington, DC: Author.
- Anvsoft Inc. (2009). Any Video Converter.
- Beck, Steer, & Brown. (1996). Beck depression inventory-II. *San Antonio*, 78(2), 490-498.
- Beck, Steer, & Carbin. (1988). Psychometric properties of the Beck Depression Inventory: Twenty-five years of evaluation. *Clinical Psychology Review*, 8(1), 77-100. doi:10.1016/0272-7358(88)90050-5
- Beck, Ward, Mendelson, Mock, & Erbaugh. (1961). An inventory for measuring depression. *Archives of general psychiatry*, 4(6), 561-571. doi:10.1001/archpsyc.1961.01710120031004

- Bernieri, & Rosenthal. (1991). Interpersonal coordination: Behavior matching and interactional synchrony. In R. Feldman & B. Rimé (Eds.), *Fundamentals of nonverbal behavior* (Vol. xiv, pp. 511 pp.). New York, NY, US: Cambridge University Press; Paris, France: Editions de la Maison des Sciences de l'Homme.
- Boker, Rotondo, Xu, & King. (2002). Windowed cross-correlation and peak picking for the analysis of variability in the association between behavioral time series. *Psychological Methods*, 7(3), 338. doi:10.1037/1082-989X.7.3.338
- Brown, Barlow, & DiNardo. (1994). *Anxiety Disorders Interview Schedule for DSM-IV (ADIS-IV): Client Interview Schedule*: Graywind Publications Incorporated.
- Clark, & Wells. (1995). A cognitive model of social phobia. *Social phobia: Diagnosis, assessment, and treatment*, 41(68), 00022-00023.
- Corning, & Malofeeva. (2004). The Application of Survival Analysis to the Study of Psychotherapy Termination. *Journal of Counseling Psychology*, 51(3), 354.
- Dijk, Fischer, Morina, van Eeuwijk, & van Kleef. (2018). Effects of Social Anxiety on Emotional Mimicry and Contagion: Feeling Negative, but Smiling Politely. *Journal of Nonverbal Behavior*, 42(1), 81-99.
- Eskildsen, Hougaard, & Rosenberg. (2010). Pre-treatment patient variables as predictors of drop-out and treatment outcome in cognitive behavioural therapy for social phobia: A systematic review. *Nordic journal of psychiatry*, 64(2), 94-105. doi:10.3109/08039480903426929
- Farabee, Ramsey, & Cole. (1993). Social anxiety and speaker gaze in a persuasive atmosphere. *Journal of Research in Personality*, 27(4), 365-376.
- Galbusera, Finn, & Fuchs. (2016). Interactional synchrony and negative symptoms: An outcome study of body-oriented psychotherapy for schizophrenia. *Psychother Res*, 1-13. doi:10.1080/10503307.2016.1216624

-
- Grammer, Kruck, & Magnusson. (1998). The courtship dance: Patterns of nonverbal synchronization in opposite-sex encounters. *Journal of Nonverbal Behavior*, 22(1), 3-29. doi:10.1023/A:1022986608835
- Hans, & Hiller. (2013). A meta-analysis of nonrandomized effectiveness studies on outpatient cognitive behavioral therapy for adult anxiety disorders. *Clinical Psychology Review*, 33(8), 954-964. doi:10.1037/a0025749
- Heerey, & Kring. (2007). Interpersonal consequences of social anxiety. *Journal of abnormal psychology*, 116(1), 125.
- Ho, Imai, King, & Stuart. (2011). MatchIt: nonparametric preprocessing for parametric causal inference. *Journal of Statistical Software*, 42(8), 1-28. doi:10.18637/jss.v042.i08
- Horowitz, Strauss, Thomas, & Kordy. (2016). *IIP-D Inventar zur Erfassung Interpersonaler Probleme - Deutsche Version. [Inventory for the assessment of interpersonal problems - German version]* (Vol. 3. überarbeitete Auflage [3rd edition]): Hogrefe.
- Hoyer, Wiltink, Hiller, Miller, Salzer, Sarnowsky, . . . Leibing. (2014). Baseline patient characteristics predicting outcome and attrition in cognitive therapy for social phobia: Results from a large multicentre trial. *Clinical Psychology & Psychotherapy*.
- Hynan. (1990). Client reasons and experiences in treatment that influence termination of psychotherapy. *Journal of Clinical Psychology*, 46(6), 891-895.
- Iacoboni. (2009). Imitation, empathy, and mirror neurons. *Annual review of psychology*, 60, 653-670. doi:10.1146/annurev.psych.60.110707.163604
- Johansson, & Eklund. (2006). Helping alliance and early dropout from psychiatric out-patient care. *Social psychiatry and psychiatric epidemiology*, 41(2), 140-147.
- Kessler. (2003). The impairments caused by social phobia in the general population: implications for intervention. *Acta Psychiatrica Scandinavica*, 108(s417), 19-27.
- Koole, & Tschacher. (2016). Synchrony in Psychotherapy: A Review and an Integrative Framework for the Therapeutic Alliance. *Frontiers in Psychology*, 7(862). doi:10.3389/fpsyg.2016.00862

- Lakin, Jefferis, Cheng, & Chartrand. (2003). The chameleon effect as social glue: Evidence for the evolutionary significance of nonconscious mimicry. *Journal of Nonverbal Behavior*, 27(3), 145-162.
- Leichsenring, Beutel, & Leibing. (2007). Psychodynamic psychotherapy for social phobia: a treatment manual based on supportive–expressive therapy. *Bulletin of the Menninger Clinic*, 71(1), 56-84. doi:10.1521/bumc.2007.71.1.56
- Leichsenring, Hiller, Weissberg, & Leibing. (2006). Cognitive-behavioral therapy and psychodynamic psychotherapy: Techniques, efficacy, and indications. *American Journal of Psychotherapy*, 60(3), 233.
- Leichsenring, Salzer, Beutel, Herpertz, Hiller, Hoyer, . . . Poehlmann. (2013). Psychodynamic therapy and cognitive-behavioral therapy in social anxiety disorder: a multicenter randomized controlled trial. *American Journal of Psychiatry*. doi:10.1176/appi.ajp.2013.12081125
- Leichsenring, Salzer, Beutel, von Consbruch, Herpertz, Hiller, . . . Joraschky. (2009). SOPHONET–Forschungsverbund zur Psychotherapie der sozialen Phobie. *PPmP–Psychotherapie· Psychosomatik· Medizinische Psychologie*, 59(03/04), 117-123.
- Louwerse, Dale, Bard, & Jeuniaux. (2012). Behavior matching in multimodal communication is synchronized. *Cognitive science*, 36(8), 1404-1426. doi:10.1111/j.1551-6709.2012.01269.x
- Luborsky. (1995). Einführung in die analytische Psychotherapie: ein Lehrbuch. 2. Aufl. Göttingen, Zürich: Vandenhoeck und Ruprecht.
- Lutz, Schiefele, Wucherpfennig, Rubel, & Stulz. (2016). Clinical effectiveness of cognitive behavioral therapy for depression in routine care: A propensity score based comparison between randomized controlled trials and clinical practice. *Journal of Affective Disorders*, 189, 150-158. doi:10.1016/j.jad.2015.08.072
- Marsh, Richardson, & Schmidt. (2009). Social connection through joint action and interpersonal coordination. *Topics in Cognitive Science*, 1(2), 320-339.

-
- Mennin, Fresco, Heimberg, Schneier, Davies, & Liebowitz. (2002). Screening for social anxiety disorder in the clinical setting: using the Liebowitz Social Anxiety Scale. *Journal of anxiety disorders, 16*(6), 661-673. doi:10.1016/S0887-6185(02)00134-2
- Mogan, Fischer, & Bulbulia. (2017). To be in synchrony or not? A meta-analysis of synchrony's effects on behavior, perception, cognition and affect. *Journal of Experimental Social Psychology, 72*, 13-20. doi:10.1016/j.jesp.2017.03.009
- Mohl, Martinez, Ticknor, Huang, & Cordell. (1991). Early dropouts from psychotherapy. *Journal of Nervous and Mental Disease.*
- Paulick, Deisenhofer, Ramseyer, Tschacher, Boyle, Rubel, & Lutz. (2017). Nonverbal Synchrony: A new approach to better understand psychotherapeutic processes and drop-out. *Journal of Psychotherapy Integration.* doi:10.1037/int0000099
- Paulick, Rubel, Deisenhofer, Schwartz, Thielemann, Altmann, . . . Lutz. (2018). Diagnostic features of nonverbal synchrony in psychotherapy: Comparing depression and anxiety. *Cognitive Therapy and Research*, Advanced online publication. doi:10.1007/s10608-018-9914-9
- Paxton, & Dale. (2013a). Argument disrupts interpersonal synchrony. *The Quarterly Journal of Experimental Psychology, 66*(11), 2092-2102. doi:10.1080/17470218.2013.853089
- Paxton, & Dale. (2013b). Frame-differencing methods for measuring bodily synchrony in conversation. *Behavior Research Methods, 45*(2), 329-343. doi: 10.3758/s13428-012-0249-2
- R Core Team. (2017). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from <https://www.R-project.org/>
- Ramseyer, & Tschacher. (2011). Nonverbal Synchrony in Psychotherapy: Coordinated Body Movement Reflects Relationship Quality and Outcome. *Journal of Consulting and Clinical Psychology, 79*(3), 284-295. doi:10.1037/a0023419

- Rasting, & Beutel. (2005). Dyadic affective interactive patterns in the intake interview as a predictor of outcome. *Psychotherapy Research*, 15(3), 188-193. doi:10.1080/10503300512331335039
- Robinson, Herman, & Kaplan. (1982). Autonomic responses correlate with counselor–client empathy. *Journal of Counseling Psychology*, 29(2), 195. doi:10.1037/0022-0167.29.2.195
- Ruscio, Brown, Chiu, Sareen, Stein, & Kessler. (2008). Social fears and social phobia in the USA: results from the National Comorbidity Survey Replication. *Psychological Medicine*, 38(1), 15-28.
- Saxon, Barkham, Foster, & Parry. (2017). The contribution of therapist effects to patient dropout and deterioration in the psychological therapies. *Clinical Psychology & Psychotherapy*, 24(3), 575-588. doi:10.1002/cpp.2028
- Schoenherr, Paulick, Deisenhofer, Schwartz, Rubel, Lutz, . . . Altmann. (2018). Identification of movement synchrony: Validation of time series analysis methods Manuscript submitted.
- Schoenherr, Paulick, Worrack, Strauss, Rubel, Schwartz, . . . Altmann. (2018). Quantification of nonverbal synchrony using linear time series analysis methods: Lack of convergent validity and evidence for facets of synchrony. *Behavior Research Methods*, 1-23. doi:10.3758/s13428-018-1139-z
- Stangier, Clark, & Ehlers. (2006). Soziale Phobie: Fortschritte der Psychotherapie. *Göttingen, Hogrefe-Verlag*.
- Stekhoven. (2011). Using the missForest Package. *R package*, 1-11.
- Stekhoven, & Buhlmann. (2012). MissForest--non-parametric missing value imputation for mixed-type data. *Bioinformatics*, 28(1), 112-118. doi:10.1093/bioinformatics/btr597
- Stel, & Vonk. (2010). Mimicry in social interaction: Benefits for mimickers, mimicees, and their interaction. *British Journal of Psychology*, 101(2), 311-323. doi:10.1348/000712609X465424

-
- Sterling, Gottheil, Weinstein, & Serota. (2001). The effect of therapist/patient race-and sex-matching in individual treatment. *Addiction*, *96*(7), 1015-1022. doi:10.1046/j.1360-0443.2001.967101511.x
- Swift, Callahan, & Levine. (2009). Using clinically significant change to identify premature termination. *Psychotherapy: Theory, Research, Practice, Training*, *46*(3), 328.
- Swift, & Greenberg. (2012). Premature discontinuation in adult psychotherapy: A meta-analysis. *Journal of Consulting and Clinical Psychology*, *80*(4), 547-559. doi:10.1037/a0028226
- Therneau. (2015). A Package for Survival Analysis in S. R package version 2.37-7. 2014. URL <http://CRAN.R-project.org/package=survival>.
- Thomas, Brähler, & Strauß. (2011). IIP-32: Entwicklung, Validierung und Normierung einer Kurzform des Inventars zur Erfassung interpersonaler Probleme. *Diagnostica*. doi:10.1026/0012-1924/a000034
- Vicaria, & Dickens. (2016). Meta-analyses of the intra-and interpersonal outcomes of interpersonal coordination. *Journal of Nonverbal Behavior*, *40*(4), 335-361. doi:10.1007/s10919-016-0238-8
- Voncken, & Bögels. (2008). Social performance deficits in social anxiety disorder: Reality during conversation and biased perception during speech. *Journal of anxiety disorders*, *22*(8), 1384-1392.
- Vrijzen, Lange, Becker, & Rinck. (2010). Socially anxious individuals lack unintentional mimicry. *Behaviour Research and Therapy*, *48*(6), 561-564.
- Vrijzen, Lange, Dotsch, Wigboldus, & Rinck. (2010). How do socially anxious women evaluate mimicry? A virtual reality study. *Cognition and Emotion*, *24*(5), 840-847.
- Wang, & Gorenstein. (2013). Psychometric properties of the Beck Depression Inventory-II: a comprehensive review. *Revista Brasileira de Psiquiatria*, *35*(4), 416-431. doi:10.1590/1516-4446-2012-1048

- White, & Argo. (2011). When imitation doesn't flatter: The role of consumer distinctiveness in responses to mimicry. *Journal of Consumer Research*, 38(4), 667-680.
- Wintersteen, Mensinger, & Diamond. (2005). Do gender and racial differences between patient and therapist affect therapeutic alliance and treatment retention in adolescents? *Professional Psychology: Research and Practice*, 36(4), 400. doi:10.1037/0735-7028.36.4.400
- Wittchen, Wunderlich, Gruschwitz, & Zaudig. (1997). Skid-i. *Strukturiertes klinisches Interview für DSM-IV*.
- Zimmermann, Rubel, Page, & Lutz. (2017). Therapist Effects on and Predictors of Non-Consensual Dropout in Psychotherapy. *Clinical Psychology & Psychotherapy*, 24(2), 312-321. doi:10.1002/cpp.2022

Table 1

Descriptive statistics of the sample after imputation of missing data

| | Full sample (N=267) | Completers (N=219) | Dropouts (N=48) | Comparison completer vs. dropout |
|---|------------------------|-----------------------|--------------------|-------------------------------------|
| Gender female | 54% | 53% | 58% | $\chi^2_{df=1} = .54, p = .46$ |
| Age, years | 34.38 (11.77) | 34.64 (11.90) | 33.19 (11.22) | $F_{1,265} = .60, p = .44$ |
| Current relationship | 51% | 48% | 60% | $\chi^2_{df=1} = 2.27, p = .13$ |
| High school diploma or higher | 59% | 61% | 50% | $\chi^2_{df=1} = 2.04, p = .15$ |
| CBT manual-guided | 38% | 40% | 50% | $\chi^2_{df=1} = 4.32, p = .04$ |
| PDT manual-guided | 20% | 41% | 25% | $\chi^2_{df=1} = .83, p = .36$ |
| CBT naturalistic | 42% | 19% | 25% | $\chi^2_{df=1} = 1.65, p = .20$ |
| BDI T0 mean (sd) | 13.67 (8.04) | 13.36 (7.96) | 15.10 (8.35) | $F_{1,265} = 1.84, p = .18$ |
| IIP T0 mean (sd) | 1.80 (0.43) | 1.81 (0.44) | 1.76 (0.39) | $F_{1,265} = .46, p = .50$ |
| Number of comorbid diagnoses | 1.27 (1.03) | 1.21 (0.97) | 1.56 (1.22) | $F_{1,265} = 4.83, p = .03$ |
| Social anxiety disorder (1 st diagnosis) | 76% | 78% | 71% | $\chi^2_{df=1} = 1.16, p = .28$ |
| Affective disorder (2 nd diag) | 32% | 33% | 27% | $\chi^2_{df=1} = 1.03, p = .31$ |
| Other disorder (2 nd diag) | 36% | 34% | 47% | $\chi^2_{df=1} = 3.03, p = .08$ |
| none | 32% | 33% | 26% | $\chi^2_{df=1} = 1.03, p = .31$ |
| Affective disorder (1 st diag)* | 17% | 16% | 19% | $\chi^2_{df=1} = .15, p = .70$ |
| Other anxiety disorder (1 st diag) * | 5% | 4% | 8% | $\chi^2_{df=1} = 1.52, p = .22$ |
| Other disorders (1 st diag)* | 1% | 2% | 2% | $\chi^2_{df=1} = .14, p = .71$ |

Note. CBT = cognitive behavioral therapy; PDT = psychodynamic-orientated therapy, BDI = Beck Depression Inventory, IIP = Inventory of Interpersonal Problems, diag = diagnosis; * with social anxiety disorder as the 2nd diagnosis, T0 = first measurement time point, session 1.

Table 2

Results (coefficients, p-values) of logistic/multinomial regression and survival analyses

| Predictor | Model 1 with synchrony total as predictor | | | Model 2 with synchrony led by patient as predictor | | | Model 3 with synchrony led by therapist as predictor | | | | | |
|--|---|-----------|----------|--|----------|-----------|--|-----------|----------|-----------|----------|-----------|
| | <i>b</i> | <i>SE</i> | <i>p</i> | <i>OR</i> | <i>b</i> | <i>SE</i> | <i>p</i> | <i>OR</i> | <i>b</i> | <i>SE</i> | <i>p</i> | <i>OR</i> |
| Logistic regression with dropout as a dichotomous criterion | | | | | | | | | | | | |
| Synchrony | -0.06* | 0.03 | .02 | 0.94 | -0.13* | 0.05 | .01 | 0.88 | -0.08 | 0.05 | .10 | 0.92 |
| N_diag | 0.25 | 0.23 | .27 | 1.28 | 0.27 | 0.23 | .23 | 1.31 | 0.23 | 0.22 | .31 | 1.26 |
| Multinomial regression with categorical dropout (regular end, early dropout, late dropout) as criterion, reference group : regular end | | | | | | | | | | | | |
| Synchrony on early dropout | -0.22* | 0.06 | <.001 | 0.80 | -0.46* | 0.13 | <.001 | 0.63 | -0.36* | 0.11 | .001 | 0.70 |
| Synchrony on late dropout | 0.04 | 0.04 | .33 | 1.04 | 0.07 | 0.08 | .36 | 1.07 | 0.08 | 0.08 | .28 | 1.08 |
| N_diag on early dropout | 0.58 | 0.40 | .14 | 1.79 | 0.62 | 0.41 | .12 | 1.86 | 0.60 | 0.40 | .13 | 1.82 |

| | | | | | | | | | | | | |
|---|--------|------|-------|------|--------|------|-----|------|--------|------|-------|------|
| N_diag on late dropout | -0.14 | 0.33 | .67 | 0.87 | -0.15 | 0.34 | .66 | 0.86 | -0.13 | 0.33 | .70 | 0.88 |
| Logistic regression with clinically significant change as criterion | | | | | | | | | | | | |
| Synchrony | 0.08 | 0.05 | .10 | 1.08 | 0.07 | 0.09 | .45 | 1.07 | 0.22* | 0.09 | .02 | 1.25 |
| N_diag | -0.17 | 0.38 | .65 | 0.84 | -0.13 | 0.38 | .71 | 0.88 | -0.18 | 0.38 | .64 | 0.84 |
| Multinomial regression with clinically significant change x dropout categories as criterion (regular end + improvement, regular end without improvement, dropout + improvement, dropout without improvement), reference group: regular end + improvement | | | | | | | | | | | | |
| Synchrony on reg_nonimprov | -0.07 | 0.05 | .16 | 0.93 | -0.07 | 0.10 | .53 | 0.93 | -0.20* | 0.10 | .05 | 0.82 |
| Synchrony on drop_improv | -0.08 | 0.50 | .87 | 0.92 | -0.21 | 0.90 | .82 | 0.81 | -0.08 | 1.18 | .95 | 0.92 |
| Synchrony on drop_nonimprov | -0.20* | 0.00 | <.001 | 0.82 | -0.31* | 0.12 | .01 | 0.73 | -0.41* | 0.15 | <.001 | 0.66 |
| N_diag on reg_nonimprov | 0.41 | 0.46 | .38 | 1.51 | 0.37 | 0.47 | .43 | 1.45 | 0.41 | 0.46 | .38 | 1.51 |
| N_diag on drop_improv | 1.40 | 5.91 | .81 | 4.06 | 1.81 | 7.28 | .80 | 6.11 | 1.13 | 4.91 | .82 | 3.10 |
| N_diag on drop_nonimprov | 0.53* | 0.00 | <.001 | 1.70 | 0.48 | 0.59 | .42 | 1.62 | 0.53 | 0.66 | .41 | 1.70 |

| Survival analysis | | | | | | | | | | | | |
|-------------------|--------|------|-----|------|--------|------|-----|------|-------|------|-----|------|
| Synchrony | -0.05* | 0.02 | .04 | 0.95 | -0.11* | 0.90 | .02 | 0.90 | -0.07 | 0.05 | .13 | 0.93 |
| N_diag | 0.26 | 0.20 | .20 | 1.30 | 0.28 | 1.32 | .17 | 1.32 | 0.25 | 0.20 | .22 | 1.28 |

Note. synchrony = movement synchrony session 3, N_diag = number of diagnoses, reg_nonimprov = regular end and no improvement, drop_improv = dropout with improvement, drop_nonimprov = dropout without improvement, *b* = unstandardized coefficients, * indicates significance using an alpha level of 5%, OR= odds ratio, OR regarding survival analysis = hazard ratio.

Publikation 4

Movement synchrony and attachment related anxiety and avoidance in social anxiety disorder

Désirée Schoenherr¹, Bernhard Strauss¹, Jane Paulick², Anne-Katharina Deisenhofer², Brian Schwartz², Julian A. Rubel², Kaitlyn Boyle², Wolfgang Lutz², Ulrich Stangier³, Uwe Altmann¹

¹University Hospital Jena, Institute of Psychosocial Medicine and Psychotherapy, Jena, Germany

²Trier University, Department of Clinical Psychology and Psychotherapy, Trier, Germany

³Goethe University Frankfurt/Main, Department of Clinical Psychology and Psychotherapy, Frankfurt/Main, Germany

Acknowledgment

The current study was funded by the German Research Foundation (funding number GZ: STR 306/28-1 and LU 660/8-1). The analyzed videos were partly recorded in the SOPHO-NET project which was funded by the German Federal Ministry of Education and Research (BMBF, FKZ 01GV0607). We cordially thank U. Willutzki, St. Herpertz, J. Hoyer, P. Joraschky, W. Hiller, E. Leibing, M. Beutel and F. Leichsenring for providing video recordings of the SOPHO-NET project. We thank our student assistants Lukas A. Knitter and Alisa Shugaley for proof reading and help with data analysis.

Abstract

Attachment influences the individual's ability of self-regulation within interpersonal interactions, making it an important variable in psychotherapy. Movement synchrony as the association of people's movements is related to interpersonal relationships such as the therapeutic alliance and interpersonal problems. This study examines the association between attachment, social anxiety/avoidance and movement synchrony. One-hundred patients with social anxiety disorder were examined, of which $n = 57$ were treated with cognitive-behavioral therapy (CBT) and $n = 43$ with psychodynamic-oriented therapy. Symptoms were assessed using the Liebowitz Social Anxiety Scale, attachment-related anxiety and avoidance using the Experience in Close Relationships Questionnaires (ECR-R). Movements were extracted from videos of the 8th session by motion energy analysis. The frequency and time lag of synchrony were determined by time series analysis methods. Statistical analyses were performed by correlation analyses and mixed linear models. Results show that initial attachment anxiety and avoidance were not related to synchrony. However, synchrony was able to predict the extent of attachment anxiety at the end of therapy negatively. The episodes in which the therapist initiated synchrony were particularly predictive, especially in CBT. Higher attachment avoidance was associated with a larger time lag in establishing synchrony. Our results show that frequent movement synchrony has a positive effect on the attachment anxiety. It is beneficial if the therapist leads synchrony since the patient might have the feeling of the therapist being a secure base. Since there was no connection between anxiety symptoms and synchrony, synchrony seem to be rather a proximal variable to interpersonal constructs.

Keywords: attachment, nonverbal synchrony, motion energy analysis, windowed cross-lagged correlation

Public Significance Statement: The present study examined the associations of different indices of movement synchrony with anxiety symptoms and attachment characteristics using computer vision technology that makes it possible to investigate psychotherapy on a micro-process-level. Results show that movement synchrony led by the therapist in the first third of psychotherapy is a valid predictor of patient attachment anxiety at the end of therapy in a sample of patients with social anxiety disorder. Higher time lags of movement synchrony were found in patients with higher attachment avoidance. However, movement synchrony was not associated with social anxiety symptoms, possibly indicating that the dyadic variable is more closely related to interpersonal variables (attachment, interpersonal problems).

Introduction

Infants are born with behaviors designed to establish contact with their caregivers. Early experiences are then internalized as an inner working model, which partly shapes future relationships (Bowlby, 1973). An individual's pattern of attachment determines self-regulation in relation to the external environment and in interaction with other persons (Dales & Jerry, 2008). Individuals can be characterized by the two dimensions attachment-related anxiety and avoidance. An anxiously attached individual has a strong desire for close relationships, but a fear that the partner is unavailable. Often excessive reassurance seeking is present. Attachment-related anxiety is related to hyperactivating emotion regulation strategies that increase the monitoring of cues, which are related to threats and attachment representations (Mikulincer, Gillath, & Shaver, 2002). Thus, persons with high attachment-related anxiety react with higher attachment-related thoughts and worries even to neutral situations (Mikulincer, Birnbaum, Woddis, & Nachmias, 2000). Individuals with high scores of avoidance, on the other hand, strive to keep an emotional distance and often deny the need for close relationships (Cameron, Finnegan, & Morry, 2012). Attachment avoidance is linked with deactivation emotion regulation strategies meaning, for example, the inhibition of proximity seeking by down-regulating the attachment-related needs. Mikulincer et al. (2002) showed that the accessibility of attachment system is similar to anxiously attached individuals in threat situations. The difference is that attachment-related avoidance leads to a repression of the proximity seeking by dissociating preconscious activation of attachment-representations from conscious thoughts and behaviors (Mikulincer et al., 2002).

Studies have shown that, among many other variables, an insecure attachment style is associated with an increased likelihood of depression and hopelessness, whereas a secure attachment style is related to an increased emotional intelligence (Khledian, Garosi, Khairkhah, & Ghalandari, 2013). Patients with secure attachment style seem to improve better during

psychotherapy (Levy, Kivity, Johnson, & Gooch, 2018) and have better therapeutic alliances (Bernecker, Levy, & Ellison, 2014). Insecure attachment styles might work as a moderator of therapeutic outcome (Strauss et al., 2017).

References to an individual's attachment style can be found in their nonverbal behavior. This is plausible, since Ainsworth's Strange Situation test is also strongly based on the evaluation of nonverbal behaviors (approximation, gestures, facial expressions, movements) (Ainsworth, Blehar, Waters, & Wall, 1978). The type of movement, for example, is associated with attachment anxiety and attachment avoidance (Feniger-Schaal & Lotan, 2017). Furthermore, individuals with a secure attachment style and low values of attachment avoidance display more explorative and free movement than insecure individuals (Feniger-Schaal, Hart, Lotan, Koren-Karie, & Noy, 2018). Since attachment is a construct closely related to interpersonal relationships, a variable that is even more important might be movement synchrony related to another person.

Movement synchrony as one form of nonverbal synchrony is the coordination of movement dynamics of two interaction partners. For example, when the patient underlines his/her speech with gestures and the therapist subsequently changes his/her sitting posture, these movements are considered as movement synchronous. During the last years, research has mainly focused on the strength of nonverbal synchrony (Galbusera, Finn, & Fuchs, 2016; Paulick et al., 2017; Paulick et al., 2018; Ramseyer & Tschacher, 2011). The use of local time series analysis methods such as windowed cross-lagged correlation (WCLC) makes it possible to investigate also the frequency of synchrony as well as leading changes within synchrony (Altmann et al., submitted; Ramseyer & Tschacher, 2011). Therefore, it is possible to separate the frequency of synchrony that was initiated by the patient (patient moves first, therapist follows) and the frequency of synchrony that is initiated by the therapist (therapist moves first, patient follows). Moreover, synchrony can be evaluated regarding the amount of time needed

until synchrony is established. If a person moves, it can take 1 second or 5 seconds for the interaction partner to engage in the behavior. The delay between the initial and echoed behavior is characterized by the time lag. The average (Feldman, Greenbaum, & Yirmiya, 1999; Kleinbub, 2016) and the range (Tschacher, Ramseyer, & Koole, 2018) of the different time lags can be used for a more detailed investigation of the nonverbal interaction. Although the time lag of synchrony may provide important information, it has rarely been investigated so far. Apart from studies investigating physiological synchrony (Kleinbub, 2017), only two studies suggest that a short time lag is positive in the sense of being associated with lower depression (Zlochower & Cohn, 1996) and greater self-control (Feldman et al., 1999). In contrast, the range of the time lag seems to be influenced by attachment characteristics: Tschacher, Ramseyer, and Koole (2017) showed that a higher time lag range was associated with greater attachment avoidance in healthy participants measured with the Measure of Attachment Qualities (MAQ).

Linking nonverbal synchrony with attachment, several studies have shown that higher levels of synchrony are associated with a secure attachment style in childhood (e.g., Beebe et al., 2012; Harrist & Waugh, 2002; Leclère et al., 2014; Lindsey & Caldera, 2015). Moreover, a high degree of synchrony between caregiver and child has been shown to have a positive influence on regulatory capacity, empathy, and general development (Feldman, 2012; Feldman, Magori-Cohen, Galili, Singer, & Louzoun, 2011).

Beyond the positive influences of nonverbal synchrony in infancy and childhood, it can also be assumed that synchrony has a positive effect on the patient's attachment pattern within psychotherapy. A positive experience with the therapist may therefore influence a patient's insecure attachment pattern (Strauss et al., 2018; Taylor, Rietzschel, Danquah, & Berry, 2015). The In-Sync model of psychotherapy suggests that nonverbal synchrony modulates emotion regulation abilities and thus also attachment characteristics of patients by forming a connection

between therapist and patient that is established by mutual coordination on a perceptual-motor level (Koole & Tschacher, 2016). Supporting evidence for the connection between higher synchrony and secure attachment in adults has been shown for movement synchrony (Ramseyer & Tschacher, 2011), affect attunement (Håvås, Svartberg, & Ulvenes, 2015), and face-rubbing mimicry (Hall, Millings, & Bouças, 2012). Palmieri et al. (2018) revealed that after an attachment prime, the therapists synchronized much more frequently on a physiological level; thus the therapist conveys the feeling of a secure base.

Besides the association between adult attachment and nonverbal interaction, there is a relation between attachment and the prevalence of anxiety symptoms (Colonnesi et al., 2011). For social anxiety disorder (SAD), which is symptomatically characterized by anxiety and avoidance of social situations (American Psychiatric Association, 2013), attachment style might be an important factor influencing process and outcome of psychotherapy. Patients with SAD are impaired in interpersonal actions (Heerey & Kring, 2007), which can also be traced back to deficits in the regulation of the self in relation to others. Patients with SAD show higher levels of attachment anxiety and lower attachment security in comparison to healthy controls (Eng, Heimberg, Hart, Schneier, & Liebowitz, 2001; Manes et al., 2016). Additionally, they show an altered cognitive experience (e.g., self-focused attention bias) (Clark & Wells, 1995), and also increased verbal self-talk as well as special nonverbal characteristics such as the avoidance of eye contact (Heerey & Kring, 2007; Kang & Gratch, 2012), lower level of movement (Levitan et al., 2012) and a lower level of synchrony in interactions with strangers (Vrijssen, Lange, Becker, & Rinck, 2010). Some studies have also found a higher level of facial synchrony in anxious patients (Dijk, Fischer, Morina, van Eeuwijk, & van Kleef, 2018). In addition, patients with higher levels of movement synchrony report a better therapeutic alliance as well as better outcomes (Altmann et al., submitted; Ramseyer & Tschacher, 2011). The kind of treatment seems to moderate the synchrony-outcome association (Altmann et al., submitted).

Taken together, attachment style, anxiety symptoms and interpersonal variables such as nonverbal synchrony seem to be closely linked with each other. To fully understand psychotherapy process, it is necessary to understand these associations more detailed.

Research Questions

Our research question is twofold: First, we want to examine whether a) self-reported social anxiety and avoidance and b) self-reported attachment anxiety and avoidance obtained at the beginning of therapy is associated with various aspects of movement synchrony in the first third of therapy (total synchrony, synchrony-led by patient/therapist, mean/range time lag (total, led by the patient/therapist)). Second, we want to investigate whether these aspects of movement synchrony are associated with the questionnaire data (social anxiety/avoidance, attachment anxiety/avoidance) obtained at the end of therapy. In addition, we want to examine if the specific therapy (i.e., CBT or PDT) is a moderator of this association.

Method

Data source

Data source was the video-analysis project “Timing of nonverbal patient-therapist-interaction and therapeutic success of social phobic patients (TimPaThIn)” conducted from 2016 to 2018. In this project, data from the multi-site SOPHO-NET research network and an outpatient clinic in the southwest of Germany were combined for secondary data analyses. The SOPHO-NET research network (Leichsenring et al., 2013, 2014) was an initiative to study patients in psychotherapy for social anxiety with an RCT comparing different therapy approaches at five different study centers as the core project. Therapies in these centers were either manual-guided short-term outpatient cognitive behavioral therapy (CBT) or manual-guided short-term outpatient psychodynamic-oriented therapy (PDT). At the study center in southwest Germany, naturalistic CBT (CBT_{nat}) was conducted either as short or long-term therapy. For a detailed description of the original study, see Schoenherr, Paulick, Strauss, et al.

(2019). As some measures were assessed by the SOPHO-NET research project only, we focused on the manual-guided CBT and PDT treatments, while omitting details about the CBT_{nat} (see the participants section for details of study flow). The local Ethics Committees of all involved centers approved the study (4189-08/14), which was conducted according to the Guidelines for Good Clinical Practice.

Participants

Primary clinical inclusion criteria of the SOPHO-NET study were the presence of a social anxiety disorder as the primary diagnosis, diagnosed with the Structural Clinical Interview for DSM-IV (SCID; Wittchen, Wunderlich, Gruschwitz, & Zaudig, 1997) and a score of above 30 on the Liebowitz Social Anxiety Scale (LSAS; Mennin et al., 2002). Exclusion criteria were the presence of a cluster A or B personality disorder, substance abuse, psychotic symptoms, organic mental disorders, risk of self-harm, or other psychological or psychiatric interventions. Technical inclusion criteria for the video analysis study were applied after removing patients with clinical exclusion criteria (sample size $n = 469$). These were in particular: the presence of a video of the 3rd, 8th, 20th or 29th therapy session (± 1 session), where both interacting persons (therapist and patient) were visible separately and acted in a verbal interaction (exclusion of $n = 114$). Exclusion criteria included another person in the room, instable light conditions, lack of a stable background (e.g., moving cars behind windows), an overlap of the bodies of therapist and patient in the video recording, the application of role-plays, or filling-in questionnaires (exclusion of $n = 88$).

Regarding the present study, we also excluded the naturalistic CBT arm ($n = 110$), because the Experience in Close Relationships Questionnaire (ECR) and the Liebowitz Social Anxiety Scale (LSAS) were not assessed in this study center. As we focused on movement synchrony in the 8th therapy session, we also excluded patients who had no video of this session

($n = 57$), resulting in a final sample of $n=100$ patients (CBT: $n = 57$, PDT: $n = 43$). All patients provided informed consent for study participation.

Treatment and setting

Both therapy arms (CBT & PDT) had a maximal duration of 30 sessions (mean duration: CBT: 26.0 ($SD = 4.9$) sessions; PDT: 21.7 ($SD = 7.3$) sessions). The first five sessions were “probatory” sessions, which are compulsory in the German health care system. These sessions are typically used to clarify administrative issues and to determine indication for psychotherapy. All therapy sessions were conducted on a weekly basis by 38 therapists who were either advanced postgraduates or licensed therapists. Postgraduate therapists were supervised by experienced therapists every fourth session. Therapists had a mean age of 34.25 years ($SD = 8.86$) and 67 % were female. Each therapist treated on average 2.63 patients (ranging from 1 to 8). Cognitive behavioral therapy was based on the manual by Stangier, Clark, and Ehlers (2006). The psychodynamic therapy was based on Luborsky’s supportive-expressive treatment model (Luborsky, 1995) that was specifically modified for patients with social anxiety disorders (Leichsenring et al., 2013, 2014). Videos were recorded with one-hand cameras during therapy after patients gave their consent. Therapists and patients were sitting facing each other.

Measures

Experiences in Close Relationships Questionnaire (ECR). The Experiences in Close Relationships Questionnaire is a self-report measure that assesses attachment anxiety (ECR-ANX) and attachment avoidance (ECR-AVO) of individuals with respect to their feelings in close relationships. It contains 36 items (18 items assessing anxiety, 18 items assessing avoidance), which are answered on a 7-point Likert scale ranging from 1 (not at all) to 7 (very much) (Brennan, Clark, & Shaver, 1998). High values indicate a high degree of attachment

anxiety or avoidance. Cronbach's alphas for both scales were reported as high, ranging from $\alpha = .91$ to $.94$ (Brennan et al., 1998).

Liebowitz Social Anxiety Scale (LSAS). The Liebowitz Social Anxiety Scale (Mennin et al., 2002) is a scale which can be applied as the clinician-administered or self-report version. Both versions consist of 24 situations (13 performance, 11 social situations) that have to be rated with respect to the past week on a 0-3 Likert scale. Ratings are done with respect to anxiety and avoidance resulting in six scales: anxiety/avoidance of performance, anxiety/avoidance of social interaction, and total anxiety/avoidance. We are using the total score (LSAS-TOT), total anxiety (LSAS-ANX) and total avoidance (LSAS-AVO) scales of the clinician-administered version. Cronbach's alpha of the scales and the total score has been shown to be high, ranging from $\alpha = .81$ to $.96$ (Heimberg et al., 1999). Inter-rater-reliability was high (Kendall's $W = 0.98$, $df = 2$, $p < .001$) (Leichsenring et al., 2013).

Body movements. The movements of the interaction partners (therapist, patient) were automatically coded using motion energy analysis (MEA; Altmann, 2013; <https://github.com/10101-00001/MEA>). Twenty videos were previously subjected to a preliminary analysis to determine a data-driven cut-off value for the necessary intensity changes required to count a pixel-change as *movement*. This is realized by pixels in the background (where no movement takes place). MEA of these pixels was computed and then the 99% quantile of intensity changes that occur in the background was determined. In our case, we set the threshold to 12. This is in line with the default values of another MEA software that range between 10 and 15 (Ramseyer & Tschacher, 2011). We computed MEA for the first 15 min of the sequence (after administrative issues, filling-in of questionnaires, and started latest after 10 min), as this segment has been shown to be representative for the amount of movement synchrony of the entire session in previous studies (Paulick et al., 2017; Ramseyer & Tschacher, 2011). MEA is performed for two different regions of interest (ROIs): For this

purpose, the upper body of the person of interest was framed with the freehand tool contained in the script. To make movement of the interaction partners comparable, the motion energy (ME) values were divided by the respective ROI pixel size and multiplied by 100. A value of 20 therefore means the activation of 20% of the upper body ROI. The MEA output provides 4 time series: 2 for the specified ROI and 2 in the background, which were used for filtering image noise. In addition to filtering via the background ROIs, all ME values that reflect movements that were too fast to correspond to human movements (resting state - movement of 15% of the upper body - resting state within 0.12 seconds) were also deleted. The resulting missing data in the time series were filled by linear interpolation. Small irregularities in the signal were smoothed with a moving median with the bandwidth 5. Before running the time series analysis, the ME time series of patient and therapist were logarithmized to make the heights of the movement-peaks comparable. These settings have been validated by Schoenherr, Paulick, Deisenhofer, et al. (2019).

Movement synchrony. Movement synchrony was also identified automatically. For this purpose, windowed cross-lagged correlation was used in combination with a specific peak-picking algorithm (GitHub: https://github.com/10101-00001/sync_ident; Altmann, 2011; Altmann, 2013; Schoenherr et al., 2018). The result of the WCLC is a matrix that contains a correlation (in our case R^2) at every time point of the time series and every time delay up to the maximum time lag. We used the following settings for the WCLC: bandwidth 125 frames (= 5 seconds), no time series smoothing, maximum time lag 125 frames (= 5 seconds), step: 1 frame (= 0.04 seconds) = overlapping windows (Schoenherr, Paulick, Deisenhofer, et al., 2019). Afterwards, the resulting matrix is filtered by the peak-picking algorithm. Sequences are extracted that show a strong cross-correlation with the same time lag (± 1 frame = 0.04 seconds) over a period of time (at least 10 frames = 0.4 seconds). If there are overlapping synchronization intervals, the interval with the higher average R^2 is selected. The extracted intervals are stored

in a list of synchrony intervals (LOSI). In this list, each interval has a start and end point, a time lag as well as average synchrony strength (R^2). A positive time lag indicates that the patient initiated synchrony whereas a negative time lag indicates that the therapist initiated synchrony. In the current study, the measures of interest were the frequency of synchrony, average time lag, and the range of the time (total, led by the patient/therapist). Frequency was calculated by summing the duration of the synchronization intervals and then dividing the result by the total duration of the video (15min). For better interpretation, the value was multiplied by 100. Therefore, it represents the percentage of the time during which the interaction was synchronous. The mean time lag is the mean value of the absolute time lags, while the range is operationalized by the range of the absolute time lags. Analogue procedures were applied to calculate patient- and therapist-led synchrony measures, except that only those intervals were used that had either a positive (patient leading) or negative time lag (therapist leading).

Statistical analyses

Imputation. Missing values in our dataset were due to incomplete filling-in of questionnaires and premature terminations. They ranged from 0 to 25%, with most variables having low amount of missing values (mean rate of missing values: 8.5%, median 4%). We used demographic and clinical variables to predict missing values using the missForest package in R (R Core Team, 2017; Stekhoven & Bühlmann, 2012). The little MCAR test suggested that missings were completely at random ($\chi^2 = 288$, $df = 784$, $p > .99$). We used the following settings for imputation: seed = 94, maximum number of iterations = 20, number of random trees = 1000. Fit indices suggested successful imputation (normalized root-mean-square error = .18, proportion of false classified = .07).

Matching. In the original SOPHO-NET research network study, patients were randomized to CBT or PDT. Due to the fact, that we only used those patients who had a video of the 8th therapy session, CBT and PDT are not comparable with regard to demographic

variables. Therefore, we conducted propensity score matching (full matching) using patient age, gender, current relationship, and education. Matching was conducted with the MatchIt package (Ho, Imai, King, & Stuart, 2007, 2011) in R. Standard mean differences ranged between -.13 and .11 (i.e., $<.25$) indicating a good match (Rubin, 2001).

Data analysis. Descriptive statistics of the sample and the outcome variables LSAS and ECR are presented for CBT and PDT separately. Referring to the outcome statistics, pre-post comparisons were calculated using *t*-tests for paired samples. Note, that the comparison of CBT and PDT in terms of efficacy was not the aim of the present paper; for the major results of the SOPHO-NET study see Leichsenring et al. (2013). The amount of therapist and patient-led synchrony was compared using a *t*-test for paired samples.

To answer our first research question, two-sided Pearson product-moment correlations were conducted between the synchrony measures (synchrony, mean/range time lag (total, patient leading, therapist leading)) and the initial questionnaire data LSAS-TOT, LSAS-ANX, LSAS-AVO, and ECR-ANX, ECR-AVO. Research suggests the existence of an interaction effect with therapy approach (Altmann et al., submitted). Therefore, we also analyzed the data using a stepwise backwards mixed effects regression including the propensity scores (PS) as regressions weights with all interactions effects of synchrony measures or questionnaire data with the therapy approach. We used the synchrony measure as a criterion and included all questionnaire measures simultaneously as predictors as well as the therapy approach and all interactions with therapy approach. According to Altmann et al. (submitted) random effects for the therapists were included, because one therapist treated more than one patient (random intercept for patients within therapists). Multi-level analysis was only conducted when correlations pointed towards an association.

To answer our second research question, the procedure was similar as described for examining the first research question: We started with correlations and conducted PS-weighted

mixed effects linear models if the correlations pointed towards a meaningful association. LSAS and ECR scales assessed at the end of therapy were the criteria, while synchrony measures, initial questionnaire values (LSAS, ECR), therapy approach, and all interactions with therapy approach were used as predictors.

To investigate the therapy approach as moderator, we included interaction terms with the therapy approach in the regression models as described, and conducted a multi-group model for regression models showing a significant interaction term between the synchrony measure and the therapy approach.

Results

Descriptive statistics

Descriptive sample statistics are displayed in Table 1. Pre-post differences were significant for ECR avoidance ($t = 2.76, p = .01$), LSAS anxiety ($t = 12.57, p < .01$), and LSAS avoidance ($t = 11.37, p < .01$). The pre-post difference for ECR anxiety showed a statistical trend ($t = 1.64, p = .10$). The average amount of synchrony at session 8 was: synchrony total: 20.77% ($SD = 9.14$), synchrony patient-led: 10.59% ($SD = 4.90$), synchrony therapist-led: 10.01% ($SD = 4.55$), mean time lag: 63.77 frames ($SD = 4.14$), mean time lag therapist (patient leading): 63.05 frames ($SD = 6.76$), mean time lag patient (therapist leading): 63.50 frames ($SD = 6.64$), range time lag: 122.09 frames ($SD = 2.00$), range time lag therapist (patient leading): 118.72 frames ($SD = 7.33$), and range time lag patient (therapist leading): 119.61 frames ($SD = 4.06$). Patients lead synchrony significantly more often than therapists do at session 8 ($t = -1.98, p = .05$).

Prediction of synchrony measures

Table 2 shows correlations between the initial questionnaire data and synchrony measures. We found significant positive correlations between synchrony total and LSAS pre scales, as well as synchrony led by the therapist and LSAS-TOT, LSAS-ANX. With respect to the time lag, only the mean time lag (total and therapist leading) was significantly negatively associated with the ECR-AVO pre. As correlations only display the association between two variables, disregarding other important influences, regressions including all important predictors were also carried out when correlations pointed towards a meaningful association. All pre-therapy questionnaire data, the therapist, and therapy approach (+ interactions with questionnaire data) were used to predict synchrony measures at s8.

For the prediction of synchrony total and synchrony led by the therapist, the therapy approach was the only significant predictor. The random intercept accounted for about 16.5% of variance for synchrony total (10.7% for synchrony led by the therapist). Therapy approach had a moderate influence on the synchrony variable (synchrony total: $\beta = -.52, p < .01$; synchrony led by the therapist: $\beta = -.54, p < .01$). Thereby, a negative coefficient indicates lower synchrony in PDT. No significant predictor was found for the prediction of the mean time lag of synchrony (total). For the mean time lag of synchrony of the patient (therapist leading), ECR-AVO pre was a significant predictor ($\beta = .32, p < .01$).

Prediction of post treatment measures

Correlations between synchrony measures and outcome questionnaire data (post) are displayed in Table 2. The mean time lag (total, patient/therapist leading) was significantly associated with ECR-AVO post. In addition, two measures of synchrony (synchrony total, synchrony led by the therapist) were associated with ECR-ANX post. Regarding the regressions, pre-therapy questionnaire data, the synchrony measure, the therapist, and therapy

approach (+ interactions with questionnaire data and synchrony) were used to predict the post-therapy questionnaire data.

Thus, five regressions were conducted for the post-therapy questionnaire data: 1-2) ECR-ANX as criterion with total synchrony or therapist-led synchrony as the predictor, 3-5) ECR-AVO as criterion with the mean time lag total/mean time lag of the patient (therapist leading) or mean time lag of the therapist (patient leading) as the predictor. All other meaningful predictors that remained in the models are displayed in Table 3.

Treatment approach as moderator

Since synchrony was differentially associated with attachment dimensions in PDT and CBT as indicated by the significant interaction, we also applied a multi-group model for synchrony total/led by the therapist and the mean time lag of the patient (therapist leading) as criteria. Results are presented in Table 4. Synchrony was especially relevant for the change in ECR-ANX when it is led by the therapist in CBT; for PDT we did not find a significant influence. The mean time lag of the patient is associated with attachment avoidance in CBT, too. The influence of the therapist as cluster variable varied between CBT and PDT (synchrony total: $ICC_{CBT} = .40$, $ICC_{PDT} = .07$; synchrony led by the therapist: $ICC_{CBT} = .40$, $ICC_{PDT} = .09$, mean time lag therapist leading: $ICC_{CBT} = .03$, $ICC_{PDT} = .00$).

Discussion

The pre-post comparisons of the examined variables showed that anxiety symptoms as well as attachment characteristics change through psychotherapy. Anxiety symptoms improved significantly over the course of therapy. Attachment changed as well, especially with regard to the amount of attachment anxiety. It has been shown in other studies that attachment-related avoidance and anxiety change during therapy (Strauss et al., 2018; Taylor et al., 2015). The improvement of attachment avoidance can only marginally be found in the sample we examined. In the entire sample of the SOPHO-NET trial (495 patients) post-therapy, patients

treated with CBT revealed significant changes in attachment anxiety and avoidance, whereas patients treated with PDT showed no significant changes. Changes between post treatment and the two follow-ups were significant in both treatment conditions (Strauss et al., 2018).

At the beginning of therapy the therapist shows more following that might be a sign of giving the patient room for his/her explanations of their symptoms. The therapist might induce the feeling of being understood for the patient by reacting on his/her movements.

Regarding the prediction of synchrony measures based on the initial questionnaire data, we found that synchrony total and therapist-led synchrony in the first third of therapy is neither dependent on initial anxiety/avoidance symptoms nor on attachment anxiety/avoidance, but on the person of the therapist and the kind of treatment. That we did not find the expected relation of attachment to synchrony (Feniger-Schaal & Lotan, 2017) and outcome and synchrony (Paulick et al., 2018) might be due to our homogeneous sample of patients with SAD, the used synchrony algorithm or the different contexts while assessing synchrony. Several studies showed the context-dependency of the effects of synchrony (Wallot, Mitkidis, McGraw, & Roepstorff, 2016). In comparison to the therapist effect of 8% demonstrated by Ramseyer and Tschacher (2011) and 9% in Paulick et al.'s (2017) study, in the current study the therapist effect was higher (16.5%). In contrast to both mentioned studies, our study measured the frequency of synchrony by applying WCLC, and a peak-picking algorithm. Thus, intervals with no synchrony were neglected. In the studies by Ramseyer and Tschacher (2011) and Paulick et al. (2017), these intervals are also aggregated when calculating the global synchrony score. This may be an explanation for obtaining a higher amount of variance explained by the therapists in our study. Nonverbal synchrony is influenced by the regulatory competences of both individuals (Koole & Tschacher, 2016) and closely related to empathy (Chartrand & Bargh, 1999; Finset & Ornes, 2017; Kleinbub et al., 2012). Therefore, it is plausible that therapists differ with regard to their ability to initiate or establish synchrony.

The fact that movement synchrony is a dyadic construct that cannot be predicted by characteristics of only one person might explain that synchrony measures were independent of attachment pattern and anxiety symptoms. In this context, Kivlighan Jr and Marmarosh (2018) showed that the estimated therapeutic alliance and, above all, the congruence of the patient- and therapist-rated alliance, which has been shown to be related to treatment outcome (Rubel, Bar-Kalifa, Atzil-Slonim, Schmidt, & Lutz, 2018) depends largely on the therapist's attachment style. A connection might be found if the therapist's attachment characteristics, anxiety or the fit of both attachment styles are taken into account.

With respect to the mean time lag, we find that the level of attachment-related avoidance can predict the mean time lag of the patient (therapist leading). Persons with higher levels of attachment-related avoidance showed higher reaction times on attachment-related primes (Mikulincer et al., 2002). Therefore, the additional inhibition of proximity seeking of the patients might lead to higher time lags in synchronizing with the therapist.

Regarding the prediction of the outcome, we found a connection between attachment avoidance and a higher time lag of synchrony, which is consistent with Feldman et al. (1999) and Zlochower and Cohn (1996). Patients with higher attachment avoidance engage in movement led by the therapist in a delayed manner. The higher time lag may indicate an inhibition of establishing contact with the therapist as a deactivation strategy. Highly avoidant individuals lack social skills (e.g., self-efficacy, emotional awareness), leading to interpersonal problems (Mallinckrodt, 2000). Avoidant patients manage to engage with the therapist, as they do not differ with regard to the amount of synchrony. However, a deactivation of the attachment system and low emotional awareness seem to lead to an increase in the time it takes to establish contact with the therapist. Thus, a larger time lag in synchronizing might be an indicator of higher attachment-related avoidance during therapy.

Frequent synchrony was associated with lower values of patient attachment anxiety at the end of therapy. The positive experience of synchronous movement could give the patient the feeling of being understood, allowing the therapist to become a secure base (Bowlby, 1973). Attachment anxiety is reduced by experiencing an appropriate close relationship. This is consistent with findings on the association between the therapeutic alliance and outcome. It has been shown that therapists who form a stronger therapeutic alliance early in treatment have patients who better improve during therapy (Baldwin, Wampold, & Imel, 2007). Following the In-Sync model of psychotherapy, the patient's regulatory competencies might have been increased by therapist-led synchrony of movement. The lacking association of movement synchrony and the attachment avoidance scale suggests that the avoidant attached individuals are not that responsive, which might be due to the repression of proximity seeking and being in contact with the therapist.

Contrary to Gueguen (2011), we found no association between social anxiety and avoidance and movement synchrony by controlling for attachment anxiety and avoidance. This might indicate that movement synchrony as assessed with our algorithm is rather a proximal variable to interpersonal constructs and not directly associated with symptomatology.

Looking at the results of the multi-group model, the effect of synchrony on attachment-related anxiety is particularly present in CBT, which is consistent with Altmann et al. (submitted) showing an association between movement synchrony and the reduction of interpersonal problems. In PDT, the correlation is not significant. Particularly in the psychodynamic setting, a number of studies on mimic synchrony have indicated that increased synchrony might have a negative effect on conflict tension. Therefore, it might have negative consequences on the treatment outcome (Bänninger-Huber & Huber, 2017; Rasting & Beutel, 2005). It can be assumed that nonverbal synchrony is clearly dependent on the content of the sessions and interventions. Evidence for the context dependency of nonverbal synchrony has

been provided by studies in naturalistic and experimental settings (Allsop, Vaitkus, Marie, & Miles, 2016; Duran & Fusaroli, 2017; Paxton & Dale, 2017; Tschacher, Rees, & Ramseyer, 2014; Wallot et al., 2016). Thus, the effects of nonverbal synchrony during cognitive restructuring, behavioral analysis or relationship-episode interviews may have very different effects on outcome. Instead of aggregating synchrony across interventions, future studies should investigate synchrony in more standardized settings. We think that the effects were present especially in CBT relates to the fact that CBT sessions are more structured by the therapist that might have a positive influence on the attachment anxiety of socially anxious individuals.

Strengths & Limitations

One strength of the current study is the investigation of several movement synchrony characteristics (e.g., mean time lag, total frequency, patient-led synchrony) in different contexts (CBT vs. PDT). This approach allows the investigation of differential principles of synchrony. Furthermore, the conclusions of the correlation analysis were supported by more sophisticated models including propensity score matching and multi-level modeling. As a result, spurious correlations were avoided. The automated assessment of movements and their synchrony with publicly accessible scripts also increases transparency and replicability. Attachment dimensions and anxiety/ avoidance variables were included in one analysis to disentangle the different effects of movement synchrony.

A clear limitation is our aggregation of synchrony over time and interventions. Furthermore, synchrony was investigated disregarding the content. In order to further investigate the differential effects of synchrony, larger samples and standardized settings are necessary. In addition, attachment characteristics were assessed using self-report measures, which might lead to biases.

Practical implications

With regard to the patient's attachment style, it seems beneficial as a therapist to offer various movements to increase the probability of movement synchrony. Guidance seems to convey to the patient a sense of a safe haven. If lacking synchrony is noticed, it can be actively addressed in the therapy process. A large time lag until synchrony is established might point towards attachment-related avoidance. However, it should not be consciously imitated, as this has a negative effect on the therapeutic relationship (White & Argo, 2011). In order to increase the perception of nonverbal aspects, information on synchrony should be explicitly integrated into psychotherapist training.

Conclusion

In summary, the results show that lower attachment anxiety at the end of therapy was associated with frequent movement synchrony in the first third of therapy. To avoid spurious correlations, it is important that all effects are integrated into one (multi-level) regression model. By doing so, we found that the therapist might serve as a secure base for the patient by leading movement synchrony. Attachment avoidance was predicted by the mean time lag. Following our results, a high time lag is an indicator of a patient with higher values of attachment avoidance. However, there was no association of synchrony with symptom anxiety or avoidance. The experienced interpersonal relationship influences attachment patterns and interpersonal problems at the end of therapy; however, it seems to have no direct influence on symptoms.

References

- Ainsworth, M. D., Blehar, M. C., Waters, E., & Wall, S. (1978). *Patterns of attachment: Assessed in the strange situation and at home*. Hillsdale, NJ: Erlbaum.
- Allsop, J. S., Vaitkus, T., Marie, D., & Miles, L. K. (2016). Coordination and collective performance: Cooperative goals boost interpersonal synchrony and task outcomes. *Frontiers in Psychology*, 7. doi:10.3389/fpsyg.2016.01462
- Altmann, U. (2011). Investigation of movement synchrony using windowed crosslagged regression. *Analysis of Verbal and Nonverbal Communication and Enactment: The Processing Issue*, 344-354. doi:10.1007/978-3-642-25775-9_31
- Altmann, U. (2013). *Synchronisation nonverbalen Verhaltens: Weiterentwicklung und Anwendung zeitreihenanalytischer Identifikationsverfahren. [Synchronization of nonverbal behavior: Development and Application of time series analysis methods]*: Springer-Verlag.
- Altmann, U., Schoenherr, D., Paulick, J., Deisenhofer, A.-K., Schwartz, B., Rubel, J., . . . Strauss, B. (submitted). Timing of nonverbal patient-therapist-interaction and therapeutic success of social phobic patients. *Psychotherapy Research*.
- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders (5th ed.)*. Washington, DC: Author.
- Baldwin, S. A., Wampold, B. E., & Imel, Z. E. (2007). Untangling the alliance-outcome correlation: Exploring the relative importance of therapist and patient variability in the alliance. *Journal of Consulting and Clinical Psychology*, 75(6), 842.
- Bänninger-Huber, E., & Huber, E. (2017). Nonverbal Processes in Psychotherapeutic Interaction. *Journal of Psychology and Psychotherapie Research*, 4 (1), 1-8. doi:10.12974/2313-1047.2017.04.01.1

- Beebe, B., Lachmann, F. M., Markese, S., Buck, K. A., Bahrnick, L. E., Chen, H., . . . Jaffe, J. (2012). On the origins of disorganized attachment and internal working models: Paper II. An empirical microanalysis of 4-month mother–infant interaction. *Psychoanalytic dialogues*, 22(3), 352-374.
- Bernecker, S. L., Levy, K. N., & Ellison, W. D. (2014). A meta-analysis of the relation between patient adult attachment style and the working alliance. *Psychotherapy Research*, 24(1), 12-24.
- Bowlby, J. (1973). *Attachment and loss, vol. II: Separation* (Vol. 2): Basic Books New York.
- Brennan, K. A., Clark, C. L., & Shaver, P. R. (1998). Self-report measurement of adult attachment: An integrative overview.
- Cameron, J. J., Finnegan, H., & Morry, M. M. (2012). Orthogonal dreams in an oblique world: A meta-analysis of the association between attachment anxiety and avoidance. *Journal of Research in Personality*, 46(5), 472-476.
- Chartrand, T. L., & Bargh, J. A. (1999). The chameleon effect: The perception–behavior link and social interaction. *Journal of Personality and Social Psychology*, 76(6), 893. doi:10.1037/0022-3514.76.6.893
- Clark, D. M., & Wells, A. (1995). A cognitive model of social phobia. *Social phobia: Diagnosis, assessment, and treatment*, 41(68), 00022-00023.
- Colonesi, C., Draijer, E. M., Jan JM Stams, G., Van der Bruggen, C. O., Bögels, S. M., & Noom, M. J. (2011). The relation between insecure attachment and child anxiety: A meta-analytic review. *Journal of Clinical Child & Adolescent Psychology*, 40(4), 630-645.
- Dales, S., & Jerry, P. (2008). Attachment, affect regulation and mutual synchrony in adult psychotherapy. *American Journal of Psychotherapy*, 62(3), 283-312. doi:10.1176/appi.psychotherapy.2008.62.3.283

-
- Dijk, C., Fischer, A. H., Morina, N., van Eeuwijk, C., & van Kleef, G. A. (2018). Effects of Social Anxiety on Emotional Mimicry and Contagion: Feeling Negative, but Smiling Politely. *Journal of Nonverbal Behavior*, *42*(1), 81-99.
- Duran, N. D., & Fusaroli, R. (2017). Conversing with a devil's advocate: Interpersonal coordination in deception and disagreement. *PLoS One*, *12*(6), e0178140. doi:10.1371/journal.pone.0178140
- Eng, W., Heimberg, R. G., Hart, T. A., Schneier, F. R., & Liebowitz, M. R. (2001). Attachment in individuals with social anxiety disorder: the relationship among adult attachment styles, social anxiety, and depression. *Emotion*, *1*(4), 365.
- Feldman, R. (2012). Parent–infant synchrony: A biobehavioral model of mutual influences in the formation of affiliative bonds. *Monographs of the Society for Research in Child Development*, *77*(2), 42-51.
- Feldman, R., Greenbaum, C. W., & Yirmiya, N. (1999). Mother–infant affect synchrony as an antecedent of the emergence of self-control. *Developmental Psychology*, *35*(1), 223.
- Feldman, R., Magori-Cohen, R., Galili, G., Singer, M., & Louzoun, Y. (2011). Mother and infant coordinate heart rhythms through episodes of interaction synchrony. *Infant Behavior and Development*, *34*(4), 569-577.
- Feniger-Schaal, R., Hart, Y., Lotan, N., Koren-Karie, N., & Noy, L. (2018). The Body Speaks: Using the Mirror Game to Link Attachment and Non-verbal Behavior. *Frontiers in Psychology*, *9*(1560). doi:10.3389/fpsyg.2018.01560
- Feniger-Schaal, R., & Lotan, N. (2017). The embodiment of attachment: Directional and shaping movements in adults' mirror game. *The Arts in Psychotherapy*, *53*, 55-63. doi:10.1016/j.aip.2017.01.006

- Finset, A., & Ornes, K. (2017). Empathy in the Clinician-Patient Relationship: The Role of Reciprocal Adjustments and Processes of Synchrony. *Journal of Patient Experience*, 4(2), 64-68. doi:10.1177/2374373517699271
- Galbusera, L., Finn, M. T., & Fuchs, T. (2016). Interactional synchrony and negative symptoms: An outcome study of body-oriented psychotherapy for schizophrenia. *Psychother Res*, 1-13. doi:10.1080/10503307.2016.1216624
- Gueguen, N. (2011). The mimicker is a mirror of myself: Impact of mimicking on self-consciousness and social anxiety. *Social Behavior and Personality: an international journal*, 39(6), 725-728.
- Hall, N. R., Millings, A., & Bouças, S. B. (2012). Adult Attachment Orientation and Implicit Behavioral Mimicry. *Journal of Nonverbal Behavior*, 36(4), 235-247. doi:10.1007/s10919-012-0136-7
- Harrist, A. W., & Waugh, R. M. (2002). Dyadic synchrony: Its structure and function in children's development. *Developmental Review*, 22(4), 555-592. doi:10.1016/S0273-2297(02)00500-2
- Håvås, E., Svartberg, M., & Ulvenes, P. (2015). Attuning to the unspoken: The relationship between therapist nonverbal attunement and attachment security in adult psychotherapy. *Psychoanalytic psychology*, 32(2), 235. doi:10.1037/a0038517
- Heerey, E. A., & Kring, A. M. (2007). Interpersonal consequences of social anxiety. *Journal of abnormal psychology*, 116(1), 125.
- Heimberg, R. G., Horner, K., Juster, H., Safren, S., Brown, E., Schneier, F., & Liebowitz, M. (1999). Psychometric properties of the Liebowitz social anxiety scale. *Psychological Medicine*, 29(1), 199-212.

-
- Ho, D. E., Imai, K., King, G., & Stuart, E. A. (2007). Matching as nonparametric preprocessing for reducing model dependence in parametric causal inference. *Political analysis*, 15(3), 199-236. doi:10.1093/pan/mpi013
- Ho, D. E., Imai, K., King, G., & Stuart, E. A. (2011). MatchIt: nonparametric preprocessing for parametric causal inference. *Journal of Statistical Software*, 42(8), 1-28. doi:10.18637/jss.v042.i08
- Kang, S.-H., & Gratch, J. (2012). *Understanding the nonverbal behavior of socially anxious people during intimate self-disclosure*. Paper presented at the International Conference on Intelligent Virtual Agents.
- Khledian, M., Garosi, M. R., Khaikhah, Z., & Ghalandari, S. (2013). The Relationship between Attachment Lifestyle with Depression, Hopefulness and Emotional Intelligence. *Open Journal of Social Science Research*, 1(2), 15-21.
- Kivlighan Jr, D. M., & Marmarosh, C. L. (2018). Counselors' attachment anxiety and avoidance and the congruence in clients' and therapists' working alliance ratings. *Psychotherapy Research*, 28(4), 571-580.
- Kleinbub, J. R. (2016). The rhythm of therapy: psychophysiological synchronization in clinical dyads.
- Kleinbub, J. R. (2017). State of the Art of Interpersonal Physiology in Psychotherapy: A Systematic Review. *Frontiers in Psychology*, 8. doi:10.3389/fpsyg.2017.02053
- Kleinbub, J. R., Messina, I., Bordin, D., Voci, A., Calvo, V., Sambin, M., & Palmieri, A. (2012). *Synchronization of skin conductance levels in therapeutic dyads* (Vol. 86).
- Koole, S. L., & Tschacher, W. (2016). Synchrony in Psychotherapy: A Review and an Integrative Framework for the Therapeutic Alliance. *Frontiers in Psychology*, 7(862). doi:10.3389/fpsyg.2016.00862

- Leclère, C., Viaux, S., Avril, M., Achard, C., Chetouani, M., Missonnier, S., & Cohen, D. (2014). Why synchrony matters during mother-child interactions: a systematic review. *PLoS One*, *9*(12), e113571.
- Leichsenring, F., Salzer, S., Beutel, M. E., Herpertz, S., Hiller, W., Hoyer, J., . . . Poehlmann, K. (2013). Psychodynamic therapy and cognitive-behavioral therapy in social anxiety disorder: a multicenter randomized controlled trial. *American Journal of Psychiatry*. doi:10.1176/appi.ajp.2013.12081125
- Leichsenring, F., Salzer, S., Beutel, M. E., Herpertz, S., Hiller, W., Hoyer, J., . . . Poehlmann, K. (2014). Long-term outcome of psychodynamic therapy and cognitive-behavioral therapy in social anxiety disorder. *American Journal of Psychiatry*.
- Levitan, M., Crippa, J., Bruno, L., Pastore, D., Freire, R., Arrais, K., . . . Nardi, A. (2012). Postural balance in patients with social anxiety disorder. *Brazilian Journal of Medical and Biological Research*, *45*(1), 38-42.
- Levy, K. N., Kivity, Y., Johnson, B. N., & Gooch, C. V. (2018). Adult attachment as a predictor and moderator of psychotherapy outcome: A meta-analysis. *Journal of Clinical Psychology*, *74*(11), 1996-2013.
- Lindsey, E. W., & Caldera, Y. M. (2015). Shared affect and dyadic synchrony among secure and insecure parent-toddler dyads. *Infant and Child Development*, *24*(4), 394-413.
- Luborsky, L. (1995). Einführung in die analytische Psychotherapie: ein Lehrbuch. 2. Aufl. Göttingen, Zürich: Vandenhoeck und Ruprecht.
- Mallinckrodt, B. (2000). Attachment, social competencies, social support, and interpersonal process in psychotherapy. *Psychotherapy Research*, *10*(3), 239-266.
- Manes, S., Nodop, S., Altmann, U., Gawlytta, R., Dinger, U., Dymel, W., . . . Petrowski, K. (2016). Social anxiety as a potential mediator of the association between attachment

- and depression. *Journal of Affective Disorders*, 205, 264-268.
doi:10.1016/j.jad.2016.06.060
- Mennin, D. S., Fresco, D. M., Heimberg, R. G., Schneier, F. R., Davies, S. O., & Liebowitz, M. R. (2002). Screening for social anxiety disorder in the clinical setting: using the Liebowitz Social Anxiety Scale. *Journal of anxiety disorders*, 16(6), 661-673.
doi:10.1016/S0887-6185(02)00134-2
- Mikulincer, M., Birnbaum, G., Woddis, D., & Nachmias, O. (2000). Stress and accessibility of proximity-related thoughts: Exploring the normative and intraindividual components of attachment theory. *Journal of Personality and Social Psychology*, 78(3), 509.
- Mikulincer, M., Gillath, O., & Shaver, P. R. (2002). Activation of the attachment system in adulthood: threat-related primes increase the accessibility of mental representations of attachment figures. *Journal of Personality and Social Psychology*, 83(4), 881.
- Palmieri, A., Kleinbub, J. R., Calvo, V., Benelli, E., Messina, I., Sambin, M., & Voci, A. (2018). Attachment-security prime effect on skin-conductance synchronization in psychotherapists: An empirical study. *Journal of Counseling Psychology*.
- Paulick, J., Deisenhofer, A.-K., Ramseyer, F., Tschacher, W., Boyle, K., Rubel, J., & Lutz, W. (2017). Nonverbal Synchrony: A new approach to better understand psychotherapeutic processes and drop-out. *Journal of Psychotherapy Integration*. doi:10.1037/int0000099
- Paulick, J., Rubel, J., Deisenhofer, A.-K., Schwartz, B., Thielemann, D., Altmann, U., . . . Lutz, W. (2018). Diagnostic features of nonverbal synchrony in psychotherapy: Comparing depression and anxiety. *Cognitive Therapy and Research*, Advanced online publication. doi:10.1007/s10608-018-9914-9
- Paxton, A., & Dale, R. (2017). Interpersonal Movement Synchrony Responds to High- and Low-Level Conversational Constraints. *Front Psychol*, 8, 1135.
doi:10.3389/fpsyg.2017.01135

- R Core Team. (2017). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from <https://www.R-project.org/>
- Ramseyer, F., & Tschacher, W. (2011). Nonverbal Synchrony in Psychotherapy: Coordinated Body Movement Reflects Relationship Quality and Outcome. *Journal of Consulting and Clinical Psychology, 79*(3), 284-295. doi:10.1037/a0023419
- Rasting, M., & Beutel, M. E. (2005). Dyadic affective interactive patterns in the intake interview as a predictor of outcome. *Psychotherapy Research, 15*(3), 188-193. doi:10.1080/10503300512331335039
- Rubel, J. A., Bar-Kalifa, E., Atzil-Slonim, D., Schmidt, S., & Lutz, W. (2018). Congruence of therapeutic bond perceptions and its relation to treatment outcome: Within-and between-dyad effects. *Journal of Consulting and Clinical Psychology, 86*(4), 341.
- Rubin, D. B. (2001). Using propensity scores to help design observational studies: application to the tobacco litigation. *Health Services and Outcomes Research Methodology, 2*(3-4), 169-188.
- Schoenherr, D., Paulick, J., Deisenhofer, A.-K., Schwartz, B., Rubel, J., Lutz, W., . . . Altmann, U. (2019). Identification of movement synchrony: Validation of windowed cross-lagged correlation and -regression with peak-picking algorithm. *PLoS One, 14*(2). doi:10.1371/journal.pone.0211494
- Schoenherr, D., Paulick, J., Strauss, B., Deisenhofer, A.-K., Schwartz, B., Rubel, J., . . . Altmann, U. (2019). Nonverbal synchrony predicts premature termination of psychotherapy for social phobic patients. *Psychotherapy*.
- Schoenherr, D., Paulick, J., Worrack, S., Strauss, B., Rubel, J. A., Schwartz, B., . . . Altmann, U. (2018). Quantification of nonverbal synchrony using linear time series analysis

- methods: Lack of convergent validity and evidence for facets of synchrony. *Behavior Research Methods*, 1-23. doi:10.3758/s13428-018-1139-z
- Stangier, U., Clark, D. M., & Ehlers, A. (2006). Soziale Phobie: Fortschritte der Psychotherapie. *Göttingen, Hogrefe-Verlag*.
- Stekhoven, D. J., & Buhlmann, P. (2012). MissForest--non-parametric missing value imputation for mixed-type data. *Bioinformatics*, 28(1), 112-118. doi:10.1093/bioinformatics/btr597
- Strauss, B., Altmann, U., Manes, S., Tholl, A., Koranyi, S., Nolte, T., . . . Hiller, W. (2018). Changes of attachment characteristics during psychotherapy of patients with social anxiety disorder: Results from the SOPHO-Net trial. *PLoS One*, 13(3), e0192802.
- Strauss, B., Koranyi, S., Altmann, U., Nolte, T., Beutel, M. E., Wiltink, J., . . . Joraschky, P. (2017). Partner-related attachment as a moderator of outcome in patients with social anxiety disorder—a comparison between short-term cognitive-behavioral and psychodynamic therapy. *Psychotherapy*, 54(4), 339.
- Taylor, P., Rietzschel, J., Danquah, A., & Berry, K. (2015). Changes in attachment representations during psychological therapy. *Psychotherapy Research*, 25(2), 222-238.
- Tschacher, W., Ramseyer, F., & Koole, S. L. (2017). Sharing the Now in the Social Present: Duration of Nonverbal Synchrony Is Linked With Personality. *Journal of personality*.
- Tschacher, W., Ramseyer, F., & Koole, S. L. (2018). Sharing the now in the social present: duration of nonverbal synchrony is linked with personality. *Journal of personality*, 86(2), 129-138. doi:10.1111/jopy.12298
- Tschacher, W., Rees, G., & Ramseyer, F. (2014). Nonverbal synchrony and affect in dyadic interactions. *Frontiers in Psychology*, 5(1323). doi:10.3389/fpsyg.2014.01323

- Vrijesen, J. N., Lange, W.-G., Becker, E. S., & Rinck, M. (2010). Socially anxious individuals lack unintentional mimicry. *Behaviour Research and Therapy, 48*(6), 561-564.
- Wallot, S., Mitkidis, P., McGraw, J. J., & Roepstorff, A. (2016). Beyond synchrony: joint action in a complex production task reveals beneficial effects of decreased interpersonal synchrony. *PLoS One, 11*(12), e0168306.
- White, K., & Argo, J. J. (2011). When imitation doesn't flatter: The role of consumer distinctiveness in responses to mimicry. *Journal of Consumer Research, 38*(4), 667-680.
- Wittchen, H., Wunderlich, U., Gruschwitz, S., & Zaudig, M. (1997). Skid-i. *Strukturiertes klinisches Interview für DSM-IV*.
- Zlochower, A. J., & Cohn, J. F. (1996). Vocal timing in face-to-face interaction of clinically depressed and nondepressed mothers and their 4-month-old infants. *Infant Behavior and Development, 19*(3), 371-374.

Table 1

Descriptive sample statistics

| | N=100 | CBT N = 57 | PDT N = 43 |
|---|---------------|---------------|---------------|
| Gender female | 61% | 60% | 63% |
| Age in years | 34.88 (12.46) | 34.98 (12.53) | 34.74 (12.51) |
| Current relationship | 44% | 35% | 56% |
| High school diploma or higher | 63% | 54% | 74% |
| LSAS anxiety pre (sd) | 38.41 (11.18) | 40.29 (11.99) | 35.93 (9.59) |
| LSAS anxiety post (sd) | 23.15 (10.90) | 22.73 (11.24) | 23.72 (10.56) |
| LSAS avoidance pre (sd) | 31.03 (12.86) | 32.71 (13.43) | 28.83 (11.87) |
| LSAS avoidance post (sd) | 16.74 (11.33) | 16.38 (10.89) | 17.21 (12.01) |
| ECR anxiety pre (sd) | 3.24 (1.15) | 3.09 (1.97) | 3.44 (1.24) |
| ECR anxiety post (sd) | 3.09 (1.10) | 2.76 (1.03) | 3.55 (1.04) |
| ECR avoidance pre (sd) | 2.87 (1.10) | 2.79 (1.03) | 2.97 (1.18) |
| ECR avoidance post (sd) | 2.66 (1.10) | 2.54 (.84) | 2.80 (.91) |
| Number of comorbid diagnoses | .95 (.85) | .77 (.78) | 1.19 (.88) |
| SAD (1 st diagnosis) | 100% | 100% | 100% |
| affective (2 nd diagnosis) | 32% | 32% | 33% |
| other anxiety (2 nd diagnosis) | 12% | 5% | 21% |
| other (2 nd diagnosis) | 22% | 21% | 21% |
| none (2 nd diagnosis) | 34% | 42% | 23% |

Note. CBT = cognitive behavioral therapy; PDT = psychodynamic-orientated therapy, LSAS = Liebowitz Social Anxiety Scale (clinician-administered rating), ECR = Experiences in Close Relationship Inventory, * $p < .05$.

Table 2

Correlations between questionnaire data and synchrony measures.

| | Synchrony | | | Mean TL | | | Range TL | | |
|---------------------|-----------|------|-------|---------|------|-------|----------|------|------|
| | total | pat | ther | total | pat | ther | total | pat | ther |
| ECR anxiety pre | -.06 | -.03 | -.09 | -.02 | -.01 | .09 | .05 | -.13 | .12 |
| ECR avoidance pre | -.09 | -.08 | -.10 | .21* | .13 | .30** | -.08 | -.03 | .04 |
| LSAS anxiety pre | .21* | .15 | .25* | -.12 | -.04 | -.04 | .11 | .12 | .10 |
| LSAS avoidance pre | .20* | .19 | .19 | -.14 | -.08 | -.01 | .12 | .15 | .09 |
| LSAS total pre | .21* | .18 | .22* | -.14 | -.06 | -.03 | .12 | .14 | .10 |
| ECR anxiety post | -.21* | -.16 | -.25* | .13 | .06 | .21* | -.07 | -.12 | -.08 |
| ECR avoidance post | -.09 | -.08 | -.10 | .27** | .21* | .29** | -.04 | -.04 | .00 |
| LSAS anxiety post | .04 | .08 | -.01 | -.06 | .05 | .13 | -.06 | .17 | -.03 |
| LSAS avoidance post | .09 | .12 | .05 | -.05 | .10 | .10 | -.03 | .15 | .05 |
| LSAS total post | .06 | .10 | .02 | -.06 | .07 | .12 | -.05 | .16 | .01 |

Note. LSAS = Liebowitz Social Anxiety Scale (clinician-administered rating), ECR = Experiences in Close Relationship Inventory, Sync = movement synchrony, pat = patient leading, ther = therapist leading, + = indicates significance on a 5% alpha level, ** indicates significance on a 1% alpha level.

Table 3

Statistics of the final model of the step-wise backward regression for the prediction of ECR-ANX and ECR-AVO post.

| | Criterion ECR anxiety post | | | Criterion ECR avoidance post | | |
|-------------------------|----------------------------|-------------------------|------------------------|------------------------------|----------------------------|--|
| | Sync total s8 | Sync ther leading s8 | Mean TL s8 | Mean TL pat lead s8 | Mean TL ther lead s8 | |
| Intercept | $\beta = .00, p = .10$ | $\beta = .00, p = .07$ | $\beta = .00, p = .14$ | $\beta = .00, p = .72$ | $\beta = .00, p = .01$ | |
| ICC random intercept | $r = .25$ | $r = .26$ | removed | removed | removed | |
| ECR anxiety pre | $\beta = .60, p < .01$ | $\beta = .58, p < .01$ | removed | removed | removed | |
| ECR avoidance pre | removed | removed | $\beta = .65, p < .01$ | $\beta = .68, p < .01$ | $\beta = .69, p < .01$ | |
| LSAS anxiety pre | $\beta = .25, p < .01$ | $\beta = .27, p < .01$ | removed | removed | $\beta = .21, p < .01$ | |
| LSAS avoidance pre | removed | removed | $\beta = .22, p < .01$ | $\beta = .20, p < .01$ | removed | |
| Therapy approach | $\beta = -.21, p = .31$ | $\beta = -.25, p = .22$ | removed | removed | $\beta = 1.73, p < .01$ | |
| Sync measure | $\beta = -.66, p = .01$ | $\beta = -.75, p < .01$ | $\beta = .17, p = .02$ | $\beta = .11, p = .09$ | $\beta = .57, p < .01$ | |
| Sync x therapy approach | $\beta = .53^*, p = .02$ | $\beta = .58, p < .01$ | removed | removed | $\beta = -1.79^*, p = .01$ | |

Note. ECR = Experiences in Close Relationships Questionnaire, LSAS = Liebowitz Social Anxiety Scale, pre = session 1, removed = predictor was removed during stepwise backwards regression, pat lead = patient leading, ther = therapist, TL = time lag, β = standardized coefficient.

Table 4

Results of the post hoc multi-group model.

| | Sync total s8 | Sync ther leading s8 | Mean TL ther lead s8 |
|----------------|-------------------------|-------------------------|-------------------------|
| Intercept CBT | $\beta = .39, p = .09$ | $\beta = .39, p = .09$ | $\beta = -.81, p = .08$ |
| Intercept PDT | $\beta = .09, p = .70$ | $\beta = .07, p = .77$ | $\beta = .85, p = .09$ |
| CBT x sync | $\beta = -.34, p = .07$ | $\beta = -.36, p < .05$ | $\beta = 1.07, p < .05$ |
| PDT x sync | $\beta = .18, p = .16$ | $\beta = .19, p = .13$ | $\beta = -.56, p = .27$ |
| CBT x ECR_ANX | $\beta = .90, p < .01$ | $\beta = .88, p < .01$ | removed |
| PDT x ECR_ANX | $\beta = 1.01, p < .01$ | $\beta = 1.00, p < .01$ | removed |
| CBT x ECR_AVO | removed | removed | $\beta = .99, p < .01$ |
| PDT x ECR_AVO | removed | removed | $\beta = 1.04, p < .01$ |
| CBT x LSAS_ANX | $\beta = .43, p = .03$ | $\beta = .49, p = .02$ | $\beta = .39, p < .05$ |
| PDT x LSAS_ANX | $\beta = .47, p = .01$ | $\beta = .49, p = .01$ | $\beta = .40, p < .05$ |

Note. CBT = cognitive behavioral therapy, PDT = psychodynamic oriented therapy, ECR = Experiences in Close Relationships Questionnaire, LSAS = Liebowitz Social Anxiety Scale, TL = time lag, sync = synchrony measure, removed = predictor was removed during stepwise backwards regression, pat lead = patient leading, ther = therapist, TL = time lag, β = standardized coefficient, ANX = anxiety, AVO = avoidance.

Publikation 5

Associations between movement synchrony and outcome in patients with social anxiety disorder: Evidence for treatment specific effects

Uwe Altmann¹, Désirée Schoenherr¹, Jane Paulick², Anne-Katharina Deisenhofer², Brian Schwartz², Julian A. Rubel², Ulrich Stangier³, Wolfgang Lutz^{2,a}, Bernhard Strauss^{1,a}

¹ Jena University Hospital, Jena, Germany

² University Trier, Trier, Germany

³ Johann-Wolfgang-von-Goethe University Frankfurt am Main, Frankfurt am Main, Germany

^a both senior authors contributed equally to this manuscript

Author Note

The study was funded by the German Research Foundation (funding number GZ: STR 306/28-1 and LU 660/8-1). Part of the analyzed videos was recorded in the SOPHO-NET project which was funded by the German Federal Ministry of Education and Research (BMBF, FKZ 01GV0607). We cordially thank U. Willutzki, St. Herpertz, J. Hoyer, P. Joraschky, W. Hiller, E. Leibing, M. Beutel and F. Leichsenring for providing video recordings from the SOPHO-NET project.

Abstract

Background: Studies with heterogeneous samples in naturalistic treatment settings suggest that movement synchrony between therapists and patients correlates with therapeutic success. To reduce heterogeneity based on different psychological disorders, we examined a homogeneous sample of patients with social anxiety disorder and investigated whether movement synchrony in session 3 and 8 is associated with therapy outcome and alliance, and whether this association depends on the therapeutic approach.

Methods: Our sample included $N=267$ patients who were treated with either manual-guided cognitive behavior therapy (CBT), manual-guided psychodynamic therapy (PDT), or naturalistic CBT. The Helping Alliance Questionnaire (HAQ), the Inventory of Interpersonal Problems (IIP) and the Beck-Depression-Inventory (BDI) were used as measures. Body motions were coded with motion energy analysis. Movement synchrony was quantified using time series analysis methods.

Results: Movement synchrony was observed more frequently in both CBT conditions than in PDT. In both CBT groups, more synchrony was predictive of lower IIP scores at the end of therapy. If the patient leads synchrony more often than the therapist, higher IIP and BDI scores were observed at the end of treatment. PDT showed the largest effect size for the synchrony-alliance-association.

Conclusion: Movement synchrony and therapeutic success are associated. The leading when synchronizing body motions might be relevant for future research.

Keywords: social anxiety disorder, therapeutic process, movement synchrony, mimicry, nonverbal communication

In interpersonal interactions we can observe short sequences in which the nonverbal behavior of a person seems to be highly interrelated, coordinated, attuned, aligned, or synchronized with the nonverbal behavior of his/her interaction partner. Such synchronization phenomena are, for example, imitations of emotional facial expressions, convergence of vocal parameters, adaptation of body posture, or synchronous body movements (Altmann, 2011; Delaherche et al., 2012; Feldman, 2007). Meta-analyses suggest a relationship between nonverbal synchrony and higher social bonding, social attitudes and social behavior (Mogan, Fischer, & Bulbulia, 2017; Rennung & Göritz, 2016; Vicaria & Dickens, 2016).

In psychotherapy research, synchronization of patients' and therapists' nonverbal behavior is understood as an indicator of empathy, rapport, and alliance (Koole & Tschacher, 2016) and seems to be especially helpful in therapies of patients with deficits in verbal expression and social behavior (Topel & Lachmann, 2007). However, the conclusions are only based on a few studies which are inconsistent in some aspects. Ramseyer and Tschacher (2011) and Paulick et al. (2017), for example, both examined outpatient cognitive behavioral therapy (CBT) of disorder heterogeneous samples. Whereas Ramseyer and Tschacher (2011) reported a significant relationship between higher movement synchrony early in therapy and a reduction of interpersonal problems, Paulick et al. (2017) found no significant association when using the same motion assessment and multi-level approach. However, when categorizing the patients in "improved", "non-improved/consensual termination", and "non-improved/drop-out", the improved group showed a medium level of movement synchrony, the dropout group showed low levels, and the consensual termination group high levels of synchrony (Paulick et al., 2017). In contrast, symptom improvement and movement synchrony were not correlated in a disorder homogenous sample of patients with schizophrenia (Galbusera, Finn, & Fuchs, 2016). Moreover, in psychosocial counseling for university students vocal synchrony was related to greater distress (Reich, Berman, Dale, & Levitt, 2014). The findings for the

association between nonverbal synchrony and therapeutic alliance are also heterogeneous. Whereas Ramseyer and Tschacher (2011, 2014) reported that more movement synchrony is related to a better therapeutic alliance, Babl, Ramseyer, Westermann, and Caspar (2016) and Reich et al. (2014) observed the opposite regarding movement synchrony and vocal synchrony (with a leading therapist), respectively. Paulick et al. (2017) found no association between movement synchrony and alliance.

The heterogeneous results can be explained in multiple ways. First, differences in therapeutic attitudes and approach-specific interventions might affect the nonverbal behavior of therapist or his/her reactions on the nonverbal behavior of patient's respectively. On the one side psychodynamic therapist might be more neutral and abstinent (Bänninger-Huber & Huber, 2017). On the other side the therapeutic alliance and corresponding nonverbal behaviors are of importance both, in CBT (Castonguay, Constantino, McAleavey, & Goldfried, 2010) and psychodynamic therapy (PDT) (Fosha, 2001). Empirical findings are rare. According to Schoenherr, Paulick, Strauss, et al. (2019), premature therapy termination is associated with infrequent movement synchrony in the 3rd session. Moderator analyses indicated a (marginal significant) effect of the therapeutic approach on synchrony-dropout-association. So far, the impact of the therapeutic approach on synchrony is still unclear.

Second, the decision of who is the leader (*zeitgeber*, drive) of nonverbal synchrony and who is the follower (mimicking, driven) might affect the synchrony-outcome-association. Whereas patients' self-efficacy was related to movement synchrony when patients were leading, the quality of therapeutic alliance was associated with movement synchrony with a leading therapist (Ramseyer & Tschacher, 2011). In another study, vocal synchrony with a leading patient predicted an increase of symptoms, whereas synchrony with a leading therapist and outcome were not significantly associated (Reich et al., 2014). Third, the synchronized behavior modality can be important. As mentioned above, more movement synchrony seems

to be related to improvement (Ramseyer & Tschacher, 2011) whereas synchrony in vocal pitch (under some circumstances) was associated with later deterioration (Reich et al., 2014).

Fourth, the results could be depending on the outcome variable. Reich et al. (2014) found an association between vocal synchrony and the deterioration of depression, whereas Imel et al. (2014) observed a significant relationship between vocal congruence and higher empathy. Also, Paulick et al. (2017), Ramseyer and Tschacher (2011) and Ramseyer and Tschacher (2014) reported that movement synchrony did not correlate with each of the outcomes studied in the studies. Finally, the mental disorder might affect nonverbal synchrony: Research has revealed correlations between movement synchrony and symptom intensity (Kupper, Ramseyer, Hoffmann, & Tschacher, 2015) and has shown disorder specific patterns of synchrony development over the course of the treatment (Galbusera et al., 2016; Paulick et al., 2018). As a consequence, averaging results over different diagnostic groups may result in spurious associations reflecting relevant subgroups within a diagnostically heterogeneous sample. Focusing on a specific disorder could foster our understanding of synchrony-outcome associations on a more fine-grained level.

Social anxiety disorder (SAD) is particularly suitable to study these kinds of associations: Experiments revealed that socially anxious individuals have problems to perceive nonverbal synchrony (Vrijssen, Lange, Dotsch, Wigboldus, & Rinck, 2010) – possibly due to higher levels of self-focused attention and lower attention in situations. They showed less unintentional movement synchrony which could be an expression of a decreased desire to affiliate (Abbott, Kocovski, & Obhi, 2018; Vrijssen, Lange, Becker, & Rinck, 2010) or of uncomfortable feelings (Likowski et al., 2011). Healthy individuals reported lower social anxiety if their body posture was imitated by their interaction partner (Gueguen, 2011). Regarding facial synchrony, it was observed that socially anxious individuals reacted more often with a polite smile than persons with low social anxiety, probably because smiling

increases the acceptance of the interaction partner and avoids confrontations (Dijk, Fischer, Morina, van Eeuwijk, & van Kleef, 2018).

SAD is one of the most prevalent psychological disorders (lifetime prevalence of 13.3% and 12-month-prevalence of 7%, Kessler, Petukhova, Sampson, Zaslavsky, and Wittchen (2012)) and is related to excessive fear and avoidance of social situations (e.g., getting in touch with strangers), low self-esteem and fear to be evaluated (American Psychiatric Association, 2013). In social situations, the individuals fear that the interlocutor is aware of the physical symptoms of anxiety and makes a negative judgment. Since anxiety symptoms are perceived as unpleasant, a distinct avoidance and safety behavior is developed (Stravynski, Kyparissis, Amado, Hoffmann, & DiBartolo, 2010). Moreover, persons with SAD perceive and interpret social information more negatively and have dysfunctional interpersonal pattern which might result in problems in interpersonal relationships (Alden & Taylor, 2004).

Research questions

The current state of research suggests focusing on a disorder-specific sample to exclude heterogeneity resulting from different psychopathologies. Due to the high prevalence of SAD and the link between synchrony and social anxiety, we examined patients with SAD. In accordance with other studies, we considered body movements assessed with computer vision methods and quantified movement synchrony using time series analysis methods. Based upon the results summarized above, we hypothesized that frequent movement synchrony in an early therapy session predicts therapeutic outcome (less interpersonal problems and better therapeutic alliance). Due to the lack of empirical studies, we further explored whether leading when synchronizing body motions is related to the therapeutic outcome and whether the therapeutic approach affects the frequency of synchrony and the synchrony-outcome-association.

Methods

Background of the study

This multi-site study has a quasi-experimental three-armed longitudinal design. One study center (Jena, Germany) provided patients from the SOPHO-NET trial (Leichsenring et al., 2013) which compared manual-guided CBT or manual -guided PDT in a randomized controlled study. In the other study center (Trier, Germany), patients with SAD were treated with CBT under naturalistic conditions. Patients were not randomized regarding manual-guided and naturalistic therapy, since this study comprises a secondary data analysis. Furthermore, we did not aim at a comparison of therapy effectiveness. There was no blinding regarding the examination condition. Patients and therapists gave informed consent for video recordings but did not know the study aim of examining nonverbal synchrony. The study was approved by the local ethics committees.

Recruiting, inclusion criteria, and sample

In the SOPHO-NET trial, a total of $N=1450$ patients were screened for SAD using the Structured Clinical Interview for Axis I Disorders (SCID-I) (Wittchen, Wunderlich, Gruschwitz, & Zaudig, 1997). Inclusion criteria were a primary diagnosis of SAD, and a score in the Liebowitz Social Anxiety Scale of >30 (Leichsenring et al., 2013). Exclusion criteria were comorbid psychotic or acute substance-related disorder, cluster A and B personality disorder, organic mental disorders, prominent risk of self-harm, severe medical conditions, and concurrent psychotherapeutic or psychopharmacological treatment. The included $N=495$ patients were randomized to three conditions: PDT, CBT, and a waitlist group. During waiting time $N=21$ patients dropped out. The remaining waitlist persons were also randomized to CBT or PDT.

For this study, in a next step, $N=209$ patients were excluded because therapy sessions were not video recorded ($N=144$) or the videos had a sub-optimal quality ($N=65$). In the study center Trier, $N=3200$ patients were screened with SCID-I. $N=134$ patients with SAD as primary or secondary diagnosis and the absence of a comorbid psychosis and acute substance

dependency were selected. Due to therapist-transfer ($N=2$) and sub-optimal video quality ($N=21$), $N=23$ patients were excluded. In total, $N=154$ patients of the SOPHO-NET trial and $N=111$ patients from study center Trier were included in the final sample (see Figure 1).

- Figure 1 -

Treatment and therapists

The manual-guided CBT (Stangier, Clark, & Ehlers, 2006) focused on a change of maintaining factors identified in Clark and Wells's model of SAD (Clark & Wells, 1995). Safety behaviors and self-focused attention are of particular importance. Typical interventions were: establishing a personal version of the model of SAD, role-play-based behavioral experiments, practicing external focus of attention, and restructuring of distorted self-image by video feedback or behavioral experiments.

The manual-guided PDT (Leichsenring, Beutel, & Leibing, 2007) was an adaptation of Luborsky's supportive-expressive treatment approach (SET; Luborsky, 1984) for SAD. The therapy encompasses supportive interventions (e.g., establishment of the therapeutic alliance) and expressive interventions (e.g., clarification that symptoms are caused by and expressions of a core conflictual relationship theme (CCRT)). The CCRT with the components wish, response to others, and response to self was clarified, addressed and worked through with regard to past and present relationships of the patient.

The naturalistic CBT can be described as an integrative CBT approach including interpersonal and emotion-focused elements. Therapists were not supposed to follow a certain manual. Individual interventions were adapted on patient characteristics and therapeutic progress. Typical interventions were psycho-education, relaxation training, cognitive restructuring, and situational exposure (Lutz, Schiefele, Wucherpfennig, Rubel, & Stulz, 2016). Therapists received psychometric feedback on patient's progress after each session.

All therapies comprised five preparatory sessions (for diagnostic and administrative issues) which are common for outpatient psychotherapy in Germany. Furthermore, all therapy sessions were provided as individual 50-minute sessions usually on a weekly basis. Both

manual-guided therapies were short-term therapies, whereas naturalistic CBT could be short-term or long-term therapy depending on indication and therapeutic process. The treatment length was on average $M=26.0$ session ($SD=4.9$) for manual-guided CBT, $M=27.1$ sessions ($SD=7.3$) for manual-guided PDT, and $M=42.7$ sessions ($SD=18.8$) for naturalistic CBT.

All $N=119$ therapists had a university degree as clinical psychologist or physician. Their psychotherapeutic training was completed or in an advanced stage. The average age was 31.5 years and 78% were female. A therapist provided treatments only in one therapy arm (1-10 patients per therapist, median=2). In naturalistic CBT, therapists were supervised after each 4th session by a senior therapist. In both manual-guided therapies, continuous on-site and cross-site supervisions were applied to assure adherence to the manuals (Leichsenring et al., 2013).

Instruments

Socio-demographic characteristics of patients, such as gender, age, diagnosis, school education, and current relationship status were assessed at the beginning of therapy. Our primary outcome measure was the German version of the Inventory of Interpersonal Problems (IIP-32; Horowitz, Strauß, Thomas, & Kordy, 2016). We used the total score (average of all 32 items) that reflects the degree of distress associated with interpersonal problems. The Beck-Depression-Inventory (BDI; Beck, Ward, Mendelson, Mock, & Erbaugh, 1961) assessed the severity of depression. In the study centers, different versions of the BDI were applied. Therefore, we constructed a new scale based on 16 “overlapping” items. The new scale correlated high with total scores (with BDI-I $r=.982^{**}$, with BDI-II $r=.984^{**}$). With the Helping Alliance Questionnaire (HAQ-I; Bassler, Potratz, & Krauthauser, 1995; HAQ-II; Luborsky et al., 1996) patients’ ratings of relationship satisfaction were assessed. Due to the fact that different versions were applied in the study centers, we computed a new score based on the five overlapping items. The new score correlated with the original scores with $r=.75-.81$. Cronbach’s α of all scales was $>.8$. Further questionnaires were used only in one of the two centers and could therefore not be considered in the presented analysis.

Video recording, motion and synchronization assessment

Videos were selected from therapy sessions 3 (S3), 8 (S8), 20 (S20) and 29 (S29) because these time points were close to the measurement time points of the SOPHO-NET study. If a video was either not recorded or had a sub-optimal video quality, a video of one session before or after the concerned time point was used. Exclusion criteria for a video were: sub-optimal light conditions (e.g., light reflections of window), a person was not or only partly visible, a person covered the other within the video frame, a person stood up, a third person was in the room, and session reflected assessment of questionnaires, role plays, and family constellation. Up to four videos were chosen for each patient (in total $N=646$ videos).

The therapy videos of the SOPHO-NET sample were recorded using a single camcorder (camera angle: $\frac{3}{4}$ view to profile). In the study center Trier, a multi-camera system provided split-screen videos. All video recordings were converted into AVI format (size=640x480, frame rate=25fps, bit rate=2000Kbits/second) using Any Video Converter 3.0.

The body motions were coded separately for patient and therapist using motion energy analysis (MEA; Altmann, 2013) (MATLAB© scripts available at <https://github.com/10101-00001/MEA>). MEA is a frame differencing method which measures the intensity of body movements frame by frame by counting the number of pixels whose gray intensity changed noticeably (see Figure 2). According to Altmann (2013), we used the 99% quantile of background ROI as the cut-off value which was 12 in our data. Tests with one video and MEA implementations of Ramseyer and Tschacher (2011) and Paxton and Dale (2012) showed a high agreement between time series of different implementations ($r(\text{Ramseyer, Paxton}) = .895$, $r(\text{Ramseyer, Altmann}) = .996$, and $r(\text{Paxton, Altmann}) = .894$ with $p < .001$). To distinguish motions of patient and therapist, we defined so-called regions of interest (ROI) which covered the individual upper body in the video frame. Additionally, we applied two ROI in the background with the size of 10x10 pixels which were used subsequently to filter video errors. To reduce the time exposure for coding, we considered the first 15min of the therapy session (after clarifying of administrative issues or filling in questionnaires) because Ramseyer and

Tschacher (2011) and Paulick et al. (2017) reported a high correlation of synchrony between the first 15min and the entire 50min session.

- Figure 2 -

We obtained two person time series with the length $15\text{min} * 60\text{sec}/\text{min} * 25\text{frames}/\text{sec} = 22500$ frames for each video. Each value was divided by the size of the individual ROI and multiplied with 100. As a result, the values range from 0 (no motion) to 100 (maximal measurable motion; the entire ROI was “activated”). If the motion energy of a background ROI was larger than 5 (indicator for a video error such as change of light conditions) or if the motion energy of one person changed for more than 15 points within three frames (extreme fast movement which is impossible for humans), then the motion energy value of the corresponding time point was set to missing and then linearly interpolated. Moreover, a moving median with a bandwidth of five frames was applied to smooth time series parts with high frequent variations.

Next, we used a windowed cross-lagged correlation (WCLC) (Altmann, 2011, 2013) (download: https://github.com/10101-00001/sync_ident) to assess the local interrelatedness of each time series pair. The algorithm correlates a window of patients’ time series (reference window A) and a lagged window of therapists’ time series (window B) (see Figure 3 top right). This is repeated for each possible position of a reference window and each possible time-lag including time-lag=0. Non-significant correlation coefficients were set to zero ($H_0: r=0$, significance level $\alpha=.001$). All coefficients were squared ($R^2=r*r$) because R^2 is the proportion of explained variance and, on this way, all values were positive and ranged from zero (no linear relationship between window A and window B) to one (window A can perfectly predict window B). The output of the WCLC algorithm is a R^2 -matrix (see Figure 3 middle left). An element with the “coordinates” (k, l) is a squared correlation coefficient referring to window A at frame (position) k and window B at frame $k+l$. We applied the pre-processing steps and parameter for WCLC which maximized the identification rate of synchrony and minimized the false positives in the validation study of Schoenherr, Paulick, Deisenhofer, et al. (2019): log-transformation,

window width of WCLC =125frames (respectively 5sec), no additional smoothing, and an R^2 -cut-off of .25 to control for random occurring synchrony. In accordance with Ramseyer and Tschacher (2011), Paulick et al. (2017) and Schoenherr, Paulick, Strauss, et al. (2019) a maximum time lag of ± 125 frames (equals ± 5 sec) was used.

- Figure 3 -

In the next step, the peak-pick-algorithm developed by Altmann (2011, 2013) (download: https://github.com/10101-00001/sync_ident) was applied on the R^2 -matrix. The algorithm locates synchronization intervals based on correlation coefficient maxima which are neighbored and have an equal time-lag (see Figure 3 bottom left). The output is a list of synchronization intervals (LOSI) incorporating beginning, end, time-lag, and average R^2 of each identified synchronization interval. According to Schoenherr, Paulick, Deisenhofer, et al. (2019), we excluded all synchronization intervals with $R^2 < .25$ to reduce bias from spurious cross-correlations.

Based on the LOSI of a video, the frequency of movement synchrony was computed. According to Altmann (2011, 2013), the sum of the duration of all synchronization intervals is divided by the length of the considered time series and multiplied with 100. The values can range from 0 (during the entire interaction no movement synchrony occurred) to 100 (patient and therapist were synchronized the entire time). Synchrony with leading of patient (resp. therapist imitates the patient) and synchrony with leading of therapist are computed using the time-lag information in the LOSI. The leading variable analyzed later is the difference of these two variables (negative values = therapist leads more often when synchronizing, 0 = frequency of synchrony with leading patient is equal to synchrony with leading therapist, positive values = patient leads more often when synchronizing / therapist imitates the patient more often). Frequency of movement synchrony and leading correlated weakly or insignificantly (S3: $r = .149, p = .015$; S8: $r = .095, p = .123, N = 267$; S20: $r = .108, p = .173, N = 160$; S29: $r = .058, p = .517, N = 127$). The former correlated significantly with the movement frequency of the patient ($r = .498-.559, p < .001$) and the therapist ($r = .608-.670, p < .001$).

It should be noted that synchrony assessed with our method is based on the assumption that a dyadic interaction is a dynamic system showing alternate phases of synchronized (high local interrelatedness of both person time series) and non-synchronized nonverbal behavior (less or even no interrelatedness) (Altmann, 2011, 2013; García & Di Paolo, 2018). In contrast to that, Ramseyer and Tschacher (2011) operationalized movement synchrony as the average of all $k \times l$ cross-correlation coefficients which can be interpreted as average local interrelatedness of both time series including sequences of low interrelatedness (respectively no-sync intervals). However, the frequency measure and the interrelatedness measure had a moderate concordance ($r=.79$, $p<.05$ when using equal parameter settings and $r=.55$, $p<.05$ when the recommended parameter settings by authors of algorithms are applied, see Schoenherr et al. (2018)).

Missing data

The proportion of missing data ranged from 0% to 35% (mean: 19%). Regarding patients' age and gender, data were complete. Causes for missing data were mainly premature therapy terminations and low video quality. According to Little's MCAR test ($\chi^2=539.21$, $df=511$, $p=.187$), missing values were assumed as completely at random. For imputation of relationship status, high school diploma, pre-treatment severity (IIP, BDI), impairment at end of therapy (IIP, BDI), HAQ at S20, synchronization frequency and leading at S3 and S8, the R package *missForest* (Stekhoven, 2011) was used (parameter settings: maximum number of 20 iterations, 1000 trees, seed = 47). The resulting fit indices were normalized root-mean-square error (NRMSE) =0.196 and proportion of falsely classified (PFC) =0.048. Due to a large proportion of missing values (>35%), synchrony at S20 and S29 were not imputed.

Matching of treatment groups

Since patients were not randomized to manual-guided and naturalistic therapy, we applied propensity score matching (PSM). Propensity scores with the treatment condition as binary dependent variable (manual guided=1; naturalistic=0) and demographic variables of patient (gender, age, high school diploma, current relationship) and therapist (gender, age),

initial IIP and initial BDI as predictors were computed. Then a full-matching was performed with R package *MatchIT* (Ho, Imai, King, & Stuart, 2011). No patient was excluded. The resulting propensity scores were used as regression weights. Since weights of patients from the study center Jena were set to one, results of analysis refer to a sample with the characteristics of the manual-guided therapies.

Statistical analysis

First, the entire sample and the three treatment groups were described using descriptive statistics. We compared the treatment groups regarding patient characteristic (e.g., gender, age, initial depression etc.) using χ^2 -test (for categorical variables) and ANOVA (for metric variables).

For the examination of the time course of synchrony (resp. leading) depending on the stage of treatment and therapy group, we estimated raw data averages and regression adjusted averages. For the latter, we used a growth curve model with two random effects, one for multiple patients per therapist and one for multiple measurements per patient. The dependent variable was the frequency of synchrony (resp. leading). Independent variables were initial IIP, initial BDI, treatment group, time point (session number) and the squared time point. Furthermore, we included interaction terms between time on the one side and initial IIP, initial BDI, and group on the other side as well as interaction terms of squared time on the one side and initial IIP, initial BDI, and group on the other side. Based on the coefficients of the growth curve model, we estimated adjusted averages of synchronization frequency. With “adjusted” means, we controlled for initial IIP and BDI, treatment condition and (exact) measurement time point. The adjusted averages were compared using t-tests for independent samples. Effect sizes were interpreted according to Cohen (1988) (small .2 to .5, moderate .5 to .8 and large >.8).

To examine the synchrony-outcome-association, we applied hierarchical linear models (HLM). In model 1, the therapeutic approach was a random effect; thus we got an average synchrony-outcome-association over all three treatments. Model 2 was a multi-group model in

which the synchrony-outcome-association was estimated separately for each treatment arm. Both models included a random effect for the nesting of multiple patients per therapist. The IIP scores assessed at the end of therapy (respectively BDI post and HAQ at S20) was the dependent variable. Independent variables were initial IIP-, initial BDI scores, synchrony and leading in the early therapy stage. In model 2, the therapy group was a further fixed effect. We included interaction terms between the group and every other covariate. We analyzed synchrony at S3 and S8 as predictors separately.

Results

Descriptive statistics

Table 1 displays descriptive statistics. Comparing the treatment groups regarding the baseline variables, we found different distributions for patients' and therapists' gender, proportion of high school degree, current relationship status, age of therapists, and initial BDI.

- Table 1 -

Time curve of synchronization frequency

Averages and standard deviations of synchrony and leading depending on treatment group and measurement time point are shown in Table 2. For each time point, we found that patients and therapists synchronize their movements significantly more often in manual-guided CBT than in manual-guided PDT (S3: $\text{diff}=9.7$, $SE=1.9$, $p<.001$, $d=0.87$; S8: $\text{diff}=8.9$, $SE=1.7$, $p<.001$, $d=0.90$; S20: $\text{diff}=9.3$, $SE=1.8$, $p<.001$, $d=0.86$; S29: $\text{diff}=11.6$, $SE=2.2$, $p<.001$, $d=0.87$). Similarly, in naturalistic CBT dyads synchronized more often than in manual-guided PDT except for S29 (S3: $\text{diff}=10.1$, $SE=1.8$, $p<.001$, $d=0.93$; S8: $\text{diff}=11.5$, $SE=1.7$, $p<.001$, $d=1.13$; S20: $\text{diff}=9.6$, $SE=1.8$, $p<.001$, $d=0.86$). Only at S29 the average synchronization frequency differs significantly between manual-guided CBT and naturalistic CBT (S29: $\text{diff}=8.2$, $SE=1.9$, $p<.001$, $d=0.60$). The applied Bonferroni corrected level of significance was

$\alpha=.05/12=.004$. For leading, we did not find any significant differences between adjusted group means. It should be noted that CB therapist showed more significantly movements than PD therapist (see Appendix, Table A.1).

- Table 2 -

Prediction of outcome

Next, we examined the association between synchrony (respectively leading) and pre-to-post treatment improvement when controlling for intake scores of the IIP and BDI. The statistics are displayed in Table 3 and Figure 4. In model 1, with treatment condition as random effect, a more frequent movement synchrony at S3 and S8 predicted higher therapeutic alliance in S20. The multi group model (model 2) revealed that this association is mainly present in naturalistic CBT and manual-guided PDT. Furthermore, more leading by patient in S8 correlated with lower alliance. On an average over all three approaches (model 1), more movement synchrony and less leading of the patient at S3 and S8 significantly predicted lower interpersonal problems at the end of therapy. The multi-group model (model 2) revealed that these associations were mainly present in manual guided CBT and naturalistic CBT.

The severity of depression (BDI) could not be predicted by the frequency of movement synchrony (neither the 3rd nor the 8th session in model 1 and model 2). However, more leading by the patient (respectively more imitation by the therapist at S3 and S8) predicted higher BDI scores at the end of therapy. In the multi-group model, this association was only found for manual-guided CBT.

- Table 3 - and - Figure 4 -

Please note that there were only minimal differences regarding the regression coefficients and *p*-values when not applying the regression weights (not reported in detail due to space limitations). Furthermore, we tested an HLM with gender and age of therapist. Both

were non-significant outcome predictors and had no impact on synchrony-outcome-association in all three therapy arms.

Next, we examined whether the therapy condition moderates the association between synchrony (respectively leading) and outcome. Therefore, we considered the regression coefficients of the multi-group models mentioned above. The association between HAQ at S20 and synchrony at S3 was stronger in manual-guided PDT than in manual-guided CBT ($b_{\text{man.PDT}} - b_{\text{man.CBT}} = 0.032 - 0.004 = 0.028$, $SE = 0.011$, $p = .018$). The same can be stated for naturalistic CBT and manual-guided CBT ($b_{\text{nat.CBT}} - b_{\text{man.CBT}} = 0.020 - (-0.002) = 0.022$, $SE = 0.010$, $p = .025$) regarding the association between HAQ at S20 and synchrony at S8. Furthermore, the association between synchrony at S8 and IIP post of manual-guided CBT was significantly stronger than in manual-guided PDT ($b_{\text{nat.CBT}} - b_{\text{man.CBT}} = 0.011 - (-0.019) = 0.030$, $SE = 0.012$, $p = .012$). The association between patients' leading frequency at S8 and IIP post was different in manual-guided CBT than in naturalistic CBT ($b_{\text{nat.CBT}} - b_{\text{man.CBT}} = -0.012 - 0.054 = -0.066$, $SE = 0.029$, $p = .022$). Furthermore, there was a moderation effect on the association between BDI at end of therapy and leading at S3. In manual-guided CBT, this association was stronger than in naturalistic CBT ($b_{\text{nat.CBT}} - b_{\text{man.CBT}} = -0.007 - 0.806 = -0.813$, $SE = 0.295$, $p = .006$). For depression as outcome and synchrony and leading at S8, we did not find significant moderator effects. However, many comparisons of regression coefficients reached a trend level of significance, for example, regarding the association between IIP post and synchrony at S3: manual-guided CBT versus naturalistic CBT ($b_{\text{nat.CBT}} - b_{\text{man.CBT}} = -0.004 - (-0.017) = 0.013$, $SE = 0.007$, $p = .076$) as well as manual-guided CBT versus manual-guided PDT ($b_{\text{man.PDT}} - b_{\text{man.CBT}} = 0.002 - (-0.017) = 0.019$, $SE = 0.010$, $p = .065$) or regarding the association between leading at S3 and BDI post: manual-guided PDT versus manual-guided CBT ($b_{\text{man.PDT}} - b_{\text{man.CBT}} = 0.269 - 0.806 = -0.537$, $SE = 0.029$, $p = .069$).

Discussion

The present study examined the frequency of movement synchrony in patient-therapist interactions as early indicator of therapeutic outcome. In contrast to recent studies, we investigated a disorder-specific sample of patients with SAD and the therapeutic approach as a potential moderator of the synchrony-outcome-association. According to our hypothesis, frequent movement synchrony in an early phase of therapy predicted (on average over all therapeutic approaches) a better therapeutic alliance at session 20 and reduced interpersonal problems at end of therapy. These results are in accordance with the clinical study of Ramseyer and Tschacher (2011) and the meta-analyses regarding the relationship between nonverbal synchrony, social bonding and social cognition (Mogan et al., 2017; Rennung & Göritz, 2016; Vicaria & Dickens, 2016). However, the improvement of depression was not predicted by movement synchrony. That synchrony correlates not with each outcome measure was also reported by Paulick et al. (2017), Ramseyer and Tschacher (2011) and Ramseyer and Tschacher (2014). That might mean that movement synchrony is more related to change of relationship variables (e.g. alliance, interpersonal problems, and possibly attachment) than of variables refereeing more to the affective spectrum (e.g. depression).

Moreover, our explorative analyses showed that movement synchrony is an early indicator of therapeutic success but not under all conditions: The more the patient leads synchrony (respectively the more the therapist is imitating) in session 8, the lower are their ratings of the therapeutic alliance, and the higher were the scores for interpersonal problems and depression. This result confirms the findings of Reich et al. (2014) showing a correlation between vocal synchrony led by the patient and higher symptoms at end of therapy. However, because movement synchrony led by the patient in session 3 was associated with a reduction of dropout rate (Schoenherr, Paulick, Strauss, et al., 2019), it can assumed that the therapeutic impact of leading differs depending on stage of therapy. In first sessions leading by the patient

seems to support the establishment of therapeutic alliance. In the subsequent stage, synchrony with leading by the therapist could reflect patient's trust and establishment of functional behavioral responses.

Taken together, synchronizing nonverbal behavior might help in psychotherapy to express agreement and understanding (Keller & Tschacher, 2007; Ramseyer & Tschacher, 2014), to reduce conflicts within the dyad (Bänninger-Huber & Huber, 2017) and to build, manage (Ramseyer & Tschacher, 2011; Schoenherr, Paulick, Strauss, et al., 2019) and (in case of ruptures) repair the therapeutic relationship (e.g. Safran, Muran, & Eubanks-Carter, 2011) so that patients' interpersonal pattern and emotion regulation capacities can improve (Koole & Tschacher, 2016). At the same time, therapeutic failure could emerge, if therapists imitate their patients in the middle stage of therapy more often than vice versa. Synchrony then could lead to involvement in patient's maladaptive emotional and interpersonal pattern (Anstadt, Merten, Ullrich, & Krause, 1997). As it has been shown within romantic relationships, synchrony could amplify interpersonal conflicts in combination with an ineffective emotion regulation (Butler, 2015). Accordingly, one challenge of psychotherapy is to be empathic and to get not involved in maladaptive interpersonal patterns at the same time.

Further explorative analyses provided evidence that the therapeutic approach moderated the synchrony-outcome-association. PDT showed the largest effect size for the synchrony-alliance-association. In another analysis Schoenherr, Paulick, Strauss, et al. (2019) found (on a trend level) that the association between movement synchrony in session 3 and premature therapy termination was stronger in PDT than in CBT, which might indicate a strong association between synchrony and alliance in PDT. In contrast, movement synchrony predicted an improvement of interpersonal problems in both CBT conditions. In PDT, we observed only trends regarding the prediction of depression at end of therapy and synchrony at S3 and leading at S8. The corresponding standardized regression coefficients were similar to

the significant coefficients in CBT. May be there is a power problem so that the missing significance should not be overvalued: The sample size of PDT group was only $N=54$ patients which is about half of the CBT sample size ($N_{\text{nat.CBT}}=111$, $N_{\text{man.CBT}}=102$). Furthermore, it should be considered that PD therapists showed less movement (see appendix) and less movement synchrony of patient and therapist in S3, S8, and S20 of the manual-guided PDT than in both CBT groups. The reason for that is still unclear. Possible explanations are, for example, a neutral and abstinent therapeutic attitude (Bänninger-Huber & Huber, 2017) or the impact of specific interventions of manual-guided PDT (e.g. working through a CCRT) as opposed to highly structured very structured treatment of the examined manual-guided CBT. An alternative explanation could be that in our sample the PD therapists were on average older and more frequently males which could affect the frequency (Schoenherr, Paulick, Strauss, et al., 2019) and perception of synchrony (Kraft-Todd et al., 2017). However, the application of propensity score matching (which includes therapist age and gender) should have reduced this bias. Interestingly, in naturalistic CBT synchrony decreased from session 20 to 29 so that naturalistic CBT and manual-guided PDT had similar synchrony averages at session 29 respectively that the averages of both CBT's differed significantly. A reason can be that all manual-guided therapies were short-term therapies whereas the naturalistic CBT could be short-term or long-term therapy. In other words, we might have compared different stages of the treatments.

For the development of theories there is a high need for further clinical studies. Future research should clarify, for example, the impact of therapeutic approach in relation to characteristics of therapists such as competences, empathic ability or intra-/extraversion. Furthermore, so far studies with clinical samples examined only unintentional synchrony. Possibly, intentional synchrony leads to other effects on therapeutic outcome. The experiment of Fitzpatrick et al. (2017) showed, for example, that cooperation was correlated with

spontaneous social motor synchrony, but not with intentional synchrony. Furthermore, the question should be addressed how much synchrony is needed. It can be assumed that the frequency of synchrony is different before, during and after critical therapy events (Duran & Fusaroli, 2017; Hart et al., 2016). Possibly, the time point of synchrony will affect the therapeutic alliance (Bänninger-Huber & Huber, 2017; Kodama, Hori, Tanaka, & Matsui, 2018).

Limitations

There are several limitations: First, patients were only partly randomized to the treatment conditions because the present study was a secondary data analysis. To limit the corresponding bias, we applied matching and regression adjustment. Second, in naturalistic CBT, we included also patients with SAD as secondary diagnosis to get a typical sample of psychotherapy patients under naturalistic conditions. Third, in both centers were different outcome measures assessed because this study was a secondary data analysis. Fourth, we considered only the first 15min of therapy sessions and used aggregated measures of movement (no specific motions) and synchrony (synchronization frequency) whereby the time point of synchronization within a session and the content of the interaction/conversation was not considered.

Strengths of this study comprise the large, disorder-specific sample ($N=267$ dyads), consideration of different therapeutic approaches, a highly reliable and objective assessment of motions, consideration of potential confounders and the nested data structure by applying HLM. Furthermore, the presented study is based on an extensive methodical preliminary work (Altmann, 2011, 2013; Schoenherr, Paulick, Deisenhofer, et al., 2019; Schoenherr et al., 2018).

Conclusions

Depending on the therapeutic approach, movement synchrony in an early therapy session predicts different measures of therapeutic success. However, when in the middle stage

of therapy the therapist imitates patient's movement behavior more often than vice versa, he or she might be more involved in patients' maladaptive interpersonal pattern which on the other side could hamper the change process of interpersonal problems.

References

- Abbott, K. A., Kocovski, N. L., & Obhi, S. S. (2018). Impact of Social Anxiety on Behavioral Mimicry During a Social Interaction With a Confederate. *Journal of Social and Clinical Psychology, 37*(1), 22-43. doi:10.1521/jscp.2018.37.1.22
- Alden, L. E., & Taylor, C. T. (2004). Interpersonal processes in social phobia. *Clinical psychology review, 24*(7), 857-882. doi:10.1016/j.cpr.2004.07.006
- Altmann, U. (2011). Investigation of movement synchrony using windowed cross-lagged regression. In A. Esposito, A. Vinciarelli, K. Vicsi, C. Pelachaud, & A. Nijholt (Eds.), *Analysis of Verbal and Nonverbal Communication and Enactment: The Processing Issue* (pp. 344-354). Berlin: Springer.
- Altmann, U. (2013). *Synchronisation nonverbalen Verhaltens [Synchronization of nonverbal behavior]*. Berlin: Springer.
- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders (5th ed.)*. Washington, DC: Author.
- Anstadt, T., Merten, J., Ullrich, B., & Krause, R. (1997). Affective dyadic behavior, core conflictual relationship themes, and success of treatment. *Psychotherapy Research, 7*(4), 397-417.
- Babl, A., Ramseyer, F., Westermann, S., & Caspar, F. (2016). *Do as I do, not as I say - Automatically detected nonverbal synchrony between patients and therapists in psychotherapeutic dyads assessed with Microsoft Kinect*. Paper presented at the 47th Annual Meeting of the Society for Psychotherapy Research, June 2016, Jerusalem.
- Bänninger-Huber, E., & Huber, E. (2017). Nonverbal Processes in Psychotherapeutic Interaction. *Journal of Psychology and Psychotherapie Research, 4* (1), 1-8. doi:10.12974/2313-1047.2017.04.01.1
- Bassler, M., Potratz, B., & Krauthauser, H. (1995). Der "Helping Alliance Questionnaire" (HAQ) von Luborsky. Möglichkeiten zur Evaluation des therapeutischen Prozesses von stationärer Psychotherapie. *Psychotherapeut, 40*(1), 23-32.

- Beck, A. T., Ward, C. H., Mendelson, M., Mock, J., & Erbaugh, J. (1961). An inventory for measuring depression. *Archives of general psychiatry*, 4(6), 561-571. doi:10.1001/archpsyc.1961.01710120031004
- Butler, E. A. (2015). Interpersonal affect dynamics: It takes two (and time) to tango. *Emotion Review*, 7(4), 336-341. doi:10.1177/1754073915590622
- Castonguay, L. G., Constantino, M. J., McAleavey, A. A., & Goldfried, M. R. (2010). The therapeutic alliance in cognitive-behavioral therapy. In J. C. Muran & J. P. Barber (Eds.), *The therapeutic alliance. An evidence-based guide to practice* (pp. 150-171). New York, NY: The Guilford Press.
- Clark, D. M., & Wells, A. (1995). A cognitive model of social phobia. In R. G. Heimberg, M. R. Liebowitz, D. A. Hope, & F. R. Schneier (Eds.), *Social phobia: Diagnosis, assessment, and treatment* (pp. 69-93). New York, NY: Guilford Press.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2 ed.). Hillsdale: Erlbaum Associates.
- Delaherche, E., Chetouani, M., Mahdhaoui, A., Saint-Georges, C., Viaux, S., & Cohen, D. (2012). Interpersonal synchrony: A survey of evaluation methods across disciplines. *IEEE Transactions on Affective Computing*, 3(3), 349-365. doi:10.1109/T-AFFC.2012.12
- Dijk, C., Fischer, A. H., Morina, N., van Eeuwijk, C., & van Kleef, G. A. (2018). Effects of Social Anxiety on Emotional Mimicry and Contagion: Feeling Negative, but Smiling Politely. *Journal of Nonverbal Behavior*, 42(1), 81-99. doi:10.1007/s10919-017-0266-z
- Duran, N. D., & Fusaroli, R. (2017). Conversing with a devil's advocate: Interpersonal coordination in deception and disagreement. *PLoS ONE*, 12(6), e0178140. doi:10.1371/journal.pone.0178140
- Feldman, R. (2007). Parent-infant synchrony and the construction of shared timing, physiological precursors, developmental outcomes, and risk conditions. *Journal of*

-
- Child psychology and Psychiatry*, 48(3-4), 329-354. doi:10.1111/j.1469-7610.2006.01701.x
- Fitzpatrick, P., Romero, V., Amaral, J. L., Duncan, A., Barnard, H., Richardson, M. J., & Schmidt, R. C. (2017). Social Motor Synchronization: Insights for Understanding Social Behavior in Autism. *Journal of Autism and Developmental Disorders*, 47(7), 2092-2107. doi:10.1007/s10803-017-3124-2
- Fosha, D. (2001). The dyadic regulation of affect. *Journal of clinical psychology*, 57(2), 227-242. doi:10.1002/1097-4679(200102)57:2<227::AID-JCLP8>3.0.CO;2-1
- Galbusera, L., Finn, M. T., & Fuchs, T. (2016). Interactional synchrony and negative symptoms: An outcome study of body-oriented psychotherapy for schizophrenia. *Psychotherapy Research*, 1-13. doi:10.1080/10503307.2016.1216624
- García, E., & Di Paolo, E. A. (2018). Embodied Coordination and Psychotherapeutic Outcome: Beyond Direct Mappings. *Frontiers in Psychology*, 9(1257). doi:10.3389/fpsyg.2018.01257
- Gueguen, N. (2011). The mimicker is a mirror of myself: Impact of mimicking on self-consciousness and social anxiety. *Social Behavior and Personality: an international journal*, 39(6), 725-728. doi:10.2224/sbp.2011.39.6.725
- Hart, Y., Czerniak, E., Karnieli-Miller, O., Mayo, A. E., Ziv, A., Biegon, A., . . . Alon, U. (2016). Automated video analysis of non-verbal communication in a medical setting. *Frontiers in Psychology*, 7(1130). doi:10.3389/fpsyg.2016.01130
- Ho, D. E., Imai, K., King, G., & Stuart, E. A. (2011). MatchIt: nonparametric preprocessing for parametric causal inference. *Journal of Statistical Software*, 42(8), 1-28.
- Horowitz, L. M., Strauß, B., Thomas, A., & Kordy, H. (2016). *Inventar zur Erfassung interpersonaler Probleme - deutsche Version* (3 ed.). Göttingen: hogrefe.
- Imel, Z. E., Barco, J. S., Brown, H. J., Baucom, B. R., Baer, J. S., Kircher, J. C., & Atkins, D. C. (2014). The association of therapist empathy and synchrony in vocally encoded arousal. *Journal of Counseling Psychology*, 61(1), 146.

- Keller, E., & Tschacher, W. (2007). Prosodic and gestural expression of interactional agreement. In A. Esposito, M. Faundez-Zanuy, E. Keller, & M. Marinaro (Eds.), *Verbal and nonverbal communication behaviours* (pp. 85-98). Berlin: Springer.
- Kessler, R. C., Petukhova, M., Sampson, N. A., Zaslavsky, A. M., & Wittchen, H. U. (2012). Twelve-month and lifetime prevalence and lifetime morbid risk of anxiety and mood disorders in the United States. *International journal of methods in psychiatric research*, *21*(3), 169-184. doi:10.1002/mpr.1359
- Kodama, K., Hori, K., Tanaka, S., & Matsui, H. (2018). How Interpersonal Coordination Can Reflect Psychological Counseling: An Exploratory Study. *Psychology*, *9*(5), 1128-1142. doi:10.4236/psych.2018.95070
- Koole, S. L., & Tschacher, W. (2016). Synchrony in Psychotherapy: A Review and an Integrative Framework for the Therapeutic Alliance. *Frontiers in Psychology*, *7*. doi:10.3389/fpsyg.2016.00862
- Kraft-Todd, G. T., Reinero, D. A., Kelley, J. M., Heberlein, A. S., Baer, L., & Riess, H. (2017). Empathic nonverbal behavior increases ratings of both warmth and competence in a medical context. *PLoS ONE*, *12*(5), e0177758. doi:10.1371/journal.pone.0177758
- Kupper, Z., Ramseyer, F., Hoffmann, H., & Tschacher, W. (2015). Nonverbal synchrony in social interactions of patients with schizophrenia indicates socio-communicative deficits. *PLoS ONE*, *10*(12), e0145882. doi:10.1371/journal.pone.0145882
- Leichsenring, F., Beutel, M., & Leibing, E. (2007). Psychodynamic psychotherapy for social phobia: a treatment manual based on supportive–expressive therapy. *Bulletin of the Menninger Clinic*, *71*(1), 56-84. doi:10.1521/bumc.2007.71.1.56
- Leichsenring, F., Salzer, S., Beutel, M. E., Herpertz, S., Hiller, W., Hoyer, J., . . . Poehlmann, K. (2013). Psychodynamic therapy and cognitive-behavioral therapy in social anxiety disorder: a multicenter randomized controlled trial. *American Journal of Psychiatry*, *170*(7), 759-767. doi:10.1176/appi.ajp.2013.12081125

-
- Likowski, K. U., Weyers, P., Seibt, B., Stöhr, C., Pauli, P., & Mühlberger, A. (2011). Sad and lonely? Sad mood suppresses facial mimicry. *Journal of Nonverbal Behavior, 35*(2), 101-117. doi:10.1007/s10919-011-0107-4
- Luborsky, L. (1984). *Principles of psychoanalytic psychotherapy: A manual for supportive-expressive (SE) treatment*. New York, NY: Basic Books.
- Luborsky, L., Barber, J. P., Siqueland, L., Johnson, S., Najavits, L. M., Frank, A., & Daley, D. (1996). The Revised Helping Alliance Questionnaire (HAq-II) : Psychometric Properties. *Journal of Psychotherapy Practice and Research, 5*(3), 260-271.
- Lutz, W., Schiefele, A.-K., Wucherpfennig, F., Rubel, J., & Stulz, N. (2016). Clinical effectiveness of cognitive behavioral therapy for depression in routine care: A propensity score based comparison between randomized controlled trials and clinical practice. *Journal of affective disorders, 189*, 150-158. doi:10.1016/j.jad.2015.08.072
- Mogan, R., Fischer, R., & Bulbulia, J. A. (2017). To be in synchrony or not? A meta-analysis of synchrony's effects on behavior, perception, cognition and affect. *Journal of Experimental Social Psychology, 72*, 13-20. doi:10.1016/j.jesp.2017.03.009
- Paulick, J., Deisenhofer, A.-K., Ramseyer, F., Tschacher, W., Boyle, K., Rubel, J., & Lutz, W. (2017). Nonverbal Synchrony: A new approach to better understand psychotherapeutic processes and drop-out. *Journal of Psychotherapy Integration*. doi:10.1037/int0000099
- Paulick, J., Rubel, J., Deisenhofer, A.-K., Schwartz, B., Thielemann, D., Altmann, U., . . . Lutz, W. (2018). Diagnostic features of nonverbal synchrony in psychotherapy: Comparing depression and anxiety. *Cognitive Therapy and Research, 42*(5), 539-551. doi:10.1007/s10608-018-9914-9
- Paxton, A., & Dale, R. (2012). Frame-differencing methods for measuring bodily synchrony in conversation. *Behavior Research Methods, 45*(2), 329-343. doi:10.3758/s13428-012-0249-2

- Ramseyer, F., & Tschacher, W. (2011). Nonverbal synchrony in psychotherapy: coordinated body movement reflects relationship quality and outcome. *Journal of consulting and clinical psychology, 79*(3), 284-295. doi:10.1037/a0023419
- Ramseyer, F., & Tschacher, W. (2014). Nonverbal synchrony of head- and body-movement in psychotherapy: different signals have different associations with outcome. *Frontiers in Psychology, 5*. doi:10.3389/fpsyg.2014.00979
- Reich, C. M., Berman, J. S., Dale, R., & Levitt, H. M. (2014). Vocal synchrony in psychotherapy. *Journal of Social and Clinical Psychology, 33*(5), 481-494. doi:10.1521/jscp.2014.33.5.481
- Rennung, M., & Göritz, A. S. (2016). Prosocial Consequences of Interpersonal Synchrony. *Zeitschrift für Psychologie, 224*(3), 168-189. doi:10.1027/2151-2604/a000252
- Safran, J. D., Muran, J. C., & Eubanks-Carter, C. (2011). Repairing alliance ruptures. *Psychotherapy, 48*(1), 80-87. doi:10.1037/a0022140
- Schoenherr, D., Paulick, J., Deisenhofer, A.-K., Schwartz, B., Rubel, J., Lutz, W., . . . Altmann, U. (2019). Identification of movement synchrony: Validation of time series analysis methods *PLoS ONE, 14*(2), e0211494. doi:10.1371/journal.pone.0211494
- Schoenherr, D., Paulick, J., Strauss, B., Deisenhofer, A.-K., Schwartz, B., Rubel, J., . . . Altmann, U. (2019). Nonverbal synchrony predicts premature termination of psychotherapy for social phobic patients. *Psychotherapy, accepted*.
- Schoenherr, D., Paulick, J., Worrack, S., Strauss, B., Rubel, J., Schwartz, B., . . . Altmann, U. (2018). Quantification of nonverbal synchrony using linear time series analysis methods: Convergent validity of different methods. *Behavior Research Methods, e-first*. doi:10.3758/s13428-018-1139-z
- Stangier, U., Clark, D. M., & Ehlers, A. (2006). *Soziale Phobie: Fortschritte der Psychotherapie. [Social phobia: Advances in psychotherapy]* (Vol. 28). Göttingen: Hogrefe.
- Stekhoven, D. J. (2011). Using the missForest Package. *R package*, 1-11.

-
- Stravynski, A., Kyparissis, A., Amado, D., Hoffmann, G., & DiBartolo, P. M. (2010). Social phobia as a deficit in social skills. In S. G. Hofmann & P. M. DiBartolo (Eds.), *Social anxiety: Clinical, developmental, and social perspectives* (pp. 147-181): Academic Press.
- Topel, E.-M., & Lachmann, F. M. (2007). Nonverbal dialogues: Orienting and looking behaviors between aggressive and violent children and adolescents and their therapist. *Journal of Infant, Child, and Adolescent Psychotherapy*, 6(4), 285-307. doi:10.1080/15289160701815633
- Vicaria, I. M., & Dickens, L. (2016). Meta-analyses of the intra-and interpersonal outcomes of interpersonal coordination. *Journal of Nonverbal Behavior*, 40(4), 335-361. doi:10.1007/s10919-016-0238-8
- Vrijzen, J. N., Lange, W.-G., Becker, E. S., & Rinck, M. (2010). Socially anxious individuals lack unintentional mimicry. *Behaviour research and therapy*, 48(6), 561-564. doi:10.1016/j.brat.2010.02.004
- Vrijzen, J. N., Lange, W.-G., Dotsch, R., Wigboldus, D. H., & Rinck, M. (2010). How do socially anxious women evaluate mimicry? A virtual reality study. *Cognition and Emotion*, 24(5), 840-847. doi:10.1080/13854040902833652
- Wittchen, H.-U., Wunderlich, U., Gruschwitz, S., & Zaudig, M. (1997). *SKID I. Strukturiertes Klinisches Interview für DSM-IV. Achse I: Psychische Störungen. Interviewheft und Beurteilungsheft. [SCID I. Structured clinical interview for DSM-IV. Axis I: Mental disorder. Manual.]*. Goettingen: Hogrefe.

Table 1: Descriptive statistics of the entire sample and the three treatment groups.

| | Entire sample (N=267) N in % or M (SD) | Naturalistic CBT (N=111) N in % or M (SD) | Manual guided CBT (N=102) N in % or M (SD) | Manual guided PDT (N=54) N in % or M (SD) | Comparison of approaches |
|--|--|---|--|---|--|
| Gender patients: female | 53.6% | 44.1% | 58.8% | 63.0% | $\chi^2_{df=2} = 7.012$, $p = .030$, $V = .162$ |
| Age of patients | 34.4 (11.8) | 33.8 (11.6) | 34.8 (11.8) | 34.8 (12.2) | n.s. |
| Education: high school degree | 59.2% | 47.7% | 62.7% | 75.9% | $\chi^2_{df=2} = 12.810$, $p = .002$, $V = .219$ |
| Current relationship | 50.6% | 57.7% | 39.2% | 57.4% | $\chi^2_{df=2} = 8.501$, $p = .014$, $V = .178$ |
| IIP pre | 1.8 (0.4) | 1.8 (0.5) | 1.8 (0.4) | 1.7 (0.4) | n.s. |
| BDI pre | 13.7 (8.0) | 18.0 (7.8) | 11.1 (6.5) | 9.7 (6.9) | $F_{df=2} = 34.939$, $p < .001$, $\eta^2 = 0.209$ |
| Primary diagnosis | | | | | |
| Social anxiety disorder (F40.1) | 76.4% | 43.2% | 100% | 100% | $\chi^2_{df=8} = 115.8$, $p < .001$, $V = .466$ |
| Depression (F32, F33, F34) | 16.1% | 38.7% | 0% | 0% | |
| Other anxiety disorder without SAD (F40, F41) | 3.0% | 7.2% | 0% | 0% | |
| Other disorders | 4.5% | 10.8% | 0% | 0% | |
| Secondary diagnosis | | | | | |
| Social anxiety disorder (F40.1) | 23.2% | 55.9% | 0% | 0% | $\chi^2_{df=8} = 123.6$, $p < .001$, $V = .481$ |
| Depression (F32, F33, F34) | 24.3% | 18.9% | 29.4% | 25.9% | |

| | | | | | |
|---|------------|------------|------------|-------------|--|
| Other anxiety disorder without SAD (F40, F41) | 10.1% | 6.3% | 9.8% | 18.5% | |
| Other disorders | 17.6% | 11.7% | 20.6% | 24.1% | |
| No secondary diagnosis | 24.7% | 7.2% | 40.2% | 31.5% | |
| Number of comorbid diagnosis | 1.3 (1.0) | 1.7 (1.0) | 0.9 (0.9) | 1.1 (1.0) | $F_{df=2} = 21.394$, $p < .001$, $\eta^2 = 0.139$ |
| Gender therapist: female | 74.5% | 85.6% | 80.4% | 40.7% | $\chi^2_{df=2} = 41.474$, $p < .001$, $V = .394$ |
| Age of therapists | 32.3 (7.7) | 30.8 (6.2) | 29.7 (4.1) | 40.4 (10.0) | $F_{df=2} = 52.641$, $p < .001$, $\eta^2 = 0.285$ |

Note: CBT = cognitive behavioral therapy, PDT = psychodynamic therapy, IIP pre = total score of inventory of interpersonal problems assessed at beginning of therapy, BDI pre = degree of depression assessed with Beck Depression Inventory at beginning of therapy.

Table 2: Average movement synchrony and leading depending on therapeutic approach and therapy session.

| sessio n | entire sample | | naturalistic CBT | | manual guided CBT | | manual guided PDT | |
|-------------|--|---------------|------------------|---------------|-------------------|---------------|-------------------|---------------|
| | <i>M</i> | (<i>SD</i>) | <i>M</i> | (<i>SD</i>) | <i>M</i> | (<i>SD</i>) | <i>M</i> | (<i>SD</i>) |
| | frequency of movement synchrony (raw data means) | | | | | | | |
| S3 | 24.0 | (9.2) | 25.7 | (10.2) | 26.3 | (6.8) | 16.2 | (6.6) |
| S8 | 22.0 | (8.4) | 23.1 | (8.8) | 24.2 | (7.0) | 15.2 | (6.1) |
| S20 | 19.7 | (10.4) | 20.7 | (10.8) | 21.7 | (10.5) | 14.7 | (7.4) |
| S29 | 19.1 | (9.4) | 18.1 | (8.7) | 24.7 | (9.6) | 13.2 | (5.5) |
| | leading (raw data means) | | | | | | | |
| S3 | 0.3 | (2.4) | 0.1 | (2.6) | 0.8 | (2.1) | -0.2 | (2.5) |
| S8 | 0.3 | (2.2) | 0.0 | (2.2) | 0.3 | (2.1) | 0.8 | (2.1) |
| S20 | 0.6 | (2.9) | 0.4 | (3.0) | 0.5 | (2.9) | 0.9 | (2.7) |
| S29 | 0.4 | (2.3) | 0.4 | (2.4) | 0.5 | (2.5) | 0.1 | (1.9) |
| | frequency of movement synchrony (adjusted means) | | | | | | | |
| S3 | 23.8 | (11.2) | 26.0 | (11.2) | 25.6 | (11.6) | 15.9 | (10.7) |
| S8 | 23.0 | (10.1) | 26.2 | (10.3) | 23.7 | (10.0) | 14.8 | (9.9) |
| S20 | 20.7 | (11.0) | 22.7 | (11.3) | 22.4 | (10.8) | 13.1 | (10.9) |
| S29 | 18.7 | (13.4) | 16.3 | (13.1) | 24.4 | (14.1) | 12.8 | (12.4) |
| | leading (adjusted means) | | | | | | | |
| S3 | 0.2 | (3.1) | -0.2 | (3.0) | 0.8 | (3.4) | -0.1 | (2.7) |
| S8 | 0.4 | (2.5) | 0.1 | (2.5) | 0.6 | (2.6) | 0.6 | (2.3) |
| S20 | 0.6 | (3.0) | 0.6 | (3.1) | 0.5 | (3.0) | 0.9 | (2.8) |
| S29 | 0.5 | (4.1) | 0.6 | (4.0) | 0.7 | (4.4) | -0.2 | (3.6) |

Note. *M* = mean, *SD* = standard deviation, CBT = cognitive behavioral therapy, PDT = psychodynamic oriented therapy, adjusted means based on a growth curve model in which we controlled for initial interpersonal problems, initial depression and exact therapy session, used propensity scores as regression weights, and included a random effect for multiple patients per therapist.

Table 3: Prediction of therapy outcome.

| outcome | model 1 (therapy arm as random effect) | | | | | | model 2 (multi group model) | | | | | | |
|---|---|--------------------|---------|-------------------|-------------------|---------|-----------------------------|--------------------|---------|-------------------|--------------------|---------|-------|
| | naturalistic CBT | | | manual guided CBT | | | naturalistic CBT | | | manual guided CBT | | | |
| | b | SE | β | b | SE | β | b | SE | β | b | SE | β | |
| outcome HAQ S20 | intercept | 5*** | (0.16) | -0.04 | 4.82*** | (0.23) | 0.19 | 5.26*** | (0.28) | -0.08 | 5.26*** | (0.31) | -0.02 |
| | IIP pre | 0 | (0.08) | 0.00 | 0.16 | (0.14) | 0.14 | 0.01 | (0.13) | 0.00 | -0.31 ⁺ | (0.17) | -0.26 |
| | BDI pre | -0.01 ⁺ | (0.00) | -0.12 | -0.01 | (0.01) | -0.11 | -0.01 | (0.01) | -0.10 | -0.01 | (0.01) | -0.21 |
| | synchrony S3 | 0.02*** | (0.00) | 0.29 | 0.02** | (0.01) | 0.28 | 0 | (0.01) | 0.06 | 0.05** | (0.01) | 0.57 |
| | leading S3 | -0.01 | (0.01) | -0.07 | 0 | (0.02) | 0.02 | -0.04 ⁺ | (0.02) | -0.17 | 0 | (0.02) | 0.01 |
| ICC _{therapist} = 30.6%; ICC _{approach} = 2.6% | | | | | | | | | | | | | |
| outcome HAQ S20 | intercept | 5.12*** | (0.16) | -0.03 | 4.62*** | (0.26) | 0.18 | 5.48*** | (0.27) | -0.09 | 5.39*** | (0.31) | -0.15 |
| | IIP pre | 0.01 | (0.08) | 0.01 | 0.22 ⁺ | (0.13) | 0.19 | -0.06 | (0.13) | -0.05 | -0.26 | (0.18) | -0.22 |
| | BDI pre | -0.01 | (0.00) | -0.09 | 0 | (0.01) | -0.06 | -0.01 | (0.01) | -0.08 | 0 | (0.01) | -0.07 |
| | synchrony S8 | 0.01* | (0.00) | 0.17 | 0.02** | (0.01) | 0.32 | 0 | (0.01) | -0.03 | 0.02 | (0.01) | 0.28 |
| | leading S8 | -0.04* | (0.01) | -0.15 | -0.04 | (0.03) | -0.15 | -0.02 | (0.02) | -0.09 | -0.06 ⁺ | (0.03) | -0.24 |
| ICC _{therapist} = 29.8%; ICC _{approach} = 3.1% | | | | | | | | | | | | | |
| outcome IIP _{post} | intercept | 0.47* | (0.16) | 0.00 | 0.27 | (0.19) | -0.3 | 0.54* | (0.23) | -0.09 | 0.66* | (0.26) | 0.45 |
| | IIP pre | 0.56*** | (0.07) | 0.46 | 0.41*** | (0.11) | 0.3 | 0.61*** | (0.11) | 0.51 | 0.55*** | (0.14) | 0.45 |
| | BDI pre | 0.01* | (0.00) | 0.14 | 0.02*** | (0.01) | 0.4 | 0.01 | (0.01) | 0.11 | 0 | (0.01) | -0.06 |
| | synchrony S3 | -0.01** | (0.00) | -0.16 | 0 | (0.00) | -0.1 | -0.02** | (0.01) | -0.30 | 0 | (0.01) | 0.02 |
| | leading S3 | 0.03* | (0.01) | 0.12 | 0 | (0.02) | 0.0 | 0.05** | (0.02) | 0.24 | 0.01 | (0.02) | 0.06 |
| ICC _{therapist} = 25.8%; ICC _{approach} = 14.8% | | | | | | | | | | | | | |
| outcome IIP _{post} | intercept | 0.51** | (0.15) | -0.03 | 0.43* | (0.21) | -0.29 | 0.36 | (0.22) | -0.04 | 0.57* | (0.25) | 0.53 |
| | IIP pre | 0.59*** | (0.07) | 0.48 | 0.44*** | (0.11) | 0.36 | 0.74*** | (0.11) | 0.61 | 0.49** | (0.15) | 0.40 |
| | BDI pre | 0.01 | (0.00) | 0.09 | 0.02** | (0.01) | 0.33 | 0.01 | (0.01) | 0.09 | 0 | (0.01) | -0.02 |
| | synchrony S8 | -0.01** | (0.00) | -0.20 | -0.01* | (0.01) | -0.19 | -0.02*** | (0.01) | -0.30 | 0.01 | (0.01) | 0.17 |

| | | | | | | | | | | | | |
|---------------------|---|--------|-------|---------|--------|-------|---------|--------|-------|--------------------|--------|-------|
| leading S8 | 0.03* | (0.01) | 0.11 | -0.01 | (0.02) | -0.05 | 0.05** | (0.02) | 0.23 | 0.03 | (0.02) | 0.14 |
| | ICC _{therapist} = 30.3%; ICC _{approach} = 11.2% | | | | | | | | | | | |
| outcome | | | | | | | | | | | | |
| BDI _{post} | | | | | | | | | | | | |
| intercept | 1.01 | (1.41) | 0.01 | 0.23 | (2.09) | -0.03 | -1.2 | (2.53) | -0.15 | 4.5 | (2.78) | -0.01 |
| IIP pre | 1.44 ⁺ | (0.74) | 0.11 | 1.18 | (1.23) | 0.09 | 0.46 | (1.18) | 0.03 | 2.06 | (1.54) | 0.15 |
| BDI pre | 0.3*** | (0.04) | 0.42 | 0.33*** | (0.07) | 0.46 | 0.41*** | (0.07) | 0.57 | 0.14 ⁺ | (0.08) | 0.19 |
| synchrony S3 | -0.06 ⁺ | (0.03) | -0.10 | -0.03 | (0.05) | -0.05 | 0 | (0.06) | 0.00 | -0.16 ⁺ | (0.09) | -0.26 |
| leading S3 | 0.35** | (0.12) | 0.14 | -0.01 | (0.22) | 0.00 | 0.81*** | (0.20) | 0.34 | 0.27 | (0.22) | 0.11 |
| | ICC _{therapist} = 23.8%; ICC _{approach} = 2.7% | | | | | | | | | | | |
| outcome | | | | | | | | | | | | |
| BDI _{post} | | | | | | | | | | | | |
| intercept | -0.49 | (1.49) | 0.01 | -0.55 | (2.35) | -0.05 | -2.59 | (2.47) | -0.08 | 1.84 | (2.82) | 0.20 |
| IIP pre | 1.38 ⁺ | (0.74) | 0.10 | 0.7 | (1.21) | 0.05 | 1.85 | (1.22) | 0.14 | 1.21 | (1.65) | 0.09 |
| BDI pre | 0.3*** | (0.04) | 0.41 | 0.35*** | (0.08) | 0.48 | 0.38*** | (0.07) | 0.53 | 0.14 | (0.10) | 0.19 |
| synchrony S8 | 0.01 | (0.04) | 0.02 | 0.03 | (0.06) | 0.04 | -0.01 | (0.06) | -0.02 | 0.07 | (0.12) | 0.09 |
| leading S8 | 0.33* | (0.13) | 0.12 | 0.16 | (0.25) | 0.06 | 0.45* | (0.21) | 0.17 | 0.51 ⁺ | (0.27) | 0.19 |
| | ICC _{therapist} = 23%; ICC _{approach} = 3.8% | | | | | | | | | | | |

Note: CBT = cognitive behavioral therapy, PDT = psychodynamic therapy, b = regression coefficient, SE = standard error, β = standardized regression coefficient, *** $p < .001$, ** $p < .01$, * $p < .05$, + $p < .1$, IIP = total score of inventory of interpersonal problems, BDI = sum score of beck depression inventory, HAQ = sum score of Helping Alliance Questionnaire, pre = measurement at beginning of therapy, post = measurement at end of therapy, S3, S8 and S20 = session 3, 8 and 20, synchrony = frequency of movement synchrony, leading = proportion of leading when synchronizing movement (positive values mean that patient leads more often than the therapist respectively that the therapist mimicked more often the patient), ICC = intra-class correlation.

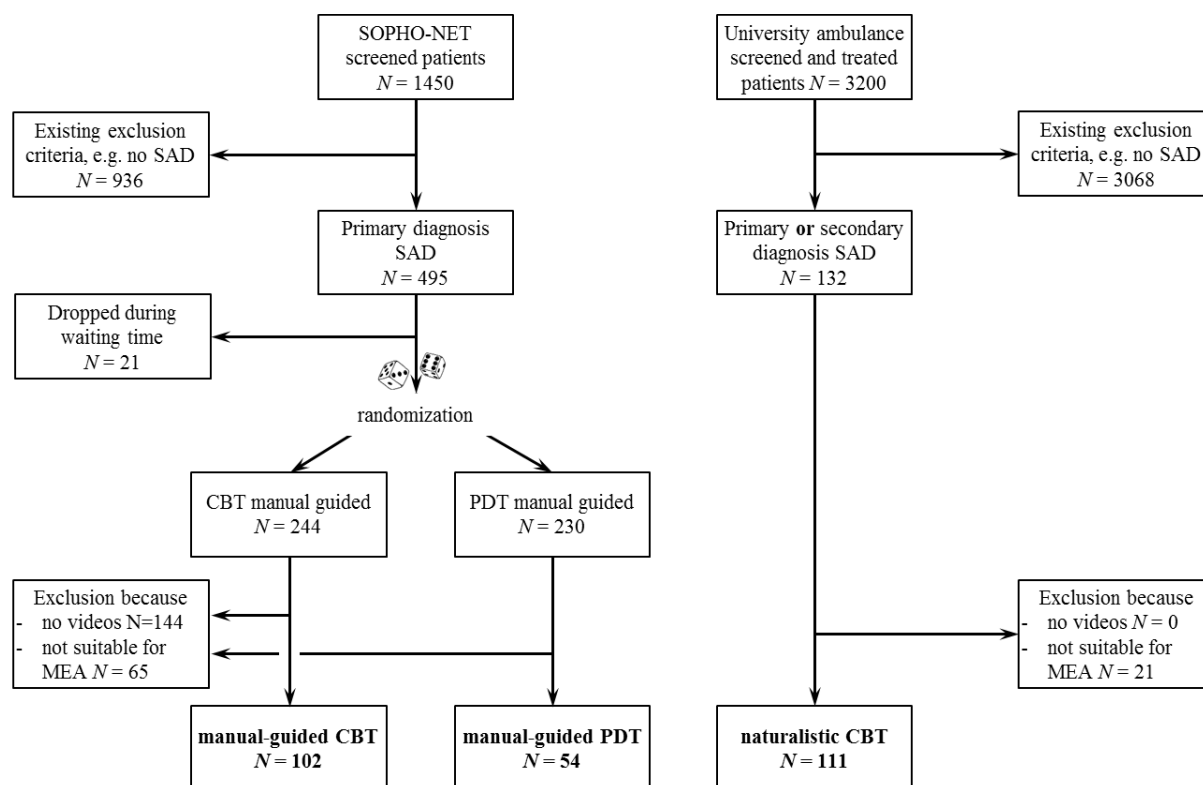


Figure 1: Flow chart (SAD = social anxiety disorder, CBT = cognitive behavioral therapy, PDT = psychodynamic therapy, MEA = Motion Energy Analysis).

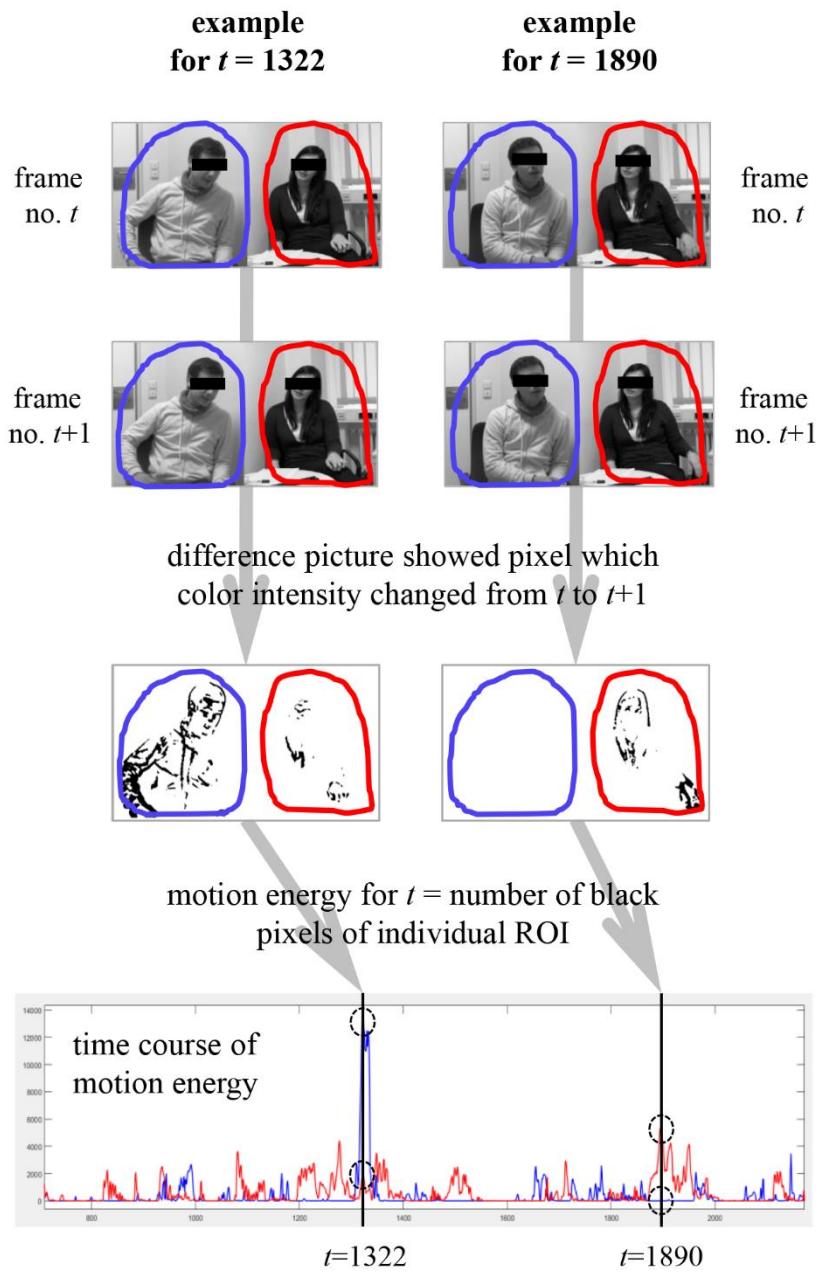


Figure 2: Illustration of motion energy analysis (for two time points are showed: two consecutive frames in which the individual regions of interest (ROI) are marked, their difference picture, and time course of both motion energy time series whereby the resulting motion energy values of both examples are marked).

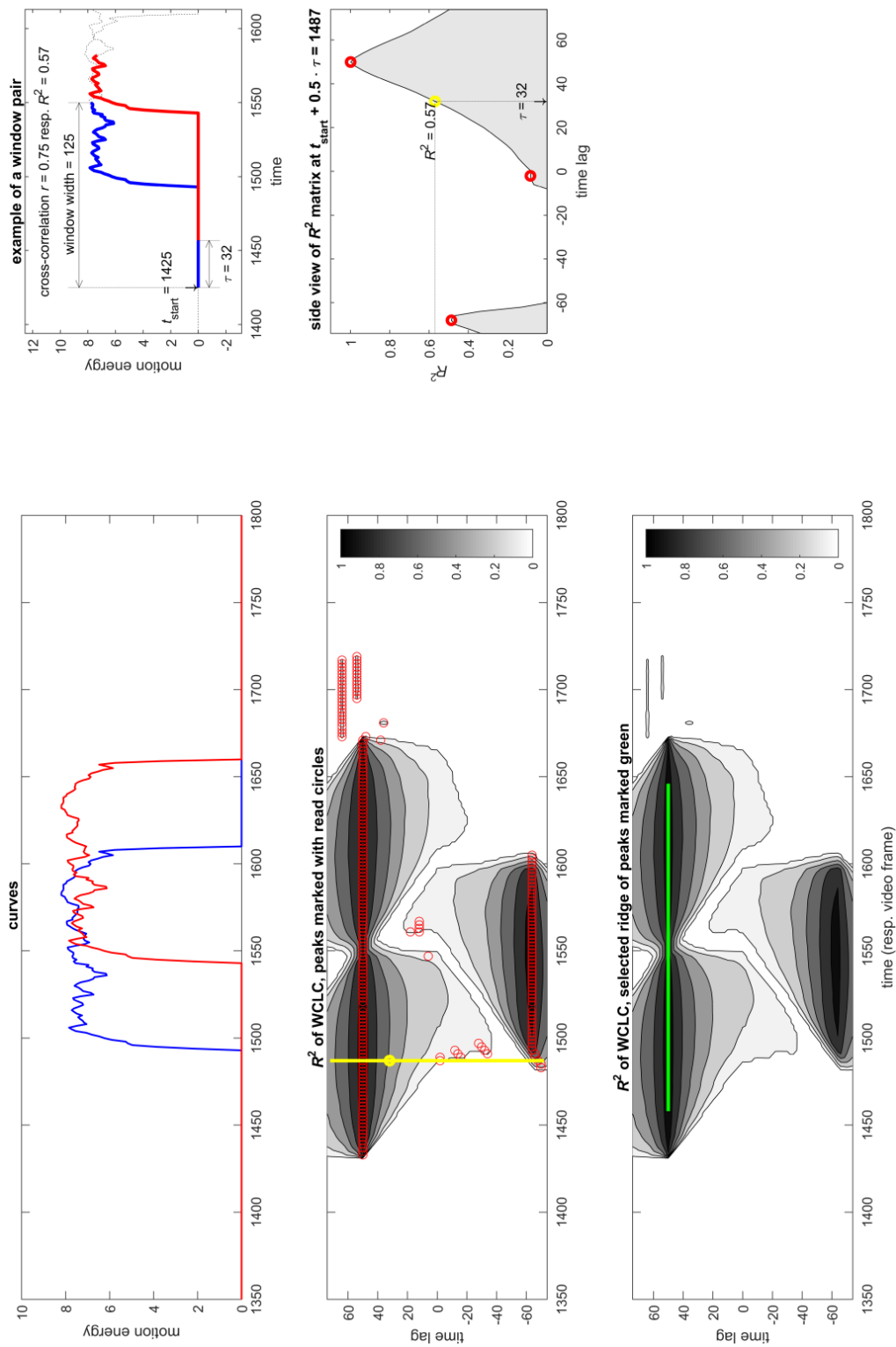


Figure 3: Illustration of windowed cross-lagged correlation (WCLC) and the subsequent peak picking of resulting R^2 -matrix, the synchronization interval identified by the algorithm is marked with green line in the lower subplot.

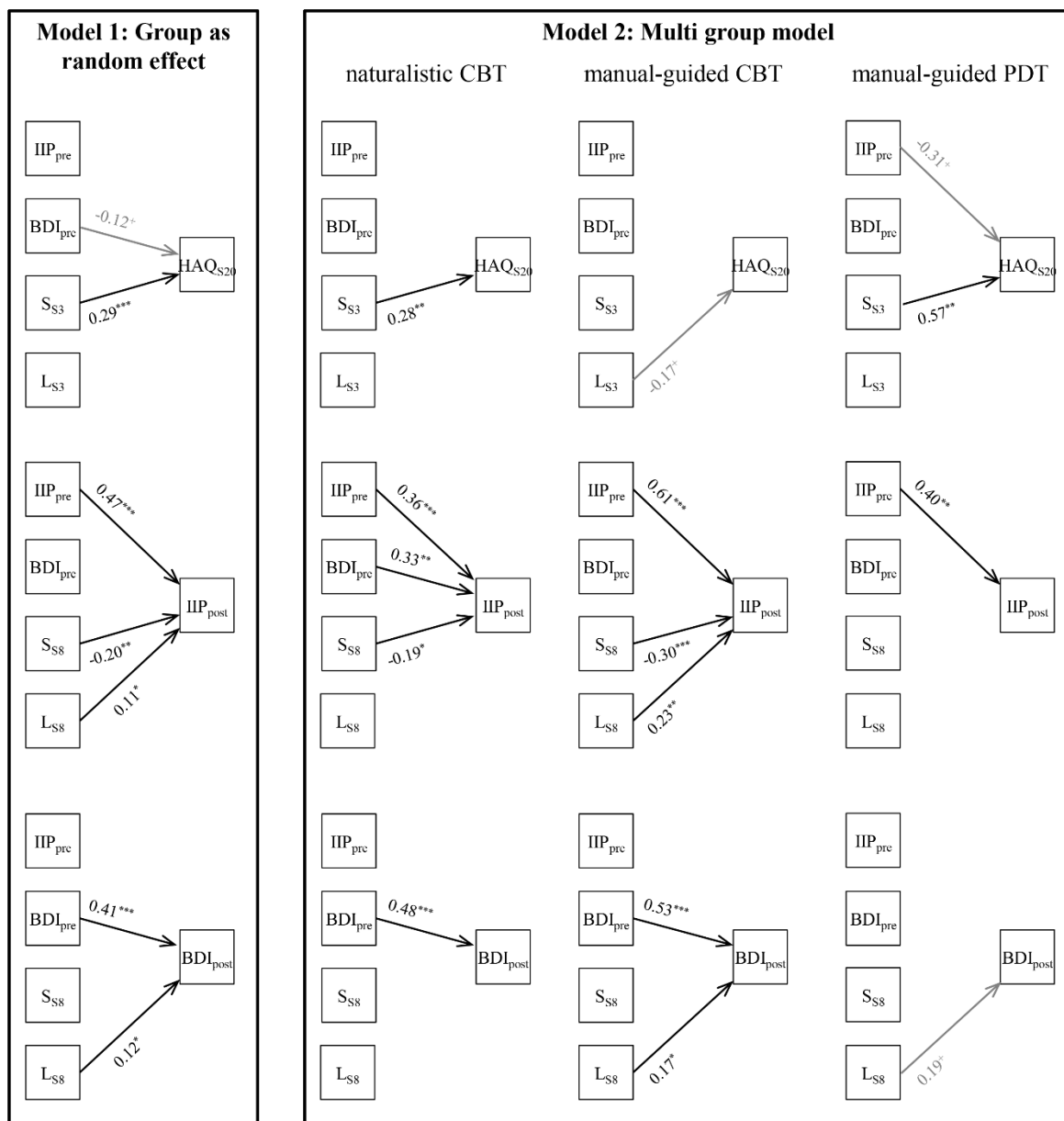


Figure 4: Associations between interpersonal problems (IIP), depression (BDI), therapeutic alliance (HAQ), frequency of movement synchrony (S) and leading by patient (L) depending on therapeutic approach (CBT = cognitive behavioral therapy, PDT = psychodynamic therapy, S3 = session 3, S8 = session 8, S20 = session 20; only significant standardized regression coefficients of HLM; *** $p < .001$, ** $p < .01$, * $p < .05$, + $p < .1$).

Appendix
Table A.1: Average movement frequency of patient and therapist depending on therapeutic approach and therapy session..

| session | entire sample | | naturalistic CBT | | manual guided CBT | | manual guided PDT | |
|--|---------------|--------|------------------|--------|-------------------|--------|-------------------|--------|
| | M | (SD) | M | (SD) | M | (SD) | M | (SD) |
| frequency of patients movements (raw data means) | | | | | | | | |
| S3 | 15.8 | (11.6) | 18.5 | (13.3) | 14.3 | (7.0) | 10.9 | (9.0) |
| S8 | 15.1 | (11.0) | 18.5 | (13.7) | 12.7 | (6.9) | 12.2 | (8.0) |
| S20 | 14.8 | (11.2) | 17.1 | (12.4) | 13.7 | (10.6) | 12.1 | (8.8) |
| S29 | 14.4 | (10.0) | 15.0 | (11.2) | 15.1 | (9.0) | 12.4 | (8.8) |
| frequency of therapists movements (raw data means) | | | | | | | | |
| S3 | 13.9 | (8.1) | 15.4 | (8.5) | 15.4 | (7.5) | 9.0 | (5.4) |
| S8 | 14.5 | (8.4) | 15.3 | (8.1) | 17.9 | (8.5) | 8.8 | (5.4) |
| S20 | 13.7 | (8.5) | 13.1 | (7.8) | 16.7 | (9.4) | 9.4 | (6.8) |
| S29 | 12.5 | (7.4) | 13.0 | (8.5) | 15.4 | (5.8) | 7.8 | (4.6) |
| frequency of patients movements (adjusted means) | | | | | | | | |
| S3 | 16.0 | (13.1) | 20.0 | (13.4) | 13.6 | (13.5) | 12.1 | (11.7) |
| S8 | 15.9 | (11.6) | 20.8 | (12.3) | 12.9 | (11.3) | 11.4 | (10.7) |
| S20 | 15.4 | (12.8) | 19.3 | (13.5) | 13.4 | (12.4) | 10.9 | (12.0) |
| S29 | 14.8 | (15.8) | 15.2 | (15.8) | 16.0 | (16.7) | 11.7 | (14.1) |
| frequency of therapists movements (adjusted means) | | | | | | | | |
| S3 | 14.4 | (10.1) | 15.6 | (9.4) | 16.1 | (10.6) | 8.8 | (10.7) |
| S8 | 14.5 | (9.3) | 15.3 | (8.6) | 16.9 | (9.5) | 8.5 | (10.3) |
| S20 | 13.5 | (10.2) | 13.5 | (9.5) | 16.5 | (10.4) | 7.9 | (11.3) |
| S29 | 11.6 | (11.8) | 11.4 | (11.0) | 14.0 | (12.5) | 7.5 | (12.0) |

Deskriptive Daten Bewegungsanalysen N =86

Table 1

Descriptive statistics of the sample after imputation of missings

| | Stichprobe Bewegungs- analysen (N=86) | Gesamtstich- probe (N =267) |
|--|---|--------------------------------|
| Gender female | 51% | 54% |
| Age in years | 34.25 (11.72) | 34.38 (11.77) |
| Current relationship | 49% | 51% |
| High school diploma or higher | 58% | 59% |
| CBT manual-guided | 20% | 38% |
| PDT manual-guided | 28% | 20% |
| CBT naturalistic | 52% | 42% |
| BDI pre mean (sd) | 13.72 (7.63) | 13.67 (8.04) |
| IIP pre mean (sd) | 1.79 (.42) | 1.80 (0.43) |
| Number of comorbid diagnoses | 1.48 (.99) | 1.27 (1.03) |
| Social anxiety disorder (1 st diagnosis) | 69% | 76% |
| Affective disorder (2 nd diagnosis) | 34% | 32% |
| Other disorder (2 nd diagnosis) | 48% | 36% |
| none | 18% | 32% |
| Affective Disorder (1 st diagnosis)* | 25% | 17% |
| Other Anxiety Disorder (1 st diagnosis) * | 2% | 5% |
| Other Disorders (1 st diagnosis)* | 4% | 1% |

Note. CBT = cognitive behavioral therapy; PDT = psychodynamic-orientated therapy, BDI = Beck Depression Inventory, IIP = Inventory of Interpersonal Problems, * with social anxiety disorder as the 2nd diagnosis.

Publikationen und Kongressbeiträge

Publikationen – angenommen/veröffentlicht

- *Schoenherr, D., Paulick, J., Strauss, B., Deisenhofer, A.-K., Schwartz, B., Rubel, J., Lutz, W., Stangier, U., Altmann, U. (zur Publikation angenommen, 06.12.2018). Nonverbal synchrony predicts premature termination of psychotherapy for social anxiety disorder. *Psychotherapy*.
- *Schoenherr, D., Paulick, J., Strauss, B., Deisenhofer, A.-K., Schwartz, B., Rubel, J., Lutz, W., Stangier, U., Altmann, U. (2019). Identification of movement synchrony: Validation of windowed cross-lagged correlation and –regression with peak-picking algorithm. *PLoS One*, 14(2). doi:10.1371/journal.pone.0211494
- *Schoenherr, D., Paulick, J., Worrack, S., Strauss, B., Rubel, J., Schwartz, B., Deisenhofer, A.-K., Lutz, W., Stangier, U., Altmann, U. (2018). Quantification of nonverbal synchrony using linear time series analysis methods: Lack of convergent validity and evidence for facets of synchrony. *Behavior Research Methods*, 1-23. doi:10.3758/s13428-018-1139-z.
- Altmann, U., Thielemann, D., Zimmermann, A., Steffanowski, A., Bruckmayer, E., Pfaffinger, I., Fernbacher, A., Strauss, B. (accepted). Outpatient psychotherapy improves symptoms and reduces health care costs in regular and prematurely terminated therapies. *Frontiers in Psychology – Clinical and Health Psychology*, 9.
- Thielemann, D., Richter, F., Strauss, B., Braehler, E., Altmann, U., Berger, U. (2018). Differential Item Functioning in Brief Instruments of Disordered Eating. *European Journal of Psychological Assessment*.
- Paulick, J., Rubel, J. A., Deisenhofer, A. K., Schwartz, B., Thielemann, D., Altmann, U., Boyle, K., Strauss, B. & Lutz, W. (2018, efirst). Diagnostic Features of Nonverbal Synchrony in Psychotherapy: Comparing Depression and Anxiety. *Cognitive Therapy and Research*, 1-13. doi: 10.1007/s10608-018-9914-9
- Preißler, S., Thielemann, D., Dietrich, C., Hofmann, G. O., Miltner, W. H., & Weiss, T. (2017). Preliminary evidence for training-induced changes of morphology and phantom limb pain. *Frontiers in Human Neuroscience*, 11, 319.
- Thielemann, D., Sengewald, M.-A., Kappler, G., Steyer, R. (2017). A probit latent state IRT model with latent item-effect variables. *European Journal of Psychological Assessment*, 33(4), 271-284. doi:10.1027/1015-5759/a000417
- Altmann, U., Thielemann, D., Strauß, B. (2016) Ambulante Psychotherapie unter Routinebedingungen: Forschungsbefunde aus dem deutschsprachigen Raum. *Psychiatrische Praxis*, 43(7), 360-366. doi: 10.1055/s-0042-115616

*In dieser Arbeit zusammengefasste Publikationen

Publikationen – in Vorbereitung/ eingereicht

- *Altmann, U., Schoenherr, D., Paulick, J., Deisenhofer, A.-K., Schwartz, B., Rubel, J., Stangier, U., Lutz, W., Strauss, B. (in Begutachtung, eingereicht am 14.02.2019) Associations between movement synchrony and outcome in patients with social anxiety disorder: Evidence for treatment specific effects. *Psychotherapy Research*.
- Altmann, U., Schoenherr, D., Paulick, J., Knitter, L.A., Worrack, S., Deisenhofer, A.-K., Schwartz, B., Rubel, J., Lutz, W., Strauss, B., (in Überarbeitung). Introduction, Practical Guide, and Validation Study for Measuring Body Movements Using Motion Energy Analysis.
- Lutz, W., Prinz, J., Schwartz, B., Paulick, J., Schoenherr, D., Deisenhofer, A.-K., Terhürne, P., Boyle, K., Altmann, U., Strauss, B., Rafaeli, E., Atzil-Slonim, D., Bar-Kalifa, E., Rubel, J., (in Begutachtung, eingereicht am 15.02.2019). Patterns of early change in interpersonal problems and their relationship to nonverbal synchrony and multidimensional outcome. *Journal of Counseling Psychology*.
- Rubel, J., Cardace, F., Paulick, J., Deisenhofer, A.-K., Schwartz, B., Merkler, M., Altmann, U., Schoenherr, D., Strauss, B., Lutz, W. (in Überarbeitung). Konvergenz von Skalen zur Erfassung sozialer Ängste: Ein Faktoranalytischer Linking-Ansatz und dessen Anwendung.
- *Schoenherr, D., Strauss, B., Paulick, J., Deisenhofer, A.-K., Schwartz, B., Rubel, J., Lutz, W., Stangier, U., Altmann, U. (in Begutachtung, eingereicht am 07.03.2019). Movement synchrony and attachment related anxiety and avoidance in social anxiety disorder. *Journal of Counseling Psychology*.

Kongressbeiträge – vorgestellt

- Preißler, S., Thielemann, D., Dietrich, C., Hofmann, G. O., Miltner, W. H. R., Weiss, T. (2014). Trainingsbedingte Veränderungen der Hirnstruktur und der Phantomschmerzsymptomatik bei Armamputationspatienten. Posterpräsentation durch D. Thielemann auf dem Deutschen Schmerzkongress 2014, Hamburg.
- Thielemann, D., Steyer, R., Sengewald, M-A. (2015). Latente personenspezifische Itemparameter als Erweiterung des Rasch Modells. Vortrag auf der 12. Tagung der Fachgruppe für Methoden und Evaluation 2015, Jena.
- Altmann, U., Thielemann, D., & Strauß, B. Does nonverbal synchronization depend on the topic of conversation and behavior modality? Vortrag, 47th Annual Meeting of the Society for Psychotherapy Research, Jerusalem, Israel, 25.06.2016
- Thielemann, D., Dittmann, J., Worrack, S., Lutz, W., Strauss, B., Altmann, U. Vergleich verschiedener Analysemethoden für nonverbale Synchronisation in menschlichen Interaktionen. Vortrag, Deutscher Kongress für Psychosomatische Medizin und Psychotherapie 2017, Berlin, 23.03.2017
- Thielemann, D., Dittmann, J., Lutz, W., Strauss, B., Altmann, U. Bewegungssynchronisation von Patient und Psychotherapeut: Eine Validierungsstudie zeitreihenanalytischer

- Verfahren zur Absicherung gegen Methoden-Artefakte. Posterpräsentation, Deutscher Kongress für Psychosomatische Medizin und Psychotherapie 2017, Berlin, 24.03.2017
- Thielemann, D., Paulick, J., Strauss, B., Deisenhofer, A.-K., Schwartz, B., Rubel, J., Lutz, W., Altmann, U. (2017) Nonverbale Synchronisation – Ein Prädiktor für Therapieabbrüche bei Patienten mit Sozialer Phobie? Vortrag im Rahmen der Fortbildungsreihe am Institut für Medizinische Psychologie des Universitätsklinikums Heidelberg, Heidelberg, 18.07.2017.
- Thielemann, D., Paulick, J., Strauss, B., Deisenhofer, A.-K., Schwartz, B., Rubel, J., Lutz, W., Altmann, U. (2017) Nonverbal synchronization – A predictor of premature termination for patients suffering from social anxiety disorder? Presentation at 4th Joint Conference Society for Psychotherapy Research UK & European Chapters, Oxford, United Kingdom, 21.09.2017.
- Thielemann, D., Strauss, B., Altmann, U. (2017) Identification of movement synchrony: Construct validation of time series analysis methods. Presentation at the 48th International Annual Meeting of the Society for Psychotherapy Research, Toronto, Canada. 22.06.2017.
- Thielemann, J. Paulick, B. Strauß, B. Schwartz, A.-K. Deisenhofer, J. Rubel, W. Lutz, U. Altmann (2018) Nonverbale Synchronisation – Ein Prädiktor für Therapieabbrüche bei Patienten mit Sozialer Phobie? Vortrag, Deutscher Kongress für Psychosomatische Medizin und Psychotherapie, Berlin, 23.03.2018
- Thielemann, D., Paulick, J., Strauss, B., Deisenhofer, A.-K., Schwartz, B., Rubel, J., Lutz, W., Altmann, U. (2018). Anwendung und Güte von Methoden zur automatisierten Erhebung von Bewegungen und Bewegungssynchronisation. Vortrag, 36. Symposium der Fachgruppe Klinische Psychologie und Psychotherapie der DGPS, Landau, 11.05.2018
- Schoenherr, D., Paulick, J, Strauss, B., Deisenhofer, A.-K., Schwartz, B., Rubel, J., Lutz, W., Altmann, U. (2018). Nonverbal synchrony – A predictor of premature termination for patients suffering from social anxiety disorder? Presentation at 49th SPR International Annual Meeting, Amsterdam, The Netherlands, 29.06.2018.
- Schönherr, D., Paulick, J., Strauss, B., Schwartz, B., Deisenhofer, A.-K., Rubel, J., Lutz, W., Stangier, U., Altmann, U. (2018). Nonverbale Synchronie von Bewegungen als Prädiktor für einen Therapieabbruch bei Patienten mit Sozialer Phobie. Vortrag auf dem Gemeinsamen Kongress der DGMP und DGMS, Leipzig, 28.09.2018.
- Schoenherr, D., Paulick, J., Worrack, S., Strauss, B., Rubel, J., Schwartz, B., Deisenhofer, A.-K., Lutz, W., Stangier, U., Altmann, U. (2019). Using linear time series analysis methods: Lack of convergent validity and evidence of different facets of movement synchrony. Presentation at 20. Herbstakademie - Synchronization in Embodied Interaction, Freiburg, 01.03.2019.
- Schoenherr, D., Strauss, B., Paulick, J., Deisenhofer, A.-K., Schwartz, B., Rubel, J., Boyle, K., Lutz, W., Stangier, U., Altmann, U. (2019). The association of movement synchrony with

attachment and outcome in social anxiety disorder. Presentation at 20. Herbstakademie - Synchronization in Embodied Interaction, Freiburg, 02.03.2019.

Kongressbeiträge – beteiligt

Paulick, J., Thielemann, D., Schwartz, B., Rubel, J., Deisenhofer, A.-K., Altmann, U., Strauß, B. & Lutz, W. (2018). Zusammenhänge zwischen nonverbaler Synchronizität und Early Response in der ambulanten Psychotherapie. Vortrag, Deutscher Kongress für Psychosomatische Medizin und Psychotherapie, Berlin, 23.03.2018.

Altmann, U., Thielemann, D., Paulick, J., Deisenhofer, A.-K., Schwartz, B., Rubel, J., Lutz, W., Strauß, B. (2018). Nonverbale Synchronisation und Therapieerfolg: Ein Vergleich von tiefenpsychologisch fundierter Therapie und Verhaltenstherapie bei sozialen Phobien. Vortrag, 36. Symposium der Fachgruppe Klinische Psychologie und Psychotherapie der DGPS, Landau, 11.0.2018

Paulick, J., Rubel, J., Thielemann, D., Deisenhofer, A.-K., Schwartz, B., Terhürne, P., Altmann, U., Strauss, B., Lutz, W. (2018). Nonverbale Synchronizität bei Patienten mit Sozialer Phobie: Ein Moderator zwischen early response und dem Therapieergebnis? Vortrag, 36. Symposium der Fachgruppe Klinische Psychologie und Psychotherapie der DGPS, Landau, 11.0.2018

Schwartz, B., Paulick, J., Thielemann, D., Deisenhofer, A.-K., Rubel, J., Altmann, U., Strauß, B., Lutz, W. (2018). Nonverbale Synchronie bei Sozialer Phobie: Zum Attunement von Patient und Therapeut. Vortrag, 36. Symposium der Fachgruppe Klinische Psychologie und Psychotherapie der DGPS, Landau, 11.05.2018

Altmann, U., Schoenherr, D., Paulick, J., Deisenhofer, A.-K., Schwartz, B., Rubel, J., Lutz, W., Strauß, B. (2018). Nonverbal synchrony during psychotherapy of patients suffering from social anxiety disorder – A comparison of cognitive-behavioral therapy and psychodynamic therapy. Presentation at 49th SPR Internation Annual Meeting, Amsterdam, The Netherlands, 29.06.2018.

Paulick, J., Rubel, J., Schoenherr, D., Deisenhofer, A.-K., Schwartz, B., Terhürne, P., Altmann, U., Strauss, B., Lutz, W. (2018). Nonverbal synchrony in social phobia: A new approach for better understanding the relations between early response and treatment outcome. Presentation at 49th SPR Internation Annual Meeting, Amsterdam, The Netherlands, 29.06.2018.

Schwartz, B., Paulick, J., Schoenherr, D., Deisenhofer, A.-K., Rubel, J., Altmann, U., Strauss, B., Lutz, W. (2018). Nonverbal synchrony in social phobia: The attunement of patient and therapist. Presentation at 49th SPR Internation Annual Meeting, Amsterdam, The Netherlands, 29.06.2018.

Altmann, U., Schoenherr, D., Paulick, J., Deisenhofer, A.-K., Schwartz, B., Rubel, J., Lutz, W., Strauß, B. (2018). Nonverbal synchrony during psychotherapy of patients suffering from social anxiety disorder – A comparison of cognitive-behavioral therapy and psychodynamic therapy. Presentation at 51st Congress of DGP, Frankfurt, 18.09.2018.

- Schwartz, B., Paulick, J., Thielemann, D., Deisenhofer, A.-K., Rubel, J., Altmann, U., Strauss, B., Lutz, W. (2018). The attunement of patient and therapist and its relation to outcome in the treatment of social phobia. Presentation at 51st Congress of DGP, Frankfurt, 18.09.2018.
- Altmann, U., Schönherr, D., Paulick, J., Deisenhofer, A.-K., Schwartz, B., Rubel, J., Stangier, U., Lutz, W., Strauss, B. - Nonverbale Synchronie von Bewegungen und Therapieerfolg: Ein Vergleich von psychodynamisch-orientierter Therapie und Verhaltenstherapie bei sozialen Phobien. Vortrag auf dem Gemeinsamen Kongress der DGMP und DGMS, Leipzig, 28.09.2018.
- Schwartz, D., Schönherr, D., & Strauß, B. Grenzverletzungen und Grenzüberschreitungen als Ursache unerwünschter Effekte in der Psychotherapie, Vortrag, Gemeinsame Jahrestagung der DGMP und DGMS, Leipzig, 28.09.2018.
- Worrack, S., Schönherr, D., Guntinas-Lichius, O., Volk, G. F., Strauß, B., Altmann, U. Nonverbale Synchronie emotionaler Expressionen in gesundheitsrelevanten Interaktionen: Prädiktoren und Zusammenhänge mit der Patientenzufriedenheit. Vortrag, Gemeinsame Jahrestagung der DGMP und DGMS, Leipzig, 28.09.2018.
- Schwartz, D., Schönherr, D., Schleu, A., Freyberger, H. J., & Strauß, B. Grenzverletzungen und Grenzüberschreitungen als Ursache unerwünschter Effekte in der Psychotherapie, Vortrag, DGPPN Kongress, Berlin, 30.11.2018.

Ehrenwörtliche Erklärung

Hiermit erkläre ich, dass mir die Promotionsordnung der Fakultät für Sozial- und Verhaltenswissenschaft der Friedrich-Schiller-Universität Jena bekannt ist. Weiterhin versichere ich, dass ich die vorliegende Arbeit selbständig verfasst und keine anderen als die angegebenen Hilfsmittel und Quellen verwendet habe. Aus Quellen übernommene Inhalte wurden als solche gekennzeichnet. Die Eigenanteile der verschiedenen Publikationen sind in Tabelle 12 gekennzeichnet. Bei der Erstellung und Supervision der Publikation waren die in den Publikationen erwähnten Ko-Autoren beteiligt (Uwe Altmann, Bernhard Strauß, Susanne Worrack, Jane Paulick, Anne-Katharina Deisenhofer, Brian Schwartz, Julian Rubel, Ulrich Stangier und Wolfgang Lutz). Darüber hinaus wurde die Hilfe eines Promotionsberaters nicht in Anspruch genommen. Zudem haben Dritte weder unmittelbar noch mittelbar geldwerte Leistungen für Arbeiten erhalten, die mit dem Inhalt der Dissertation im Zusammenhang stehen.

Diese Dissertation liegt noch nicht als Prüfungsarbeit für eine staatliche oder andere wissenschaftliche Qualifikationsleistung vor. Darüber hinaus wurde die vorliegende Arbeit weder in gleicher noch in ähnlicher Weise an einer anderen Hochschule bzw. anderen Fakultät eingereicht.

Ich versichere, dass ich nach bestem Wissen die reine Wahrheit gesagt und nichts verschwiegen habe.

Ort, Datum

Unterschrift

Danksagung

Mein herzlicher Dank gilt Prof. Dr. Bernhard Strauß für die Betreuung und Begutachtung dieser Arbeit. Durch Gespräche auf Konferenzen, in Projektbesprechungen oder Feedback für erstellte Manuskripte lernte ich viel über die selbstständige Arbeit eines Forschers und die damit verbundenen Aufgaben. Weiterhin danke ich Prof. Dr. Wolfgang Tschacher für die Übernahme einer Gutachterposition. Ich habe die Gespräche über Synchronie immer als beflügelnd und inspirierend erlebt. Außerdem danke ich Prof. Dr. Peter Noack für die Übernahme des Gutachtens dieser Arbeit.

Innerhalb meiner Arbeitsgruppe möchte ich besonders Dr. Uwe Altmann danken, der mich für das Thema der Synchronie begeisterte und mir nicht nur in Programmierungsfragen und bei der statistischen Analyse eine helfende Hand war, sondern mit detailliertem Nachfragen und Feedback meinen Kenntnisstand über Synchronie stets voranbrachte. Weiterhin möchte ich mich bei allen Koautoren der Publikationen bedanken. Ich erlebte unsere Zusammenarbeit immer als sehr aufgeschlossen, fordernd und konstruktiv. Ein besonderer Dank gilt dabei Dr. Jane Paulick: In Gesprächen mit ihr konnte ich die Erkenntnisse der Forschung stets auf die Praxis münzen und neue Ideen entwickeln.

Weiterhin gilt mein Dank meinem kompletten Kollegenkreis: Ob ich PD Dr. Jenny Rosendahl zum Vorgehen bei der Berichterung von Effektstärken fragte oder gemeinsam mit Dr. Luise Bell, Dr. Felicitas Richter, Susanne Worrack und Carina Taubert neue Herangehensweisen oder Erklärungsmodelle für meine Ergebnisse diskutierte. Ich empfand diese Gespräche stets als fruchtbringend und aufschlussreich. Dr. Felicitas Richter möchte ich ebenfalls für das Korrekturlesen der vorliegenden Arbeit danken.

Ein weiterer großer Dank geht an meine Familie: Hättet ihr nicht so bedingungslos hinter mir gestanden, wäre mein Weg über die Erziehungswissenschaft zur Psychologie mit eingeschlossenem Auslandssemester wohl nicht so reibungslos verlaufen. Egal, welche Idee ich für mein Leben hatte, ihr fragtet nie „Warum?“ sondern standet immer hinter mir und vermitteltet mir so, dass ich jeden Weg, den ich mir aussuche, bestreiten kann. Ein herzlicher Dank geht dafür an meine Mutter Simone, meinen Vater Silvio, meine Schwester Janett und ihren Mann Martin und meine Großeltern Hanni & Edgar und Ruth & Heinz. Ein besonderer Dank geht an meine Schwester, dafür, dass sie diese Arbeit Korrektur gelesen hat, mich stets unterstützt hat und in den letzten Jahren zu einer besten Freundin geworden ist, mit der ich das Interesse an der Forschung als auch schöne Erlebnisse während meiner Promotionszeit teilen und erleben konnte.

Ich danke meinem gesamten Freundeskreis für die Unterstützung, das Zuhören und Ermutigen, wenn Frustrationen auftraten. Außerdem habe ich es genossen, euch von neuen Forschungsergebnissen zu berichten.

Doch mein wohl größtes Dankeschön geht an meinen Mann, Dr. Jan Schönherr. Ohne ihn hätte die Synchronie-Forschung wohl an mancher Stelle festgesteckt. Durch detailliertes Nachfragen, wohlgemeintes Feedback und dem Streben nach Perfektion bei größtmöglicher Effizienz, hat er mich dazu gebracht, in vieler Hinsicht noch über mich hinaus zu wachsen. Waren die stundenlangen Gespräche, bei denen er mich auch nach dem kleinsten Parameter eines Algorithmus‘ fragte, zwar manchmal nervenaufreibend, aber im Nachhinein stets hilfreich, voranbringend, inspirierend und strukturierend für meine Forschungstätigkeit. Auch danke ich ihm für das „Autoren-Auge“ sowie das „Forscher-Auge“ mit dem er die vorliegende Arbeit Korrektur gelesen hat und die emotionale Unterstützung, die er mich durch so viele Aufmerksamkeiten spüren ließ.

Vielen Dank!

Denn auch gute Forschung entsteht nur in Interaktionen und mit Kommunikation - wenn Menschen auf Menschen treffen und Synchronisationsphänomene erleben.