*Original Research*

DIGITAL HEALTH

Digital Health
Volume 8: 1–17
© The Author(s) 2022
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/20552076221129098
journals.sagepub.com/home/dhj
SAGE

# Movement-based patient-therapist attunement in psychological therapy and its association with early change

**Brian Schwartz[1], Julian A. Rubel[2], Anne-Katharina Deisenhofer[1] and Wolfgang Lutz[1]**

## Abstract

**Objective:** Attunement is a novel measure of nonverbal synchrony reflecting the duration of the present moment shared by two interaction partners. This study examined its association with early change in outpatient psychotherapy.

**Methods:** Automated video analysis based on motion energy analysis (MEA) and cross-correlation of the movement time-series of patient and therapist was conducted to calculate movement synchrony for $N = 161$ outpatients. Movement-based attunement was defined as the range of connected time lags with significant synchrony. Latent change classes in the HSCL-11 were identified with growth mixture modeling (GMM) and predicted by pre-treatment covariates and attunement using multilevel multinomial regression.

**Results:** GMM identified four latent classes: high impairment, no change (Class 1); high impairment, early response (Class 2); moderate impairment (Class 3); and low impairment (Class 4). Class 2 showed the strongest attunement, the largest early response, and the best outcome. Stronger attunement was associated with a higher likelihood of membership in Class 2 ($b = 0.313$, $p = .007$), Class 3 ($b = 0.251$, $p = .033$), and Class 4 ($b = 0.275$, $p = .043$) compared to Class 1. For highly impaired patients, the probability of no early change (Class 1) decreased and the probability of early response (Class 2) increased as a function of attunement.

**Conclusions:** Among patients with high impairment, stronger patient-therapist attunement was associated with early response, which predicted a better treatment outcome. Video-based assessment of attunement might provide new information for therapists not available from self-report questionnaires and support therapists in their clinical decision-making.

## Keywords

Motor mimicry, nonverbal synchrony, early response, growth mixture modeling, motion energy analysis

Submission date: 23 July 2022; Acceptance date: 11 September 2022

## Introduction

In social interactions, dyads tend to synchronize their neural, perceptual, affective, physiological, and behavioral responses.[1] This interpersonal synchrony is observable in breathing patterns, word use, heart rate variability, hormones, and electrodermal activity, and it is a marker for being involved in an interaction.[2–5] The involuntary imitation of facial expressions and gestures has already been demonstrated in mice and chimpanzees as well as in human newborns.[6–8] Besides animals and children, adults also mimic facial expressions, autonomous reactions, and movements in order to empathize with their counterparts and better understand their emotions, intentions, and behavior. Recently, a review compiled findings from biological,

[1]Department of Psychology, University of Trier, Trier, Germany
[2]Department of Psychology, Justus-Liebig-University Gießen, Giessen, Germany

**Corresponding author:**
Brian Schwartz, Department of Psychology, University of Trier, D-54286 Trier, Germany.
Email: schwartzb@uni-trier.de

psychological, and neuroscientific studies that support the assumption that autonomous mimicry is fundamental for the development of empathy.[9] Based on these findings, the authors developed the neurocognitive model of emotional contagion that assumes that an emotional state is expressed in nonverbal behavior and autonomous reactions. When the interaction partner perceives these behaviors and reactions, neural representations are activated as if the partner were experiencing the same state. In turn, this shared neural activation activates somatic and autonomic responses that manifest as motor and autonomic mimicry. Facilitated physiological and motor feedback then induces emotions in the interaction partner, enabling them to better understand their counterpart in terms of empathy.

Prochazkova and Kret[9] distinguished between autonomic mimicry relying on unconscious signals from the autonomic nervous system and motor mimicry relying on the partly unconsciously controlled motor muscles. Motor mimicry occurs when two interaction partners coordinate their movements synchronously. This function has been found to be impaired in patients with psychological disorders, e.g. social anxiety disorder.[10] Movement synchrony was associated with more positive and less negative affect,[11,12] more prosocial behaviors, perceived social bonding, and social cognition.[13] In psychotherapy, patients who were more strongly synchronized with their therapists showed a stronger therapeutic alliance, more self-efficacy, better treatment outcomes, and less attachment anxiety.[14–16] These patients were also less likely to end treatment prematurely and if they did drop out, they did so later in treatment.[17,18] In patients with schizophrenia, movement synchrony has been associated with fewer symptoms, higher social competence, social functioning, and self-evaluation of competence.[19,20] However, some effects of synchrony could not be consistently replicated (e.g. the association between synchrony and alliance[17]) and contradictory findings are also apparent. For example, Lutz et al. found lower synchrony to predict early change in interpersonal problems.[21]

Recently, automated video analysis systems were introduced, allowing the cost-effective and objective measurement of movement synchrony.[15] In contrast to previous research using human ratings of body movements, video analysis is less time-consuming and solves the problem of low interrater reliability. Automated video analysis captures body movements and their reference to one another, instead of relying on specific gestures or postures as human ratings did. A common approach to measuring movement synchrony is Motion Energy Analysis (MEA),[15] which quantifies the movements of interaction partners. MEA and another method to measure movement synchrony (namely pose-estimation) were shown to be highly consistent in their results,[22] which is why convergent validity of the measures can be assumed. Currently, research focuses on two measures of movement synchrony based on MEA. The first measure is the strength of the association between patients' and therapists' movements, i.e. how strongly their movements are correlated.[15,23] The second measure is the length of this association over time, i.e. how long patients and therapists move synchronously.[14,24]

A novel third measure is the variation of the temporal distance between synchronous movements, i.e. how flexible the latency between synchronous movements is. We refer to this measure as the movement-based attunement of patient and therapist. It is measured by the duration of nonverbal movement coordination, with larger movement-based attunement reflecting patient and therapist synchronizing their movements over a wider range of time lags between movements.[25] The basis of attunement is the therapist being present in the session, attuned and open to the patient, and it represents the duration of the present moment shared by two interaction partners.[26] Longer attunement is supposed to reflect more extended social contact between patient and therapist and higher therapist receptivity. It has been found to be associated with the interaction partners' personality traits, including greater openness to experiences, greater avoidant attachment, and less vindictive and self-centered interpersonal styles.[25]

Patient and therapist attunement might facilitate early improvement in psychotherapy. An established approach to identify early change patterns based on multiple measurements is growth mixture modeling (GMM), which is a latent variable cluster analytic method creating subgroups of patients with a similar change pattern.[27,28] It examines individual multi-point change trajectories and classifies patients by means of their trajectories. GMM adapts latent growth models and integrates a categorical latent variable for the identification of latent classes underlying the individual change trajectories. It was recently applied to identify latent classes of adaptation patterns in grief[29] and patterns of change in youth- and parent-rated anxiety symptoms.[30] Hilbert et al.[31] showed that early change trajectories outperformed common early response classifications (i.e. receiver operating characteristic curves and empirically based cut-offs) when predicting treatment outcomes in binge-eating disorder. So far, this kind of model has repeatedly identified a class of patients with high initial impairment and early improvement.[31–34] This phenomenon has been dubbed early response.

Studies on early change suggest that especially fast responding patients are able to maintain their initial success in that they show markedly positive ultimate outcomes. Early change patterns have been shown to be predictive of outcome and treatment termination in several disorders and populations. In a sample of depressed patients, early change predicted outcome at treatment termination as well as one and a half years after treatment.[35] Among college students treated at a university counseling center, early treatment response was associated with fewer symptoms at termination and follow-up as well as maintenance of gains for up to 24 months.[36] In a naturalistic study, depression and anxiety outcomes as well as treatment duration were predicted by early change.[34] Furthermore, early change has been shown to be predictive of outcome and early treatment termination in patients suffering from panic disorder.[32] In binge-eating disorder, remission at six-months follow-up as well as the binge-eating frequency at six- and 18-months follow-up were predicted by early change profiles.[31] Furthermore, differential patterns of response to the group and individual therapy have been identified and associated with treatment outcome and follow-up.[37] A recent meta-analysis concluded that there is reliable empirical evidence for early change predicting treatment outcome.[38]

In summary, automated video analysis of therapy sessions offers new opportunities for data collection and introduces novel measures that go beyond self-report questionnaires. Patient–therapist attunement is a promising construct, measuring the extent of the contact and understanding between patient and therapist. It seems suitable to predict early change patterns in psychological therapy. Early response, in turn, is an indicator of a favorable treatment course and its absence enables therapists to identify patients at risk of treatment failure early on. Therefore, this study aimed to examine whether movement-based attunement is associated with membership in an early change class. Furthermore, the long-term effects of the change classes on treatment outcomes were examined to replicate the previously found associations and demonstrate the clinical relevance of early response in the current data. The following hypotheses guided this study: (1) using GMM, growth curves of patients' symptomatic distress can be grouped into distinct latent early change classes, (2) stronger movement-based attunement is associated with a higher probability of early response, and (3) early response predicts better treatment outcome.

## Methods

### Patients

The study is based on a naturalistic sample of $N = 161$ patients treated for at least five sessions at a university outpatient clinic. Eighty-seven patients were female (54%) and their ages ranged from 16 to 68 years ($M = 36.41$, $SD = 12.48$). Diagnoses were assessed before treatment onset with the Structured Clinical Interview for DSM-IV-TR Axis I Disorders (SCID-I)[39] and the International Diagnostic Checklist for Personality Disorders (IDCL-P).[40] Trained independent clinicians with at least one year of clinical experience conducted the interviews. They were videotaped and diagnoses were discussed in expert consensus teams that included four senior clinicians. Final diagnoses were determined by the consensual agreement of at least 75% of the team members. Primary diagnoses as per SCID-I included affective disorders (59.4%), anxiety disorders (18.2%), stress and adjustment disorders (9.4%), eating disorders (1.9%) as well as other diagnoses (11.1%). Additionally, 21.1% ($n = 34$) fulfilled the criteria for a personality disorder. Of the total sample, 65.8% had at least one further diagnosis, while 32.9% had at least two comorbid diagnoses. Table 1 provides further sample characteristics. Written informed consent was obtained before the start of treatment. All procedures contributing to this work comply with the ethical standards of the relevant national and institutional committees on human experimentation and with the Helsinki Declaration of 1975, as revised in 2008.

**Table 1.** Pre-treatment sample characteristics.

| | *M* or *n* (*SD* or %) |
|---|---|
| *Socio-demographic aspects* | |
| Gender (female) | 87 (54.00) |
| Age | 36.41 (12.48) |
| Education (> 12 years) | 74 (46.00) |
| Unemployment | 28 (17.30) |
| Unable to work | 28 (17.40) |
| In committed relationship[a] | 93 (57.80) |
| *Pre-therapy assessment* | |
| HSCL-11 | 2.14 (0.59) |
| Chronicity | 5.25 (1.04) |
| Prior psychotherapy | 2.53 (1.78) |
| Treatment expectation | 3.06 (0.72) |
| GAF | 56.82 (9.51) |
| *Treatment* | |
| Length (number of sessions) | 37.55 (16.97) |
| Duration (weeks) | 74.58 (34.38) |
| Dropout[b] | 32 (19.90) |

HSCL-11: short-form of the Hopkins Symptom Checklist; GAF: Global Assessment of Functioning.
[a]Married or longstanding partnership.
[b]Therapists evaluated whether treatment ending was consensual or patients dropped out prematurely.

### Therapists and treatment

The treatment consisted of an integrative cognitive behavioral therapy (CBT) approach. According to German health care requirements, therapists were obligated to provide case formulations at the beginning of treatment. All case formulations were examined by independent surveyors (commissioned by health insurance companies), who endorsed the suggested treatment strategies as *lege artis* CBT interventions. Treatment was provided by 39 therapists who treated 1 to 15 patients ($M = 4.13$, $SD = 3.69$). All therapists participated in a three-year (full-time) or five-year (part-time) postgraduate training program, had at least one year of clinical training before participating in the study and were supervised by senior clinicians. They were familiar with treatment manuals, though not constrained to follow a strict protocol. The attending therapists took over the patients in the third session after the initial screening and clinical interview, which were conducted by independent clinicians. Treatment consisted of weekly sessions whose number was not fixed and varied between six and 84 sessions ($M = 37.55$, $SD = 16.97$, $Mdn = 32.00$).

## Video selection

Therapy sessions were routinely recorded at the outpatient clinic for supervision and research. Two synchronized cameras focused on the patient and therapist respectively and generated split-screen recordings. A static camera position, stable light conditions as well as digitized film material ensured video quality. For each patient, one video from the first third of the therapy (but not exceeding the tenth session) was randomly chosen. Videos were unsuitable for video analysis and excluded if they were of low video quality (e.g. major lighting changes or image errors), patient or therapist did not remain seated during the video, or one person's movements reached into the other person's split screen. Session numbers of selected videos ranged from the first to the seventh session with the attending therapist.

Only the first 15 minutes of each video were analyzed, as later in the sessions patients or therapists more often left their seats and their interaction was interrupted by exercises, role-plays or the use of whiteboards. Former studies have demonstrated high correlations between movement synchrony over the whole 50-minute session and during the session's first 15 minutes ($r \geq 0.80$).[15,17]

## Movement quantification

Movement quantity was measured separately for patient and therapist via MEA.[15,41] MEA is an objective and automated video-analysis algorithm. It detects the video frame rate and computes the change in grey-scale pixels between two consecutive video frames. The difference in pixels was computed for a region of interest, which covered the upper body beginning at the seat of the chairs, because the lower body was often hidden by the table. In this study, the frame rate was ten frames per second, resulting in a time series of 9000 movement estimates for one person in a 15-minute video sequence. MEA generates movement quantity time-series, but does not capture movement quality, i.e. semantic content.

## Nonverbal synchrony

Following Schoenherr et al.'s recommendations, time series of patient and therapist movements were logarithmized and not smoothed.[42] Windowed cross-lagged correlation with overlapping windows, a window width of 50 frames (i.e. five seconds), and a time lag ranging from −50 frames (therapist moves first, patient follows after five seconds) to +50 frames (patient moves first, therapist follows after five seconds) was conducted in steps of one frame. The absolute values of Pearson correlations in the 8951 windows were averaged using Fisher's Z-transformation, resulting in one correlation coefficient per time lag. Figure 1A shows a case example of the correlations between patient and therapist movement times-series for the 100 different time lags.

To control for coincidentally occurring synchrony, a nonparametric significance test relying on so-called pseudo-interactions was implemented.[15] The patient's time-series was cut into five-second segments, which were shuffled randomly 50 times to obtain 50 surrogate time-series. The windowed cross-lagged correlation analysis was repeated for each surrogate patient time-series and original therapist time-series. The resulting mean correlation for each time lag relies on artificial combined movements that never appeared in this order, thus showing no real interaction behavior. All coincidentally found correlations are random and rely on the individual movement patterns of the two interaction partners. The 50 estimates of pseudo-interaction were averaged per time lag to obtain a stable estimation of pseudo-synchrony (Figure 1B). Synchronization exceeding averaged pseudo-synchrony was assumed significant. While parametric significance tests only test against the assumption of a zero correlation, this alternative tests more conservatively against the baseline synchronization of two individuals.

## Movement-based attunement

The attunement of a dyad was measured as the range of connected time lags with significant synchrony around time lag 0 (Figure 1C). It represents the range of time distances between patients' and therapists' synchronous movements. Due to the restriction on time lags to vary between plus and minus five seconds, movement-based attunement was able to vary between 0 seconds and 10 seconds. For example, for a dyad that synchronizes significantly with a positive time lag of 0.1 s (i.e., one frame) up to a time lag of 0.7 seconds and with a negative time lag of −0.1 seconds up to −0.8 seconds, attunement amounts to 1.5 seconds (Figure 1C).

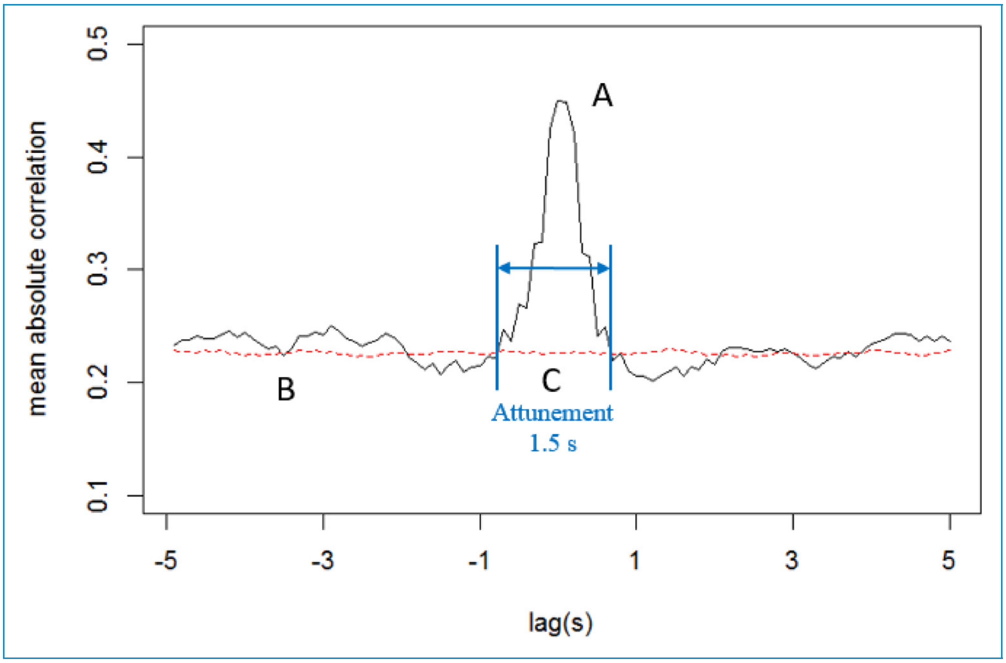


**Figure 1.** Movement synchrony, pseudosynchrony, and attunement for one exemplary patient. (A) The black line represents the mean absolute correlation between movement time-series of patient and therapist (*y*-axis) for 100 time lags (−5 to + 5; *x*-axis in seconds). (B) The red dashed line represents the average of 50 estimates of pseudosynchrony (mean absolute correlation between patient surrogate time-series and original therapist time-series; *y*-axis) for 100 time lags (*x*-axis). (C) The blue two-sided arrow represents the amount of attunement (i.e. the range of connected time lags with synchrony exceeding pseudosynchrony around time lag 0).

## Psychometric measures

*Hopkins symptom checklist–short form (HSCL-11).* The HSCL-11[43] is an 11-item short-form of the Symptom-Checklist-90-Revised (SCL-90-R)[44] that is based on the HSCL-25.[45] It assesses patients' global symptom distress on a self-rated four-point Likert scale ranging from one ("not at all") to four ("extremely"). The mean of the 11 items represents the patient's level of global symptomatic distress for the preceding week. It has a high internal consistency ($\alpha = .92$) and is highly correlated with the global severity index of the Brief Symptom Inventory ($r = .91$).[43] Internal consistency in the current sample was $\alpha = .89$ (.86, .91) for the first assessment. The HSCL-11 was routinely assessed before each session and the mean scores at sessions 1 to 10 with the attending therapist were used to model early change. Furthermore, the first HSCL-11 score was used as a covariate.

*Pre-treatment covariates.* Pre-treatment variables that were previously found to be associated with improvement on the HSCL-11[46,47] and for which previous analyses of early change were adjusted[48] were used as covariates. Pre-treatment covariates included chronicity, prior psychotherapy, patient-rated treatment expectation, and global functioning. They were taken from the Compass Tracking System, originally called the Integra Outpatient Treatment Assessment System.[49,50] This system is an early comprehensive assessment battery used to measure progress in outpatient psychotherapy. The covariates are also implemented in the comprehensive feedback and decision-support system currently routinely employed in the outpatient clinic.[47,51] All four variables were assessed with one item each. Chronicity ("How long has the problem for which you are presently seeking treatment been a concern to you?") was assessed on a six-point Likert scale from one ("less than a month") to six ("more than two years"). Prior psychotherapy ("How much psychotherapy have you had in the past?") was assessed on a six-point Likert scale ranging from one ("none") to six ("more than one year"). Treatment expectation ("How confident are you that psychotherapy will be successful in helping you with your problems?") was assessed on a four-point Likert scale ranging from one ("not at all confident") to four ("very confident"). Therapist-rated global functioning was assessed with the Global Assessment of Functioning scale (GAF),[52] a visual analog scale ranging from 1 ("constant risk of serious injury to self or others, or persistent inability to maintain minimal personal hygiene, or serious suicide attempt with a clear death wish") to 100 ("excellent performance in a wide range of activities; difficulties in life never seem to get out of control; valued by others for a variety of positive qualities; no symptoms").

## Statistical analyses

*Identifying latent classes.* Following previous studies on early change in psychotherapy, growth mixture modeling

(GMM) was applied to measure early treatment change.[32,33] Compared to simple difference values or clinically significant change criteria, GMM takes the entire course of change and not only two time points into account.[33] Furthermore, it is flexible with regard to the amount of change and intake functioning. GMM has been shown to be useful for the identification of early change patterns.[21,31]

We used the first 10 sessions with the attending therapist to model early change. Referring to the dose-effect-model of psychotherapy,[53] a logarithmic trend in the data was assumed. Mean growth curves for latent change classes as well as individual variations around these means were estimated. Variances were allowed to vary between classes. In GMM, global symptom distress assessed with the HSCL-11 was regressed on the logarithmized session number. Regression intercept and slope were calculated and the patients were grouped into latent classes based on these two coefficients. The number of classes was defined empirically using the AIC and sample size adjusted BIC as fit indices, while no class was allowed to consist of less than five percent of all cases. Furthermore, we compared the $k$-class solution with the $k$–1-class solution using the Bootstrapped Likelihood Ratio Test (BLRT). Latent class analyses were conducted with the software environment Mplus version 8.4,[54] while all further analyses were conducted with R version 4.0.1.[55]

*Data preparation for prediction models.* For all covariates and the HSCL-11 at session 10, missingness varied between 0.62% ($n = 1$; HSCL-11) and 8.07% ($n = 13$; GAF). Missing values of all covariates (first HSCL-11, chronicity, prior psychotherapy, treatment expectation, and GAF) as well as the HSCL-11 at session 10 with the attending therapist were imputed using the R package missForest version 1.4.[56] MissForest provides a nonparametric missing value imputation based on random forest that is able to handle mixed-type data for up to 30% missing values.[57] To improve imputation accuracy, movement-based attunement, latent class membership, last HSCL-11, therapist id, and number of sessions contributed additional information to the imputation procedure, but were not imputed themselves. The imputation was evaluated using the normalized root mean squared error (NRMSE).[57] In the current data, imputation was successful with an out-of-bag error of NRMSE = 0.16, which can be interpreted as a good performance of MissForest. After imputation, all predictor variables (attunement and all covariates) were grand-mean centered.

*Predicting early change.* A baseline-category logit model for multinomial responses with random effects was used to predict early change by regressing latent change classes on potential predictor variables. The random effect model (patients nested within therapists) with a random intercept and fixed slopes was estimated using the marginal quasi-likelihood (MQL) technique based on a Solomon-Cox approximation. MQL is recommended when the marginal relationships between covariates and outcome are of interest, while the penalized quasi-likelihood method is preferable when estimating random effects that are not the focus of the analyses.[58] The iterative procedure was run with a positive convergence tolerance of epsilon = 1e–08, which is the default in several control functions in R,[55,59,60] and a maximum of 200 Fisher Scoring iterations (the R package's default of a maximum of 25 iterations was increased to 200 iterations to ensure convergence).

In a first step, movement-based attunement was examined as a bivariate predictor for early change (model 1). To ensure that attunement predicted the change slopes and not only pre-treatment impairment of latent classes (intercepts), in a second step, the HSCL-11 at session 1 with the attending therapist was added to the model (model 2). Finally, to examine the effect of movement-based attunement on early change beyond pre-treatment predictors, the analysis was additionally adjusted for the assessed covariates that predicted HSCL-11 improvement in previous studies (chronicity, prior psychotherapy, treatment expectation, and GAF; model 3). As multinomial regression analysis with a non-dichotomous categorical criterion requires a reference category, all other classes are compared to this reference category only. Therefore, to compare each class with each other class, three analyses with different reference categories were conducted and results were integrated.

To examine the predictions of the model and thus evaluate their applicability and clinical usefulness, the estimated probabilities of belonging to one of the classes were plotted as a function of movement-based attunement. This graphical representation was performed for three groups of patients, namely for patients with an initial impairment above the sample average (HSCL-11 of 1 *SD* above the mean), for patients with an average impairment (HSCL-11 equal to the sample mean), and for patients with a below-average impairment (HSCL-11 of 1 *SD* below the mean). Multinomial regression analysis was conducted using the R package mclogit[59] and probabilities were plotted using the R package ggplot2.[61]

*Comparing early change classes.* For each early change class, two percentage changes were calculated. Percentages of change have been shown to be less influenced by initial impairment than other outcomes, e.g. post-treatment scores, pre-treatment to post-treatment differences or standardized individual effect sizes.[62] First, the early percentage change was calculated by subtracting HSCL-11 at session 10 from HSCL-11 at session 1 and dividing it by HSCL-11 at session 1 with the attending therapist. Second, the percentage change from pre- to post-treatment was assessed by dividing the pre-post difference on the HSCL-11 by the pre-treatment HSCL-11; pre-treatment corresponded to session 1 with the attending therapist, while post-treatment corresponded to the last observation carried forward. Percentage changes until session 10 and until the end of treatment were compared between classes using linear mixed models (patients nested within therapists) with a random intercept and fixed slopes regressing percentage change on class membership and all covariates. Since the categorical predictor class membership had to be dummy-coded, all other classes were compared to the reference category only. Therefore, again, three analyses with different reference categories were conducted and results were integrated. Additionally, treatment length measured by the number of sessions was regressed on class membership in a linear mixed model with a random intercept and fixed slopes. Linear mixed models were conducted using the R packages lme4[63] and lmerTest.[64]

## Results

### Early change classes

Growth mixture modeling revealed one model with the best fit indices, while no class consisted of less than 5% of all cases. The fit indices (Loglikelihood, AIC, sample-size adjusted BIC) improved with each additional latent class. However, only four classes were identified that included at least 5% of all patients, while the fifth class consisted of only two patients (1.24% of all cases). Additionally, the BLRT showed significant advantages in model fit for up to four classes, while the model with five classes did not fit the data significantly better than the model with four classes ($p = .122$). Therefore, the best fitting model was the model that assumed four latent classes. Class 1 included initially highly impaired patients who did not improve until session 10 ($n = 16$). Class 2 was comprised of patients who were also highly impaired in the first session with the attending therapist, but who clearly improved during the first 10 sessions ($n = 27$). Class 3 included moderately impaired patients who hardly changed ($n = 54$), while Class 4 was made up of patients with low impairment who did not change during the first 10 sessions ($n = 64$; Figure 2).

Classification probabilities for the most likely latent class membership ranged from .885 to .938. The classes did not differ significantly in the average session number analyzed (Class 1: $M = 3.63$; Class 2: $M = 4.30$; Class 3: $M = 4.30$; Class 4: $M = 4.09$; $F(3,157) = 0.893$, $p = .446$, $\omega^2 = -.002$). Table 2 presents the fit indices, results of the BLRT, and the number of patients in the smallest class for GMM solutions with one to five classes.

### Predicting early change with movement-based attunement

Movement-based attunement ranged between 0.4 and 3.3 seconds, with an average of 1.49 seconds. Average

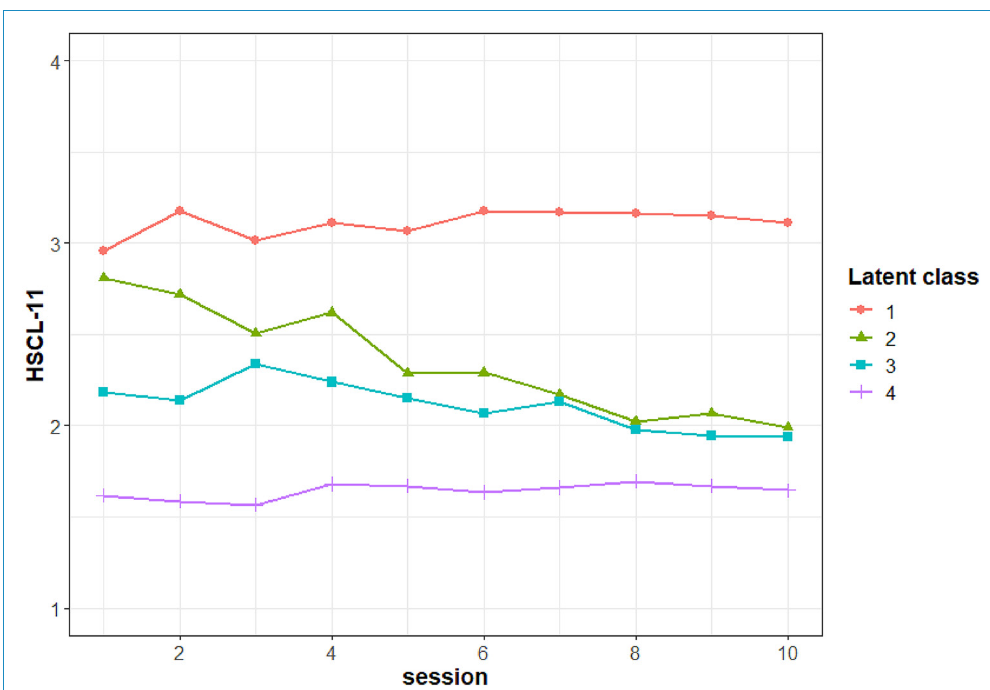


**Figure 2.** HSCL-11 symptom severity as a function of session for latent classes. Observed values of the HSCL-11 regressed on session in a growth mixture model (GMM); HSCL-11: Hopkins Symptom Checklist-11; Class 1: high impairment, no change; Class 2: high impairment, early response; Class 3: moderate impairment; Class 4: low impairment.

attunement was 1.13 seconds for Class 1 (high impairment, no early change), 1.67 seconds for Class 2 (high impairment, early response), 1.44 seconds for Class 3 (moderate impairment), and 1.54 seconds for Class 4 (low impairment). Distributions of movement-based attunement across latent change classes are depicted in Figure 3. All distributions had their mode at roughly 1.2 seconds. Class 2's distribution appears bimodal; with another peak at about 2 seconds. Class 1 showed less variance than the other classes and fewer patients with an attunement above 2 seconds, while variance was greatest in Class 2.

**Table 2.** Fit indices for GMM solutions with different numbers of classes.

| # of classes | LL | AIC | adj. BIC | BLRT *p*-value | min. *n* (%) |
|---|---|---|---|---|---|
| 1 | −795.441 | 1624.882 | 1623.449 | - | 161 (100) |
| 2 | −783.870 | 1613.74 | 1611.801 | .013 | 25 (15.53) |
| 3 | −772.296 | 1602.593 | 1600.148 | < .001 | 16 (9.94) |
| 4 | −760.700 | 1591.401 | 1588.45 | .013 | 16 (9.94) |
| 5 | −752.108 | 1586.217 | 1582.76 | .122 | 2 (1.24) |

GMM: Growth Mixture Modeling; #: number; LL: Loglikelihood H0 value; AIC: Akaike Information Criterion; adj. BIC: sample-size adjusted BIC ($n^* = (n + 2)/24$); BLRT: Bootstrapped Likelihood Ratio Test; min. *n*: number of patients in the smallest class.

In the bivariate multinomial regression models with movement-based attunement as a single predictor (model 1), patients in patient-therapist dyads with stronger movement-based attunement were more likely to belong to the highly impaired early responders (Class 2; $b = 0.249$, $p = .004$), the moderately impaired patients (Class 3; $b = 0.180$, $p = .027$), and the low impaired patients (Class 4; $b = 0.196$, $p = .015$) than to the highly impaired patients who did not change during the first 10 sessions (Class 1; Figure 2). The likelihood of belonging to Classes 2, 3 and 4 was not significantly associated with movement-based attunement (Table 3). The model converged after five Fisher Scoring iterations.

When adjusting for pre-treatment HSCL-11 (model 2), the model converged after eight iterations. The results remained stable with significant differences between Class 1 (high impairment, no change) and all other classes, while the other classes did not differ significantly from each other (Table 3). Additionally, the inclusion of all remaining covariates (chronicity, prior psychotherapy, therapist-rated treatment expectation, and GAF) in the model (model 3) did not change the results meaningfully; i.e. all significant effects remained significant. Convergence was achieved after nine Fisher Scoring iterations. While adjusting for all covariates, stronger movement-based attunement was associated with a higher likelihood to belong to Class 2 (high

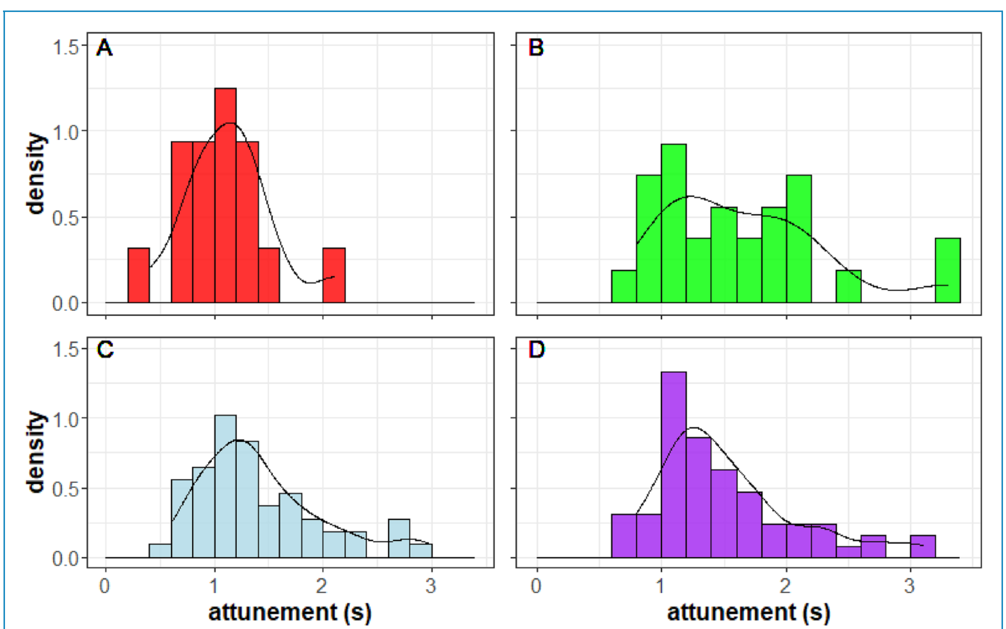


**Figure 3.** Distribution of movement-based attunement across latent classes. Attunement was measured in seconds. A: Class 1 ($n = 16$); B: Class 2 ($n = 27$); C: Class 3 ($n = 54$); D: Class 4 ($n = 64$).

impairment, early response; $b = 0.313$, $p = .007$), Class 3 (moderate impairment; $b = 0.251$, $p = .033$), and Class 4 (low impairment; $b = 0.275$, $p = .043$) compared to Class 1 (high impairment, no early improvement). Again, class membership for Classes 2, 3, and 4 was not significantly associated with movement-based attunement (Table 3).

For patients with an initial impairment of 1 *SD* below the sample mean (i.e. centered HSCL-11 = −0.59), the multilevel multinomial regression model predicted membership to Class 4 (low impairment; probability > .90), regardless of movement-based attunement, if average values were assumed for all covariates (Figure 4). Similarly, for patients with an average initial impairment (i.e. centered HSCL-11 = 0), the final model predicted membership to Class 3 (moderate impairment). While the probabilities changed slightly as a function of attunement, Class 3 was by far the most likely at each level of movement-based attunement ($p > .70$). However, for patients with an initial impairment of 1 *SD* above the sample mean (i.e. centered HSCL-11 = 0.59), the predicted probability of Class 1 (high impairment, no early response) decreased from > .80 (low attunement) to zero (strong attunement) as a function of movement-based attunement. At the same time, the probability of Class 2 (high impairment, early response) increased from < .10 (low attunement) to > .90 (strong attunement) as a function of movement-based attunement.

### Outcomes for latent classes

*Early change.* Percentage change during the first 10 sessions was 0.06 on average ($SD = 0.26$) and ranged from a deterioration of −1.00 to an improvement of 0.57. Patients with high initial impairment and early response (Class 2) showed the largest early response with an average change of 28.96% of pre-treatment impairment. Class 2 had a significantly larger percentage change until session 10 compared to Class 1 (high impairment, no change; −7.81%; $b = 0.380$, $p < .001$), and Class 4 (low impairment; −3.54%; $b = 0.196$, $p = .007$), but did not differ significantly from Class 3 (moderate impairment; 10.30%; $b = 0.112$, $p = .083$). Highly impaired patients from Class 1 also showed significantly less early change compared to patients from Class 3 (moderate impairment; $b = -0.268$, $p < .001$). All other pairwise comparisons did not reach statistical significance (Table 4).

*Treatment outcome.* Percentage change until end of treatment was 0.16 on average ($SD = 0.27$) and ranged from −1.00 to 0.70. Patients in Class 2 (high impairment and early response) showed the largest percentage change until the end of treatment (29.38% of pre-treatment impairment). Their average percentage change was significantly larger than the percentage change of patients in Class 1 (high impairment, no early change; 8.99%; $b = 0.229$, $p = .005$). However, their outcome did not differ significantly from the average percentage change in Classes 3 (moderate impairment; 19.12%; $b = 0.070$, $p = .346$) and 4 (low impairment; 9.75%; $b = 0.106$, $p = .295$). Highly impaired patients with no early change (Class 1) also differed significantly from patients with moderate (Class 3; $b = 0.299$, $p < .001$) and low impairment (Class 4; $b = 0.334$, $p = .004$). All other comparisons were not significant (Table 4).

*Treatment length.* The average number of sessions was $M = 38.94$ ($SD = 22.39$) for Class 1, $M = 40.89$ ($SD = 18.88$) for Class 2, $M = 36.39$ ($SD = 15.24$) for Class 3, and $M = 36.78$ ($SD = 16.21$) for Class 4. None of the pairwise differences were statistically significant (all $p > .25$). Table 5 presents average movement-based attunement, session number, early response and outcome for all four classes as well as for the total sample.

## Discussion

The study aimed to investigate movement-based patient–therapist attunement as a correlate of early change in outpatient psychotherapy. Attunement refers to the duration of the present moment shared by two interaction partners and was measured by the range of the temporal distance between synchronous movements. Overall, average movement-based attunement was much shorter than reported in previous studies. Tschacher et al.[25] found an average attunement of six seconds in healthy same-sex dyads. Looking at the variance in attunement, the authors found a large range of 1 to 10 seconds in the cooperative condition, which should correspond most closely to a therapeutic dyad. The sample (healthy adults vs outpatients) as well as the dyads (same-sex dyads vs same-sex and mixed patient–therapist dyads) are possible reasons for discrepancies in means. However, in another study, Tschacher et al.[65] also found a comparable duration of 5.7 seconds in conversing patient–therapist dyads. Another important influencing factor could be the setting. While the videos of the present study were recorded under routine conditions in an everyday therapy room with less conspicuous cameras mounted in the upper third of the walls, Tschacher et al.'s[65] recordings were made in a laboratory setup for the measurement of nonverbal synchrony. Furthermore, as the measurement of movement-based attunement is based on movement synchrony and there is no agreed upon procedure on how to measure it in detail, differences measuring and calculating movement synchrony may have affected the duration of the attunement measure. For our calculations, we followed recommendations by Schoenherr et al.[42] who empirically tested different algorithm specifications for the first time.

Using GMM, a class of highly impaired patients who improved early in treatment was identified (Class 2). This is in line with previous studies that usually found this class of early responders.[31,32,35] Furthermore, a second class of patients with high impairment was evident that

**Table 3.** Multilevel multinomial regression models predicting class membership.

| | Model 1 | | Model 2 | | Model 3 | |
|---|---|---|---|---|---|---|
| | estimate | *p*-value | estimate | *p*-value | estimate | *p*-value |
| *Class 1 vs Class 2* | | | | | | |
| Intercept | 0.921 | .090 | 2.040 | .040 | 2.047 | .047 |
| Attunement | 0.249 | .004 | 0.296 | .010 | 0.313 | .007 |
| HSCL-11 | | | −1.318 | .200 | −1.162 | .278 |
| Chronicity | | | | | −0.229 | .557 |
| Prior psychotherapy | | | | | −0.236 | .261 |
| Treatment expectation | | | | | −0.249 | .638 |
| GAF | | | | | −0.013 | .765 |
| *Class 1 vs Class 3* | | | | | | |
| Intercept | 1.634 | < .001 | 4.531 | < .001 | 4.757 | < .001 |
| Attunement | 0.180 | .027 | 0.257 | .028 | 0.251 | .033 |
| HSCL-11 | | | −6.807 | < .001 | −7.759 | < .001 |
| Chronicity | | | | | 0.024 | .954 |
| Prior psychotherapy | | | | | 0.018 | .937 |
| Treatment expectation | | | | | 0.437 | .507 |
| GAF | | | | | −0.004 | .934 |
| *Class 1 vs Class 4* | | | | | | |
| Intercept | 1.734 | < .001 | 2.832 | .007 | 2.545 | .029 |
| Attunement | 0.196 | .015 | 0.332 | .011 | 0.275 | .043 |
| HSCL-11 | | | −14.797 | < .001 | −17.260 | < .001 |
| Chronicity | | | | | 0.479 | .399 |
| Prior psychotherapy | | | | | 0.158 | .602 |
| Treatment expectation | | | | | 1.508 | .081 |
| GAF | | | | | 0.052 | .462 |
| *Class 2 vs Class 3* | | | | | | |
| Intercept | 0.798 | .011 | 2.539 | < .001 | 2.936 | < .001 |

(continued)

**Table 3.** Continued.

| | Model 1 | | Model 2 | | Model 3 | |
|---|---|---|---|---|---|---|
| | estimate | *p*-value | estimate | *p*-value | estimate | *p*-value |
| Attunement | −0.050 | .237 | −0.035 | .514 | −0.064 | .279 |
| HSCL-11 | | | −5.557 | < .001 | −7.094 | < .001 |
| Chronicity | | | | | 0.336 | .304 |
| Prior psychotherapy | | | | | 0.308 | .155 |
| Treatment expectation | | | | | 0.879 | .122 |
| GAF | | | | | 0.007 | .851 |
| *Class 2 vs Class 4* | | | | | | |
| Intercept | 0.884 | .003 | 0.842 | .220 | 0.724 | .396 |
| Attunement | −0.034 | .403 | 0.043 | .589 | −0.036 | .691 |
| HSCL-11 | | | −13.536 | < .001 | −16.555 | < .001 |
| Chronicity | | | | | 0.798 | .112 |
| Prior psychotherapy | | | | | 0.429 | .141 |
| Treatment expectation | | | | | 1.929 | .016 |
| GAF | | | | | 0.066 | .302 |
| *Class 3 vs Class 4* | | | | | | |
| Intercept | 0.151 | .559 | −1.687 | < .001 | −2.153 | < .001 |
| Attunement | 0.036 | .353 | 0.082 | .189 | 0.028 | .689 |
| HSCL-11 | | | −8.130 | < .001 | −9.499 | < .001 |
| Chronicity | | | | | 0.436 | .254 |
| Prior psychotherapy | | | | | 0.141 | .466 |
| Treatment expectation | | | | | 1.110 | .052 |
| GAF | | | | | 0.050 | .326 |

Class 1: high impairment, no improvement; Class 2: high impairment, early improvement; Class 3: moderate impairment; Class 4: low impairment; HSCL-11: short-form of the Hopkins Symptom Checklist; GAF: Global Assessment of Functioning.

did not change during the first 10 sessions (Class 1) as well as one class with low (Class 4) and one with moderate initial impairment (Class 3). The results of GMM have so far not been generalizable across different studies and early patterns of change vary depending on the sample,[66] which is why they must always be considered in their context. For example, in Koffmann's[66] data the one-class solution was shown to be the best-fitting. In all other studies on GMM known to us, however, an early responder class could be identified.

When predicting early change classes, movement-based attunement was only able to differentiate between change

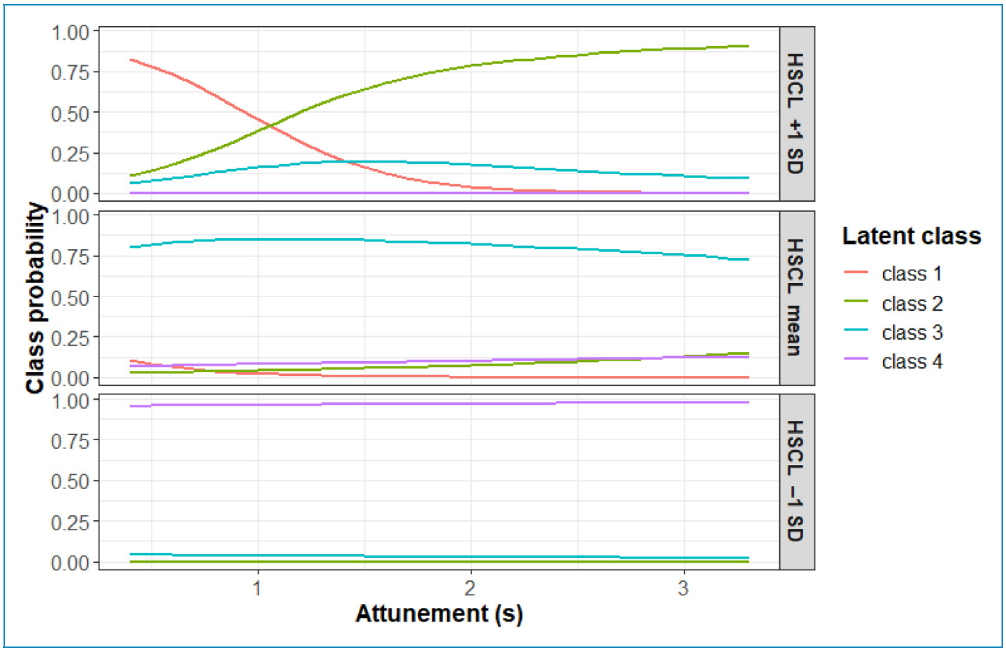


**Figure 4.** Probability of class membership as a function of attunement for mildly, moderately, and severely impaired patients. Probability of class membership (Class 1: high impairment, no early change; Class 2: high impairment, early change; Class 3: moderate impairment; Class 4: low impairment) based on the multilevel multinomial regression model with class 1 as the reference category. All other covariates were held constant at an average value of 0 (covariates were grand-mean centered). The values of the initial HSCL-11 correspond to an above-average (−1 standard deviation), average (mean) or below-average level of impairment (+1 standard deviation). Attunement was measured in seconds.

patterns of highly impaired patients, while change classes of patients with low or moderate initial impairment could not be differentiated. Among initially highly impaired patients, a stronger movement-based attunement significantly increased the probability of belonging to the class of early responders (Class 2) compared to the class of patients who did not improve during the first 10 sessions (Class 1).

Attunement seems to be particularly important for patients with high initial impairment, who may be unable to immediately engage in problem-solving work, but may profit from the therapist first applying a more motivational or relational strategy.[47,67] In contrast, for less impaired patients, attunement with the therapist may not be as important, as they are able to work on the problem directly. Since therapists show greater differences in effectiveness with highly impaired patients, whereas therapists do not differ for patients with low impairment,[68] it can also be assumed that attunement with the therapist plays a greater role in the treatment of more highly impaired patients. Change on the HSCL-11 from sessions 1 to 10 was significantly less in highly impaired patients without early change than in early responders and moderately impaired patients, but not compared to patients with low impairment. This seems due to the fact that patients with low impairment also showed no early change, while early responders and moderately impaired patients both showed meaningful improvements. This might also be the reason for early change in early responders not differing significantly from early change in moderately impaired patients, although change rates differed descriptively.

Furthermore, patient outcomes in the class with high impairment and no early change (Class 1) differed from all other classes when adjusted for initial impairment and covariates, showing the smallest percentage change until end of treatment. This indicates that primarily patients at risk of treatment failure were identifiable via early change patterns. However, although their treatment outcome did not differ significantly from low and moderately impaired patients, early responders showed the largest percentage change from pre- to post-treatment. Taking these results together, patients at risk for treatment failure are already identifiable early in therapy based on initial impairment and movement-based attunement. Highly impaired patients who are not attuned to their therapists are unlikely to respond meaningfully to treatment, while highly impaired patients whose dyad shows stronger movement-based attunement are more likely to respond early in treatment and to achieve meaningful improvement until the end of treatment. Remarkably, early responders showed the same average amount of change during the first ten sessions as until the end of treatment. As the average number of sessions did not differ significantly between classes, this does not seem due to shorter therapies among these patients. Instead, they seem to improve so much early in treatment that there is little

**Table 4.** Multilevel linear models predicting early response and treatment outcome.

| | Early response | | Outcome | |
|---|---|---|---|---|
| | estimate | *p*-value | estimate | *p*-value |
| *Model 1 (reference: Class 1)* | | | | |
| Intercept | −0.157 | .033 | −0.110 | .188 |
| HSCL-11 | 0.133 | .028 | 0.263 | < .001 |
| Chronicity | −0.034 | .055 | −0.045 | .025 |
| Prior psychotherapy | −0.013 | .222 | −0.014 | .259 |
| Treatment expectation | 0.016 | .528 | 0.005 | .871 |
| GAF | 0.001 | .561 | −0.000 | .990 |
| Class 1 vs Class 2 | 0.380 | < .001 | 0.229 | .005 |
| Class 1 vs Class 3 | 0.268 | < .001 | 0.299 | < .001 |
| Class 1 vs Class 4 | 0.184 | .064 | 0.334 | .004 |
| *Model 2 (reference: Class 2)*[a] | | | | |
| Intercept | 0.223 | < .001 | 0.118 | .081 |
| Class 2 vs Class 3 | −0.112 | .083 | 0.070 | .346 |
| Class 2 vs Class 4 | −0.196 | .025 | 0.106 | .295 |
| *Model 3 (reference: Class 3)*[a] | | | | |
| Intercept | 0.111 | < .001 | 0.188 | < .001 |
| Class 3 vs Class 4 | −0.084 | .112 | 0.036 | .559 |

Early response refers to percentage change on the HSCL-11 from session 1 to 10 with the attending therapist, outcome refers to the percentage change on the HSCL-11 from session 1 with the attending therapist to the last available observation; HSCL-11: short-form of the Hopkins Symptom Checklist; GAF: Global Assessment of Functioning; Class 1: high impairment, no improvement; Class 2: high impairment, early improvement; Class 3: moderate impairment; Class 4: low impairment.
[a]For models 2 and 3, estimates and *p*-values of all covariates (HSCL-11, chronicity, prior psychotherapy, treatment expectation, and GAF) are not reported, even though the variables were part of the models, because they were identical to the values of model 1. Similarly, the pairwise class comparisons reported in previous models (models 1 and 2, respectively) are not reported because the results were identical.

room for further improvement afterwards. Early responders are not characterized by further change, but by their ability to maintain their early improvements over the course of therapy, which is in line with previous findings.[36]

**Table 5.** Total and class means of movement-based attunement, number of sessions, early response, and outcome

| Class | Attunement | Number of sessions | Early response | Outcome |
|---|---|---|---|---|
| 1 | 1.131 | 36.78 | −0.078 | 0.090 |
| 2 | 1.674 | 40.89 | 0.290 | 0.294 |
| 3 | 1.443 | 36.39 | 0.103 | 0.191 |
| 4 | 1.541 | 38.94 | −0.035 | 0.097 |
| total | 1.489 | 37.55 | 0.061 | 0.161 |

Early response refers to percentage change on the HSCL-11 from session 1 to 10, outcome refers to the percentage change on the HSCL-11 from session 1 to the last available observation; Class 1: high impairment, no improvement; Class 2: high impairment, early response; Class 3: moderate impairment; Class 4: low impairment.

## *Limitations*

Some limitations should be considered when interpreting the results of this study. First, as the analyses were based on naturalistic data from routine mental health care, other variables that were not assessed could have influenced the effect of movement-based attunement and might be alternative explanations for the association with early change classes. However, all analyses were adjusted for possibly confounding variables to improve internal validity. Variables that were repeatedly shown to be associated with HSCL-11 improvement (initial HSCL-11, chronicity, prior psychotherapy, treatment expectation, and GAF) were used as covariates. Furthermore, multilevel models accounted for the nested structure of the data.

Second, independent of adjusting for covariates, some effects were on the borderline of significance. Despite showing a rather large difference in early percentage changes, early responders (Class 2; 28.96% early change) and moderately impaired patients (Class 3; 10.30%) as well as patients with low impairment (Class 4; −3.54%) did not differ significantly according to the multilevel linear models. This may have been due to insufficient statistical power in a sample of $N = 161$ cases. Furthermore, there may have been significant variability in the classes (more variability within than between classes), calling into question the homogeneity of the classes. However, the classes found were comparable to those found in previous studies[34] and the clinically most important ones (highly impaired patients with and without early change) showed meaningful and statistically significant differences in early change and in treatment outcome. In general, using a larger sample size could have helped to also find small effects, however, this could have also led to small and clinically irrelevant effects reaching significance.

Furthermore, in this data set movement-based attunement was only important for about 25% of patients (those with a high initial impairment) when differentiating between early improvement and no change. Thus, the findings on the influence of attunement can only be generalized and applied to highly impaired cases. However, it is critical to identify such factors, especially for this small subgroup, as these are the patients who also show poorer treatment response and outcome. Due to this subsample's smaller sample size ($n = 43$), extreme values may have had a stronger influence and the results might be less reliable overall. Because the significance tests were based on multinomial models, each comparing two groups pairwise, and the effect of interest was statistically significant, the smaller subgroup is not expected to have a meaningful impact on statistical power.

Third, movement-based attunement was computed for one session during the first 10 sessions of therapy. Analyzing a larger number of sessions could improve the measurement's reliability. Likewise, longitudinal developments in attunement could be observed and their relationship to early change investigated. Previous work on latent change classes suggests that change patterns vary systematically with the phase of therapy,[69] so our findings cannot be generalized to later phases. In addition, movement synchrony has been shown to change over the course of treatment.[14,23] Therefore, the available data only reflect the early phase of treatment.

Furthermore, the distribution of the movement-based attunement measure differed descriptively between latent change classes. For all classes but Class 1, distributions were slightly right-skewed with few larger values. Except for Class 2, all distributions were approximately normally distributed. In order not to degrade the scale level by a non-linear transformation of the variables, no transformations were performed. However, such pre-processing could potentially affect the results. The distributions fit the findings, with the early response class (Class 2) showing larger attunement scores compared to the other classes. The class with severely impaired patients without early change (Class 1) showed the fewest patients with higher attunement scores.

Moreover, movement-based attunement has been studied as a conglomerate in which variances of patient and therapist are inextricably mixed. Recent research on movement synchrony has decomposed effects of movement synchrony into effects between- and within-dyads.[70,71] Similarly, a decomposition of effects into their between- and within-therapist components would be conceivable to separate the therapist's influence from that of the patient.[72] Other efforts to better understand synchrony and its effects have focused on idiographic analyses.[73] However, in the present study, data were analyzed at the group level, as movement time-series were available for only one session per dyad and the study aimed at grouping patients into early change classes and examining correlates of class membership instead of analyzing idiographic within-patient associations.

## Conclusions

Movement-based patient-therapist attunement was shown to be a correlate of early response among patients with high initial impairment, with stronger attunement predicting early response, which was in turn associated with a better treatment outcome. The automated and resource-conscious assessment of movement-based attunement can be achieved using MEA and cross-lagged correlation analysis. In the future, early change classes as well as a measure of patient-therapist attunement might be easily incorporated into comprehensive feedback and decision-support systems to automatically identify patients at risk for treatment failure.[47,51] The only crucial requirement is that the session is video recorded at a high-quality resolution with one camera per interaction partner.

One strength of movement-based attunement as a measure in psychotherapy is that it can be assessed early in treatment (during the first sessions) and serve as an indicator for prognosis. Beyond prognosis, attunement might inform differential indication (e.g. patients in dyads with a low attunement could benefit from a more motivation- and alliance-focused treatment strategy) and adaptive indication (e.g. changing strategies based on current attunement). However, to derive any indication from initial movement-based attunement or its changes over the course of treatment, further research is needed. In addition to replicating the findings in larger data, the early response prediction model needs to be validated in independent holdout data. Furthermore, the prediction should be applied to new cases prospectively to examine the effects of providing therapists with this information on treatment progress and outcome instead of only examining the association in retrospective designs. To foster measurement-based and data-informed psychological care, therapists could be provided with information on their attunement with their patients as well as patients' pre-treatment impairment to identify patients at risk for treatment failure (high initial impairment and low movement-based attunement) or with a high probability of early response (high initial impairment and strong movement-based attunement).[74]

Alongside the evaluation of predictions, it should be investigated whether patient–therapist attunement can be actively influenced by the therapist and, if so, learned in therapist training. Clinical support tools could assist therapists with low movement-based attunement by implementing deliberate practice[75] to build behaviors associated with attunement that increase the likelihood of early change in patients with high initial impairment. Objective, video-based assessment and calculation of movement-based attunement might provide new information for therapists that is not available from self-report questionnaires or therapists' clinical impression and support therapists in their clinical decision-making.[76,77]

**Acknowledgments:** We would like to thank Kaitlyn Poster for proofreading the manuscript.

**Contributorship:** All authors made substantial contributions to the conception and design of the study. BS was responsible for data preparation, management, and analyses, and wrote the first draft of the manuscript. JR and WL consulted in data analyses and contributed to the writing of the manuscript. AD critically evaluated the manuscript and added valuable comments. All authors reviewed and edited the manuscript and approved the final version of the manuscript.

**Declaration of Conflicting Interests:** The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

**Funding:** The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article. This work was supported by the Open Access Fund of Universität Trier and German Research Foundation Deutsche Forschungsgemeinschaft (grant numbers LU 660/10-1, LU 660/8-1).

**Ethical approval:** The ethics committee of the University of Trier approved this study

**Guarantor:** BS.

**ORCID iD:** Brian Schwartz https://orcid.org/0000-0003-4695-4953

## References

1. Repp BH and Su Y-H. Sensorimotor synchronization: a review of recent research (2006–2012). *Psychon Bull Rev* 2013; 20: 403–452.
2. Coutinho J, Pereira A, Oliveira-Silva P, et al. When our hearts beat together: cardiac synchrony as an entry point to understand dyadic co-regulation in couples. *Psychophysiology* 2021; 58: e13739.
3. Prinz J, Rafaeli E, Reuter JK, et al. Physiological activation and co-activation in an imagery-based treatment for test anxiety. *Psychother Res* 2022; 32: 238–248.
4. Tschacher W and Meier D. Physiological synchrony in psychotherapy sessions. *Psychother Res* 2020; 30: 558–573.
5. Zilcha-Mano S, Goldstein P, Dolev-Amit T, et al. Oxytocin synchrony between patients and therapists as a mechanism underlying effective psychotherapy for depression. *J Consult Clin Psychol* 2021; 89: 49–57.
6. Anderson JR, Myowa-Yamakoshi M and Matsuzawa T. Contagious yawning in chimpanzees. *Proc Biol Sci* 2004; 271: S468–S470.
7. Panksepp J. *Affective neuroscience: the foundations of human and animal emotions*. Oxford: Oxford University Press, 2004.
8. Simpson EA, Murray L, Paukner A, et al. The mirror neuron system as revealed through neonatal imitation: presence from birth, predictive power and evidence of plasticity. *Philos Trans R Soc Lond B Biol Sci* 2014; 369: 20130289.
9. Prochazkova E and Kret ME. Connecting minds and sharing emotions through mimicry: a neurocognitive model of emotional contagion. *Neurosci Biobehav Rev* 2017; 80: 99–114.
10. Asher M, Kauffmann A and Aderka IM. Out of sync: nonverbal synchrony in social anxiety disorder. *Clin Psychol Sci* 2020; 8: 280–294.
11. Altmann U, et al. Investigation of movement synchrony using windowed cross-lagged regression. In: Esposito A, Vinciarelli A and Vicsi K (eds) *Analysis of verbal and nonverbal communication and enactment. the processing issues*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2011, pp.335–345.
12. Tschacher W, Rees GM and Ramseyer F. Nonverbal synchrony and affect in dyadic interactions. *Front Psychol* 2014; 5: 1323.
13. Mogan R, Fischer R and Bulbulia JA. To be in synchrony or not? A meta-analysis of synchrony's effects on behavior, perception, cognition and affect. *J Exp Soc Psychol* 2017; 72: 13–20.
14. Altmann U, Schoenherr D, Paulick J, et al. Associations between movement synchrony and outcome in patients with social anxiety disorder: evidence for treatment specific effects. *Psychother Res* 2020; 30: 574–590.
15. Ramseyer F and Tschacher W. Nonverbal synchrony in psychotherapy: coordinated body movement reflects relationship quality and outcome. *J Consult Clin Psychol* 2011; 79: 284–295.
16. Schoenherr D, Strauss B, Paulick J, et al. Movement synchrony and attachment related anxiety and avoidance in social anxiety disorder. *J Psychother Integr* 2021; 31: 163–179.
17. Paulick J, Deisenhofer A-K, Ramseyer F, et al. Nonverbal synchrony: a new approach to better understand psychotherapeutic processes and drop-out. *J Psychother Integr* 2018; 28: 367–384.
18. Schoenherr D, Paulick J, Strauss BM, et al. Nonverbal synchrony predicts premature termination of psychotherapy for social anxiety disorder. *Psychotherapy (Chic)* 2019; 56: 503–513.
19. Galbusera L, Finn MT and Fuchs T. Interactional synchrony and negative symptoms: an outcome study of body-oriented psychotherapy for schizophrenia. *Psychother Res* 2018; 28: 457–469.
20. Kupper Z, Ramseyer F, Hoffmann H, et al. Nonverbal synchrony in social interactions of patients with schizophrenia indicates socio-communicative deficits. *PLoS One* 2015; 10: e0145882.
21. Lutz W, Prinz JN, Schwartz B, et al. Patterns of early change in interpersonal problems and their relationship to nonverbal synchrony and multidimensional outcome. *J Couns Psychol* 2020; 67: 449–461.
22. Fujiwara K and Yokomitsu K. Video-based tracking approach for nonverbal synchrony: a comparison of motion energy analysis and openpose. *Behav Res Methods* 2021; 53: 2700–2711.
23. Paulick J, Rubel JA, Deisenhofer A-K, et al. Diagnostic features of nonverbal synchrony in psychotherapy: comparing depression and anxiety. *Cognit Ther Res* 2018; 42: 539–551.

24. Altmann U. *Synchronisation nonverbalen Verhaltens*. Wiesbaden: Springer Fachmedien Wiesbaden, 2013.
25. Tschacher W, Ramseyer F and Koole SL. Sharing the now in the social present: duration of nonverbal synchrony is linked with personality. *J Pers* 2018; 86: 129–138.
26. Tschacher W, Ramseyer F and Pfammatter M. The social present in psychotherapy: duration of nowness in therapeutic interaction. In: Ochs M, Borcsa M and Schweitzer J (eds) *Systemic research in individual, couple, and family therapy and counseling*. Cham: Springer International Publishing, 2020, pp.39–53.
27. Muthén B. The potential of growth mixture modelling. *Inf Child Develop* 2006; 15: 623–625.
28. Nagin DS and Odgers CL. Group-based trajectory modeling in clinical research. *Annu Rev Clin Psychol* 2010; 6: 109–138.
29. Smith KV and Ehlers A. Cognitive predictors of grief trajectories in the first months of loss: a latent growth mixture model. *J Consult Clin Psychol* 2020; 88: 93–105.
30. Skriner LC, Chu BC, Kaplan M, et al. Trajectories and predictors of response in youth anxiety CBT: integrative data analysis. *J Consult Clin Psychol* 2019; 87: 198–211.
31. Hilbert A, Herpertz S, Zipfel S, et al. Early change trajectories in cognitive-behavioral therapy for binge-eating disorder. *Behav Ther* 2019; 50: 115–125.
32. Lutz W, Hofmann SG, Rubel J, et al. Patterns of early change and their relationship to outcome and early treatment termination in patients with panic disorder. *J Consult Clin Psychol* 2014; 82: 287–297.
33. Rubel J, Lutz W, Kopta SM, et al. Defining early positive response to psychotherapy: an empirical comparison between clinically significant change criteria and growth mixture modeling. *Psychol Assess* 2015; 27: 478–488.
34. Stulz N, Lutz W, Leach C, et al. Shapes of early change in psychotherapy under routine outpatient conditions. *J Consult Clin Psychol* 2007; 75: 864–874.
35. Lutz W, Stulz N and Köck K. Patterns of early change and their relationship to outcome and follow-up among patients with major depressive disorders. *J Affect Disord* 2009; 118: 60–68.
36. Haas E, Hill RD, Lambert MJ, et al. Do early responders to psychotherapy maintain treatment gains? *J Clin Psychol* 2002; 58: 1157–1172.
37. Moggia D, Lutz W, Arndt A, et al. Patterns of change and their relationship to outcome and follow-up in group and individual psychotherapy for depression. *J Consult Clin Psychol* 2020; 88: 757–773.
38. Beard JIL and Delgadillo J. Early response to psychological therapy as a predictor of depression and anxiety treatment outcomes: a systematic review and meta-analysis. *Depress Anxiety* 2019; 36: 866–878.
39. First MB, Spitzer RL, Gibbon M, et al. Structured clinical interview for axis I DSM-IV disorders-patient edition (SCID-I/P). New York: Biometrics Research Department. New York State Psychiatric Institute 1995.
40. Bronisch T, Hiller W, Mombour W, et al. *International diagnostic checklists for personality disorders according to ICD_10 and DSM–IV-IDCL-P*. Seattle, WA: Hogrefe and Huber Publishing, 1996.
41. Ramseyer FT. Motion energy analysis (MEA): a primer on the assessment of motion from video. *J Couns Psychol* 2020; 67: 536–549.
42. Schoenherr D, Paulick J, Strauss BM, et al. Identification of movement synchrony: validation of windowed cross-lagged correlation and -regression with peak-picking algorithm. *PLoS One* 2019; 14: e0211494.
43. Lutz W, Tholen S, Schürch E, et al. Reliabilität von Kurzformen gängiger psychometrischer instrumente zur evaluation des therapeutischen fortschritts in psychotherapie und psychiatrie. *Diagnostica* 2006; 52: 11–25.
44. Derogatis LR. *Brief symptom inventory*. Baltimore, MD: Clinical Psychometric Research, 1975.
45. Coyne JC, Kessler RC, Tal M, et al. Living with a depressed person. *J Consult Clin Psychol* 1987; 55: 347–352.
46. Lutz W, Leon SC, Martinovich Z, et al. Therapist effects in outpatient psychotherapy: a three-level growth curve approach. *J Couns Psychol* 2007; 54: 32–39.
47. Lutz W, Rubel JA, Schwartz B, et al. Towards integrating personalized feedback research into clinical practice: development of the trier treatment navigator (TTN). *Behav Res Ther* 2019; 120: 103438.
48. Husen K, Rafaeli E, Rubel JA, et al. Daily affect dynamics predict early response in CBT: feasibility and predictive validity of EMA for outpatient psychotherapy. *J Affect Disord* 2016; 206: 305–314.
49. Howard KI, Moras K, Brill PL, et al. Evaluation of psychotherapy: efficacy, effectiveness, and patient progress. *American Psychologist* 1996; 51: 1059–1064.
50. Lueger RJ, Howard KI, Martinovich Z, et al. Assessing treatment progress of individual patients using expected treatment response models. *J Consult Clin Psychol* 2001; 69: 150–158.
51. Lutz W, Deisenhofer A-K, Rubel J, et al. Prospective evaluation of a clinical decision support system in psychological therapy. *J Consult Clin Psychol* 2022; 90: 90–106.
52. Aas IHM. Guidelines for rating global assessment of functioning (GAF). *Ann Gen Psychiatry* 2011; 10: 2.
53. Howard KI, Kopta SM, Krause MS, et al. The dose–effect relationship in psychotherapy. *American Psychologist* 1986; 41: 159–164.
54. Muthén LK and Muthén BO. *Mplus user's guide*. Eighth Edition. Los Angeles, CA: Muthén & Muthén, 2017.
55. R Core Team. *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing, 2020.
56. Stekhoven DJ. missForest: Nonparametric Missing Value Imputation using Random Forest: R package version 1.4, 2013.
57. Stekhoven DJ and Bühlmann P. MissForest--non-parametric missing value imputation for mixed-type data. *Bioinformatics* 2012; 28: 112–118.
58. Breslow NE and Clayton DG. Approximate inference in generalized linear mixed models. *J Am Stat Assoc* 1993; 88: 9–25.
59. Elff M. mclogit: Multinomial Logit Models, with or without Random Effects or Overdispersion: R package version 0.8.6.4, 2020.
60. Handcock MS, Chaudhuri S and Rendall MS. glmc: An R package for generalized linear models subject to constraints: R package version 0.3-1, 2018.
61. Wickham H. *ggplot2: elegant graphics for data analysis*. New York, NY: Springer, 2016.
62. Schwartz B, Cohen ZD, Rubel JA, et al. Personalized treatment selection in routine care: integrating machine learning

and statistical algorithms to recommend cognitive behavioral or psychodynamic therapy. *Psychother Res* 2021; 31: 33–51.

63. Bates D, Mächler M, Bolker B, et al. Fitting linear mixed-effects models using lme4. *J Stat Soft* 2015; 67: 1–48.
64. Kuznetsova A, Brockhoff PB and Christensen RHB. lmerTest package: tests in linear mixed effects models. *J Stat Soft* 2017; 82: 1–26.
65. Tschacher W, Ramseyer F and Bergomi C. The subjective present and its modulation in clinical contexts. *Timing Time Percept* 2013; 1: 239–259.
66. Koffmann A. Has growth mixture modeling improved our understanding of how early change predicts psychotherapy outcome? *Psychother Res* 2018; 28: 829–841.
67. Schulte D and Eifert GH. What to do when manuals fail? The dual model of psychotherapy. *Clinical Psychology: Science and Practice* 2002; 9: 312–328.
68. Saxon D and Barkham M. Patterns of therapist variability: therapist effects and the contribution of patient severity and risk. *J Consult Clin Psychol* 2012; 80: 535–546.
69. Rubel J, Lutz W and Schulte D. Patterns of change in different phases of outpatient psychotherapy: a stage-sequential pattern analysis of change in session reports. *Clin Psychol Psychother* 2015; 22: 1–14.
70. Cohen K, Ramseyer FT, Tal S, et al. Nonverbal synchrony and the alliance in psychotherapy for major depression: disentangling state-like and trait-like effects. *Clin Psychol Sci* 2021; 9: 216770262098529.
71. Prinz J, Boyle K, Ramseyer F, et al. Within and between associations of nonverbal synchrony in relation to Grawe's general mechanisms of change. *Clin Psychol Psychother* 2021; 28: 159–168.
72. Altmann U, Schwartz B, Schönherr D, et al. Therapeuteneffekte bei der ambulanten behandlung sozialer Ängste. *Psychotherapeut* 2020; 65: 444–455.
73. Ramseyer FT. Exploring the evolution of nonverbal synchrony in psychotherapy: the idiographic perspective provides a different picture. *Psychother Res* 2020; 30: 622–634.
74. Lutz W, Schwartz B and Delgadillo J. Measurement-based and data-informed psychological therapy. *Annu Rev Clin Psychol* 2022; 18: 71–98.
75. Chow DL, Miller SD, Seidel JA, et al. The role of deliberate practice in the development of highly effective psychotherapists. *Psychotherapy (Chic)* 2015; 52: 337–345.
76. Delgadillo J and Lutz W. A development pathway towards precision mental health care. *JAMA Psychiatry* 2020; 77: 889–890.
77. Lutz W. Data-informed advances and technology augmentation. *Cogn Behav Pract* 2022; 83: 749.