

in patients who were undergoing psychodynamic therapy. They found at pretreatment that none of the 35 borderline patients, who were assessed with the Adult Attachment Interview (AAI), belonged to the secure attachment category. However, at post-treatment 40% of the patients were categorized as securely attached.

Travis, Binder, Bliwise and Horne-Moyer (2001) used the Bartholomew Attachment Rating on 84 participants in outpatient psychotherapy with different diagnoses. There was a subsample of 29 patients for whom posttreatment assessments were available after time-limited psychodynamic therapy (with a mean of 21 sessions). Within this group they found dimensional pre-post increases in attachment security and decreases in several insecure dimensions. At the beginning of the treatment none of the patients was categorized as securely attached, and by post-treatment seven fell into this category.

Lawson, Barnes, Madkins, and Francois-Lamonte (2006) collected pre post self-ratings of attachment using the Adult Attachment Scale (AAS) from 33 males being treated for domestic violence. The treatment was an integrated cognitive-behavioral/psychodynamic outpatient group therapy which lasted 17 weeks. On a categorical level, the authors found increases in the number of securely attached participants (14 versus 25); however, the means of the attachment dimensions did not show any significant changes.

All results mentioned so far must be interpreted cautiously because of the small sample sizes and the lack of comparison groups.

Levy et al. (2006) were the first who investigated the changes in attachment as a result of psychotherapy using a comparison group design. They collected Adult Attachment Interviews from outpatients with Borderline Personality Disorder before and after treatment. Initially 90 patients were randomly assigned to one of three treatment conditions each of which lasted 1 year. They found (based on data available from 60 patients at post-treatment) that there was a statistically significant increase in the number of securely attached participants only in the transference-focused psychotherapy group (1 pre-versus 7 posttreatment), but not in the dialectical behavioral or psychodynamic supportive therapy groups.

Tasca, Balfour, Ritchie, and Bissada (2007) analyzed pre-post attachment changes among inpatients with a binge eating disorder. They randomly assigned 95 patients to one of the treatment groups: cognitive-behavioral or psychodynamic-interpersonal therapy. Their results, which were based on pre-post attachment self-report data (Attachment Style Questionnaire) from 66 patients, indicated a

decrease in attachment insecurity in both treatment groups.

Muller and Rosenkranz (2009) used a waitlist control group. They collected attachment self-reports (Relationship Scales Questionnaire, RSQ) from inpatients with Posttraumatic Stress Disorder undergoing the Program for Traumatic Stress Recovery. Data were available for $n = 101$ (pre post) and $n = 61$ patients (6-month follow-up). Individuals in the waitlist control group ($n = 46$) also filled out the RSQ at pre- and posttreatment. The authors' findings were inconsistent. Contrasted to the controls, the patients showed pre post improvements on only two of the six attachment scales. Within-group statistics indicated that the attachment improvements had disappeared by follow-up; however, there weren't any data from the control group available for follow-up. The results were also difficult to interpret because of the lack of randomization and because covariates were not taken into account, despite the existence of pretreatment differences between the patients and the controls.

Altogether, the existing studies suggest that attachment characteristics may be improved through different psychotherapeutic interventions in different diagnostic groups independent of the attachment measure used. However, attachment changes were predominantly investigated using diagnostically specific samples that may not be very common in psychotherapeutic routine care. Moreover, long-term attachment improvements have not yet been adequately examined.

The goal of the present study was to investigate whether inpatients undergoing group psychotherapy in routine care show changes in attachment characteristics. More specifically, we expected that the patients' attachment (increase in security, decrease in anxiety and avoidance) would improve in comparison to the untreated controls with regard to both pre-post changes and changes from pretreatment to a 1-year follow-up.

Method

Participants

We designed this study as a naturalistic observational study including a non-randomized control group. A total of $N = 265$ psychotherapy inpatients were recruited from five psychotherapeutic institutions in Germany. Two of the psychotherapy wards were from the University Hospital in Jena; one ward was psychodynamically oriented ($n = 38$) and the other ward uses CBT ($n = 35$). The other three psychodynamically oriented hospitals were from different cities: Rhein-Clinic, Bad Honnef ($n = 45$);

Asklepios Hospital, Stadtroda ($n=97$); Department of Psychotherapy and Psychosomatic Medicine, University Hospital Dresden ($n=50$). The control group subjects ($N=260$) were patients from a general practice in Weimar ($n=112$) as well as psychology undergraduate students ($n=47$) and medical students ($n=101$) from the Friedrich Schiller University in Jena. Initially we tried to use a control group that was exclusively made up of physically ill patients but due to the limited attendance we decided to also include students in the control group.

This study was part of a research project which examined changes in attachment and depression. All participants received an invitation to participate in "a study of the changes of depressive complaints" and all provided informed consent as well as their mailing address. This study was approved by the local ethics committee.

In the cooperating hospitals, all admitted patients were asked to participate in the study. The inpatient sample, therefore, consisted of patients with different primary diagnoses (comorbidity was not considered in this study), which were made on the basis of clinical judgment (unipolar affective disorders 30%, anxiety disorders 28%, eating disorders 21%, personality disorders 12%, somatoform disorders 5%, others 4%). The general practitioner in Weimar asked all patients older than 18 years to participate. As compensation for their participation, patients received feedback information about their depression scores and three patients received €100 as part of a random drawing. The same was offered to the medical students, whereas participating psychology students only received course credits for their participation.

Data were collected at three different occasions: pretreatment, posttreatment, and 1-year follow-up. The patients filled out questionnaires at the beginning of their treatment, at the end (mean treatment duration: $Md=9$ weeks; range: 3–18 weeks), and a year after finishing treatment (via mail). The questionnaires were sent to the control group individuals at identical intervals. At posttreatment, data were available from $n=188$ patients (drop-out due to treatment attrition) and $n=219$ controls. At follow-up, data were available from $n=152$ patients and $n=200$ controls.

Measures

For the assessment of adult attachment characteristics we used three different multi-item questionnaires.

Bielefeld Partnership Expectations Questionnaire (BFPE). The BFPE (Höger et al., 2008; Pollak, Wiegand-Grefe, & Höger, 2008) consists of 30 items (plus one warm-up item) that are rated on a 5-point Likert scale. The items of the BFPE evaluate the participants' expectations regarding an intimate partnership. The items are assigned to three subscales, *Fear of Rejection* (10 items, e.g., "I am afraid that my partner will withdraw from me if he/she knows the truth about my inner feelings"), *Readiness for Self-Disclosure* (12 items, e.g., "It is generally easy for me to talk to my partner about my innermost feelings"), and *Conscious Need for Care* (7 items, e.g., "When I feel neglected by my partner, I can become very depressed").

For the assignment to attachment categories, the participants' scale values can be compared to the values of the secure, ambivalent, and avoidant clusters of a large reference sample ($N=1506$; Höger et al., 2008). This therefore allows for an empirically based assignment to one of the classical attachment categories. Cronbach's alphas for the three scales ranged in our study from $\alpha=.77$ to $\alpha=.89$.

Evidence for convergent validity (associations with other attachment questionnaires) and theoretically expected associations with other concepts (such as coping or relationship quality) has been comprehensively presented by Pollak et al. (2008). The BFPE is especially suitable for a categorical assessment of attachment and we used it accordingly.

Grau's Attachment Questionnaire (GAQ). The GAQ was developed by Grau (1999). It also assesses attachment characteristics in intimate partnerships. The GAQ consists of 20 items which are rated on a 5-point Likert scale. The items are assigned to two subscales, *GAQ-Anxiety* (10 items, e.g., "I often worry that my partner could not like me enough") and *GAQ-Avoidance* (10 items, e.g., "If my partner gets too close to me, I withdraw"). The GAQ corresponds to the results of studies using exploratory factor analysis, which were based on a larger number of attachment questionnaires. In these studies orthogonal two-dimensional solutions were consistently found indicating the dimensions anxiety and avoidance (Brennan, Clark, & Shaver, 1998; Fraley, Waller, & Brennan, 2000; Stein et al., 2002). Convergent validity of the GAQ was confirmed by high associations with other attachment questionnaires, for example the BFPE (Grau, Clashausen, & Höger, 2003) and the German version of the Experiences in Close Relationships Scale (ECR; Neumann, Rohmann, & Bierhoff, 2007). Furthermore, hypothesized associations of the GAQ scales with partner relationship quality/satisfaction and

feelings of equity were validated (Grau & Doll, 2003). The reliability estimates were high in our study with internal consistencies of $\alpha = .92$ (*GAQ-Anxiety*) and $\alpha = .88$ (*GAQ-Avoidance*). By computing (*GAQ-Anxiety* + *GAQ-Avoidance*) (-1) , we provided a continuous measure of *GAQ-Security*.

Relationship Scales Questionnaire (RSQ).

The RSQ (Griffin & Bartholomew, 1994; German version from Steffanowski et al., 2001) consists of 30 items (5-point Likert scale) which assess attachment-related feelings, expectations, and motivations, not specifically for romantic partners, but for more general relationships. Originally, the items were assigned to four highly intercorrelated subscales with a fairly low reliability (Griffin & Bartholomew, 1994). However, Kurdek (2002) found the best model was a two-dimensional solution (anxiety and avoidance) which was based on a subset of only 13 of the 30 items. For our present study we administered the German version of the RSQ and used a two-dimensional factor solution with the dimensions *Anxiety* and *Avoidance* resulting from a preceding German study (Kirchmann, Fenner, & Strauss, 2007). This factor solution, which we used in the current study, consists of six items that were assigned to the dimension *RSQ-Anxiety* (e.g., "I worry that others don't value me as much as I value them") and 12 items that were assigned to *RSQ-Avoidance* (e.g., "I am somewhat uncomfortable being close to others"). The correlation of *RSQ-Anxiety* with *ECR-Anxiety* was $r = .62$, the correlation of the respective avoidance scales was $r = .71$. In our study, Cronbach's alpha for anxiety was $\alpha = .86$ and for avoidance $\alpha = .82$. We provided a continuous measure *RSQ-Security* by computing (*RSQ-Anxiety* + *RSQ-Avoidance*) (-1) .

Additionally, at the beginning of treatment, participants were asked to complete the Center for Epidemiological Studies Depression Scale (CES-D; Radloff, 1977; German version: Hautzinger & Bailer, 1993) which is a commonly used 20 item Likert-type questionnaire. The CES-D is a depression self-report with high internal consistency ($\alpha = .91$ in our study) and high correlations with other depression measures such as the Beck Depression Inventory and the Hamilton Depression Scale (Hautzinger & Bailer, 2002).

Furthermore, participants answered socio-demographic questions (age, gender, education, marital status, employment status, partner relationship).

Statistical procedures

Only in the randomized controlled trial can one expect that in a treatment and a control group

pretreatment characteristics show similar distribution statistics (frequencies, means, variances, etc.) and, therefore, that group differences in the outcome reflect unbiased treatment effects. In quasi-experimental studies confounding variables (variables related to both treatment assignment and outcome) may bias the estimates of treatment effects. Therefore it is crucial to measure and statistically control such confounders (e.g., Rosenbaum, 2002; Shadish, Cook, & Campbell, 2002; note that we call a measured confounder a *covariate*). In our present study, we used propensity score techniques for covariate modeling in combination with a generalized analysis of covariance (g-ANCOVA) provided by the statistics program *EffectLite 3.1.2*. For bivariate statistics and the computation of propensity scores we used *PASW 18.0*.

Propensity score modeling. Propensity scores (*PS*) are the conditional probabilities of a participant being in a treatment condition given a vector of covariates (for an overview see Guo & Fraser, 2010). Usually, these conditional probabilities are estimated by logistic regression or discriminant analysis. Numerous covariates can be taken into account simultaneously, yielding one single propensity score for each participant, provided that there are only two treatment conditions. A group of treated participants with the same *PS* as a group of controls also show similar distribution statistics of all the included covariates used for computing the propensity scores (Rosenbaum & Rubin, 1983). Hence, by accurately constructed *PS* a multitude of covariates (potentially all confounders) can be balanced simultaneously using only the *PS* in a statistical model. If there are differences in the outcome between treated participants and untreated controls with the same *PS*, these outcome differences can be attributed to the treatment and not to the covariates and, therefore, reflect the treatment effect.

A crucial and controversial issue in *PS* analysis is which pretreatment characteristics should be selected for *PS* computation. In the empirical literature, there are very different inclusion strategies (e.g., see the review by Weitzen, Lapane, Toledano, Hume, & Mor, 2004). There are two criteria for a proper *PS* model: the estimates should be unbiased and they should be efficient. For an unbiased estimation the inclusion of all potential confounders would be advantageous. On the other hand, the inclusion of unimportant covariates is detrimental to efficiency and, therefore, parsimonious models with only a few covariates would be preferable. In order to increase efficiency, Brookhart et al. (2006) recommended that the covariate selection for the *PS* computation should be based on the associations