This document is the pre-review version of the following article:

Konkolÿ Thege, B., Petroll, C., Rivas, C., & Scholtens, S. (in press).

The effectiveness of family constellation therapy in improving mental health: A systematic review.

Family Process, https://doi.org/10.1111/famp.12636

The revised, copy-edited and final version of this manuscript has been published in *Family Process* at:

https://onlinelibrary.wiley.com/doi/10.1111/famp.12636

The present article version does not exactly replicate the authoritative document published in *Family Process*.

Abstract

Family/systemic constellation therapy is a short-term group intervention aiming to help clients better understand and then change their conflictive experiences within a social system (e.g., family). The aim of the present systematic review was to synthetize the empirical evidence on the tolerability and effectiveness of this intervention in improving mental health. The PsycINFO, Embase, MEDLINE, ISI Web of Science, Psyndex, PsycEXTRA, ProQuest Dissertations & Theses, The Cochrane Library, Google Scholar, and an intervention-specific organization's data bases were searched for quantitative, prospective studies published in English, German, Spanish, French, Dutch, or Hungarian up until April 2020. Out of 4,197 identified records, 67 were assessed for eligibility, with 12 studies fulfilling inclusion criteria (10 independent samples; altogether 568 participants). Outcome variables were diverse ranging from positive self-image through psychopathology to perceived quality of family relationships. Out of the 12 studies, 9 showed statistically significant improvement post-intervention. The studies showing no significant treatment benefit were of lower methodological quality. The random-effect meta-analysis – conducted on 5 studies in relation to general psychopathology – indicated a moderate effect (Hedges' g of 0.531, CI: 0.387–0.676). Authors of 7 studies also investigated potential iatrogenic effects and 4 studies reported minor or moderate negative effects in a small proportion (5-8%) of participants that potentially could have been linked to the intervention. The data accumulated to date point into the direction that family constellation therapy is an effective intervention with significant mental health benefits in the general population; however, the quantity and overall quality of the evidence is low.

Keywords: family constellation, mental health, psychopathology, effectiveness, systematic review

Introduction

Family / systemic constellation therapy is a short-term group intervention aiming to help clients gain insights into and then change their inner image of a conflictual system and finally change their behavior in relation to that same system (Hunger, Bornhäuser, Link, Schweitzer, & Weinhold, 2014). The system addressed is most often the family but alternatively other systems (e.g., workplace community, ego parts, victim-perpetrator dyads) can also be the target of the intervention (in consideration of the tradition in clinical practice, the term 'family constellation therapy' is used throughout the manuscript in this broader sense, also referring to therapeutic work with systems other than the family). This form of therapy was developed in Germany in the early '90s integrating elements of – among others – psychodrama, family sculptures, contextual therapy, and certain South-African aboriginal traditions (Butollo, Franke, & Hellinger, 2017; D. B. Cohen, 2006; McQuillin & Welford, 2013; Stiefel, Harris, & Zollmann, 2002; Stones, 2006; Weber, 1993).

The intervention is typically administered in a group setting in which approximately 15-25 unrelated participants (i.e., participants are not members of the same system) meet for a one-time, 2-3-day, facilitator-led seminar/workshop. Each constellation starts with a brief interview between the facilitator and active client to clarify the individual's goal with the intervention. This is followed by a joint decision about which members of the client's system play an important role in the issue presented and these are represented by other group members during the constellation (Orban, 2008). The representatives (including the client's representative) are positioned in the room by the client initially, with spatial distances, angles, and body postures meant to correspond to the client's inner image of the system ("problem constellation"). This allows the facilitator to identify the dynamics beneath the client's presenting concern, while at the same time helps the client reflect on their internal experience from a more objective, partially external point of view (as they are observers and not participants at this point). This part of the process is non-verbal, focusing on what participants begin to experience as being part of the structure created by the active client. Next, the representatives are asked by the therapist about their physical sensations, feelings, and thoughts they had while in their positions. Rearrangements, spatial adjustments, and brief, ritualized conversations are made based on the principles of healthy functioning within a system (Hellinger, 1994; Weber, 1993) until a constellation is identified that offers a resolution for

the active participant's issue. Ideally this "solution constellation" provides a new framework for the client to feel, think, and behave in the given system (Hunger, Weinhold, Bornhäuser, Link, & Schweitzer, 2015).

Family constellation therapy has become particularly popular in Europe and South America (even becoming a part of the public health care system in certain countries; Franco de Sá, Nogueira, & De Almeida Guerra, 2019; Krüger & Schmidt-Michel, 2003; Mahr & Brömer, 2008) and is rapidly expanding in North America and Asia as well (Choi & Oh, 2018; North American Systemic Constellations, 2019a, 2019b; Pritzker & Duncan, 2019). Thousands of practitioners around the world use this method (D. B. Cohen, 2006) and with the German professional association 'Deutsche Gesellschaft für Systemaufstellungen' alone, more than 450 professionals are registered currently. Compared to its widespread use by therapists of various theoretical and professional backgrounds, little effort has been made to generate and critically evaluate empirical data regarding the effectiveness of the intervention.

Family constellation therapy has been adjusted and delivered to a large variety of client groups ranging from the general population (Broughton, 2006) through prisoners (D. B. Cohen, 2009) to different patient groups (e.g., Hausner, 2015; Jafferany et al., 2019; Nazarkiewicz & Bourquin, 2017; Ramos & Ramos, 2019). However, the number of studies using empirical methods to formally investigate the effectiveness or mechanisms of action of family constellation therapy is small and dominated by qualitative (Chu, 2008; Franke, 1996; Georgiadou, 2012; Häuser, Klein, & Schmidt-Keller, 1998; Junge, 1998) or mixed methods (Laireiter & Mitterhuemer, 2011; Mahr & Brömer, 2008; Rieger & Stückemann, 1999) studies investigating clients' satisfaction with the intervention. Despite the often rapid and significant positive changes family constellation therapy can produce in participants (Langlotz, 2005), there has been some concern among health care professionals regarding the safety of this therapeutic approach (e.g., no professional follow-up after the one-time workshop, which might be emotionally upsetting for some participants; Nelles, 2005; Reuter, 2005; Schneider, 2010; Talarczyk, 2011).

For the above reasons, synthesizing and critically evaluating the available empirical data regarding the effectiveness and tolerability of family constellation therapy is of high public health importance not just in North America but also globally. To date, two systematic reviews have been conducted on this intervention. Neither of them focused specifically on quantitative data regarding mental health outcomes and they did not emphasize data on tolerability/safety of the intervention

either. Further, the first review (Weinhold & Reinhard, 2014) summarized the research evidence up to early 2012, while several high-quality studies have been published since then. In addition, this review has been published as a book chapter written in German and is not available online – significantly limiting accessibility for a broader audience. Although the second review is more recent and was published in English, it was restricted to English language, peer-reviewed papers, which approach resulted in identifying merely 3 empirical studies to synthesize (Hurley, Koenning, & Bray, 2018). Therefore, the goal of this study was to systematically review the empirical evidence regarding family constellation therapy 1) focusing on quantitative data related to mental health outcomes, 2) considering all the evidence accumulated to date, 3) considering languages other than English as well, and 4) paying adequate attention to information related to tolerability/safety.

Methods

The protocol of the present systematic review was registered in the International Prospective Register of Systematic Reviews (PROSPERO) September 19, 2018 (# CRD42018109124). Given the preliminary stage of empirical research on family constellation therapy and the strong emphasis on locating all available evidence – including the gray literature – in the present study, the author team was not able to predict at the time of registration if enough data on the same mental health outcome would be reported. As a consequence, the study protocol included a narrative synthesis only; finally however, a meta-analysis on non-diagnosis-specific psychopathology – the single variable considered by a large enough number of studies – was also conducted.

Eligibility criteria

Studies included in the review met the following criteria: (1) quantitative studies with a longitudinal study design (including at least 2 assessment points, at least one of which was occurring before- and at least one of which was occurring after the intervention¹) (2) that evaluated the efficacy/effectiveness of family/systemic constellations on outcome measures of mental

-

¹ For instance, a study on goal attainment was excluded as it investigated the main variables of interest only at the 2-week and 4-month follow-up, while at baseline only qualitative data were collected about the participants' goals regarding the intervention (Bornhäuser & Wolff, 2014).

health. Given the limited amount of empirical data, any indicators of mental health (e.g., well-being, social functioning, psychopathology etc.) were considered as eligible outcome variables and no restrictions were made on participant populations either (e.g., general population, psychiatric in- or outpatients).

Exclusion criteria were: 1) studies without a precisely defined outcome, 2) qualitative and case studies, 3) no description of study methodology or assessment tool, 4) no available full text, and 5) study language other than English, German, Spanish, French, Dutch, or Hungarian. In the case of mixed method studies (combination of qualitative and quantitative approaches), the quantitative portion of the study was considered.

Search strategy and screening

To include both peer-reviewed and the gray literature as well, an extensive literature search was conducted including the following databases: PsycINFO, Embase, MEDLINE, ISI Web of Science, Psyndex, PsycEXTRA, ProQuest Dissertations & Theses, the Cochrane Library, and Google Scholar. Considering the date of introduction of family/systemic constellations into the clinical practice, the search was limited to studies published after the 1st of January 1993. The electronic data base searches were completed initially on 8th August 2018 and updated on the 6th of April 2020 and considered scientific works published in 6 languages (English, German, Spanish, French, Dutch, and Hungarian). The search terms included 'Family Constellation(s)', 'Systemic Constellation(s)', 'System Constellation(s)' and 'Structural Constellation(s)' as well as their grammatical variations and equivalents in the other five languages (the detailed list of search terms can be found is Supplementary Table 1). To reduce the number of irrelevant hits ('family constellation' is a common general term referring to the structure of a family), terms were searched in the title of the publications in the case of Google Scholar; while in the rest of the data bases, both the title and abstract was searched for the search terms.

In addition to traditional scientific databases, the database of the German Society of Systemic Constellations (Deutsche Gesellschaft für Systemaufstellungen; DGfS), the largest professional body devoted to the study and practice of the intervention, was also added to the pool of records to screen. Reference list of included studies and studies citing the included studies in Google Scholar were also screened for additional, potentially relevant records. The screening process – based on title and/or abstract – was completed by different members of the author team

depending on the language of the record (one assessor per item). Eligibility – based on (German or English language) full text – was assessed by the lead author, experienced in conducting systematic reviews and having content expertise specific to the research and clinical applications of the intervention.

Data extraction

Data extraction for all variables (including methodological quality) and for each eligible study was completed by two independent researchers (both with former experience in conducting systematic reviews) and discrepancies were resolved by reaching consensus. As part of the data extraction process, the following variables were considered: publication type (peer reviewed journal article, book / book chapter, thesis / dissertation, non-peer reviewed journal article, online report), study design, sample size, country of study, type of sample (e.g., normal population vs. psychiatric outpatients), sex composition (all male, all female or mixed; if mixed, percentage of female participants), and age of respondents.

The data extraction also specified detailed methodological characteristics including information regarding the control group (no-, convenience-, matched-, or randomized control group), length of intervention, length of follow-up, training level and professional background of intervention provider based on description in the article or personal website (e.g., psychiatrist with several decades of experience with family constellation therapy, social worker novice in family constellation therapy), intervention setting (private or public health care), outcome variables (construct and assessment technique), and main results. A second, simplified variable to describe overall results was also created with 2 response categories: statistically significant improvements reported or not. Finally, data were also extracted on whether study authors assessed iatrogenic effects (negative side effects not assessed, assessed and not found, assessed and found).

Study quality was assessed in a standardized way by the 2018 version of the Mixed Methods Appraisal Tool (MMAT) (Hong et al., 2018). The advantage of this rating tool is that it provides the opportunity to evaluate studies with different designs (i.e., qualitative-, quantitative randomized controlled-, quantitative non-randomized-, quantitative descriptive-, and mixed-methods studies). Each study is assessed according to two screening questions (identical across study types) and 5 design-specific items. An ad hoc supplementary question was also added to the MMAT to evaluate the quality of statistics as this aspect is not covered in the MMAT. Quality of

statistical analysis and data presentation was considered as appropriate if study authors 1) used adequate statistical tests considering the research question and type of data, 2) reported detailed results (value of test statistics, p value) of the statistical tests, and 3) reported effect size indicators as well.

A summary score (ranging from 0 to 8) was also created to facilitate the comparison of studies in terms of overall methodological quality regardless of their designs. This summary score was calculated as the simple sum of the 2 screening and 5 design-specific items of the MMAT plus the item on quality of statistics (adequate methodological characteristics on the given area coded as 1, while inadequate methodological characteristics quantified as 0).

If effect size indicators were not reported but the published descriptive data allowed the authors of the present study to calculate those, then the results of these calculations were added to the report with a reference to the fact that these data were not part of the original publication but calculated based on those. Where both effect size indicators and descriptive data allowing the calculation of those were missing (altogether or for certain subgoups), three attempts were made to gather the raw data from study authors. This effort was successful in two cases (Krüger & Schmidt-Michel, 2003; Langlotz, 2006) and unsuccessful in another two cases (Höppner, 2006; Sethi, 2009).

As a rule of thumb, we considered 0.2 as a threshold for small effect, 0.5 for moderate effect, and 0.8 for large effect in case of Cohen's d; and 0.01 as a threshold for small effect, 0.06 for moderate effect, and 0.14 for large effect in case of η^2 ; while the corresponding thresholds for r were 0.1, 0.3 and 0.5, respectively (J. Cohen, 1988). Finally, a formal statistical analysis (Mann-Whitney test) was also conducted to examine if overall methodological quality (using the summary score) was independent of the reported effectiveness of the intervention (using the simple study conclusion variable: significant positive effects were reported or not). Effect size r was calculated using the following formula: z/\sqrt{n} . The software Statistical Package for the Social Sciences, Version 25 (IBM SPSS, 2017) was used for the analysis.

Quantitative synthesis (meta-analysis)

The most frequently reported outcome indicator in the included studies was an omnibus (not diagnosis-specific) indicator of psychopathology; therefore, a meta-analysis was performed on the five studies that evaluated the effectiveness of family constellation therapy in this regard. As

different assessment tools (Global Severity Index of the SCL-90-R, Personality Assessment Inventory, Positive and Negative Symptom Scale, and Outcome Questionnaire 45.2) and so scale ranges were used in these studies, standardized difference in means (Hedges' g) was used as the effect size indicator. Where follow-up data from several assessment points were reported (Langlotz, 2006; Weinhold et al., 2013), all data points were considered when computing the effect size (for raw data entered into the analysis and time-point-level effect sizes, please see Supplementary file 'Raw meta-analysis data'). The intent of this analysis was to provide results generalizable to comparable populations; and therefore, the random-effects model was employed for the analysis. Given the significant differences across study designs, a subgroup analysis was also performed using a dichotomous (controlled vs. non-controlled) study design variable as a moderator.

Heterogeneity in effect sizes across studies was assessed with the Q and the I^2 statistic. The Q statistic provides a test of the null hypothesis that all studies in the analysis share a common effect size. If all studies shared the same effect size, the expected value of Q would be equal to the degrees of freedom. The I^2 statistic provides a percentile estimate for the proportion of variance in observed effects attributable to variance in true effects rather than sampling error.

Finally, two interval estimates were also calculated. In addition to the confidence interval for the overall effect size (i.e., precision of the estimate), the prediction interval was also calculated (based on τ as an estimate of the standard deviation of the true effect sizes) to estimate the true effect size for the universe of populations represented by the studies included in the analyses. The likelihood of publication bias was not analyzed as the low number of studies (n=5) did not make such analyses plausible. The software Comprehensive Meta-Analysis, version 3 (CMA 3; Biostat Inc., 2016) was used for these analyses.

Results

Qualitative synthesis

Background and intervention data. The traditional database search identified 1,790 records resulting in 1,283 records after deduplication. Database of the German Society of Systemic Constellations contained 2,914 entries resulting in a total of 4,197 records to screen. At this stage, 4,130 records were excluded due to being unrelated to the target intervention or not containing

empirical data, resulting in 67 records to assess for eligibility. Based on the evaluation of the full texts, further 55 studies were excluded (Figure 1).

Altogether, 12 studies met inclusion criteria representing 10 independent samples (3 papers analyzed the same sample) with a total sample size of 568. Bibliographical data and reason for exclusion for the 44 studies passing screening stage but failing to satisfy eligibility criteria are available as supplementary material to this article. The vast majority of included studies were conducted in Germany, while a single study was conducted in Australia, South Africa, and the UK each. All but one study employed a mixed sample of men and women (mean percentage of women=75.8%), with the exception of the study by Langlotz (2005) where information on the participants' gender was not reported. Most studies (n=5) were published as peer-reviewed journal articles or theses/dissertations (n=3); however, 2 books and 2 online reports were also among the included studies. The intervention was most often delivered in the format of a 2- or 3-day single workshop, with two exceptions, where shorter (1 to 4 hour) workshops were held on a repeated basis. More detailed description of the previously listed and additional variables (exact ratio of women in the sample, age characteristics, exact length of follow-up time, training background of intervention provider, and private vs. public setting of intervention delivery) are described in the original data extraction tool published as online supplementary material to this article.

Methodological data. Most studies (n=7) had a single group, pre-post design, two studies used a non-randomized controlled design, and two additional studies employed a randomized controlled design (one of them reported in 2 papers). The post-intervention follow-up time ranged from 0 (no follow-up after post-intervention assessment) to 12 months (M=16.8 weeks, SD=19.0 weeks). Inadequacy of methodological rigor was most frequent in relation to a lack of attempt to control for confounders and conducting/reporting statistical analyses (e.g., no effect size indicators). Methodological evaluation of each included study is reported in Table 2.

Outcome data. Most important characteristics of the study samples, the outcome variables and the main results are summarized in Table 3. The included studies considered a large variety of outcome variables ranging from indicators of overall psychological wellbeing and self-efficacy through interpersonal relationships (mainly with a focus on family relationships) to psychopathology (e.g., depression, overall psychopathology level). Out of the twelve studies included, authors of nine studies reported statistically significant treatment benefits in connection to participation in family constellation therapy with largely variable effect sizes (Table 3).

The analysis examining the relationship between methodological quality and study outcome indicated that those studies that reported statistically significant treatment benefits ($M_{methodology\ total\ score}$ =6.11±0.17) were of significantly higher methodological quality (Mann-Whitney U=3.00, p=0.041) than those not reporting significant, positive intervention outcomes ($M_{methodology\ total\ score}$ =4.33±0.58). The magnitude of the difference was large (r=0.59).

Safety / **tolerability.** The small majority of studies (n=7) also explicitly investigated introgenic effects emerging either attributed to the intervention by participants or merely occurring during the follow-up time. Out of these 7 studies, authors of four studies reported minor or moderate negative effects in a small proportion (5-8%) of participants that theoretically could have been linked to participation in the intervention.

Langlotz (2005) described that some participants of the study reported becoming emotionally upset, confused or exhausted during / immediately after the intervention, which these participants considered as a necessary element of the intervention process. This author also reported on intervention participants whose psychopathology scores increased significantly immediately after the intervention (at the end of the 2-day workshop), but even in these cases, at follow-up, scores decreased well below baseline scores. In another study, Langlotz (2006) reported that out of the 21 intervention participants, 1 individual (4.8%) showed clinically significant elevation in psychopathology scores immediately after the intervention. In this case, no follow-up interview was conducted to clarify if the deterioration could have been linked to the intervention or external factors (e.g., negative life event during the time of the workshop).

In Höppner's study, all participants were offered the opportunity to contact a therapist on the phone should they feel that the intervention destabilized them (Höppner, 2006). The author reported that out of the 81 participants², four individuals used this opportunity, three of whom only wished to further elaborate on their interpretation of the intervention, while one participant (1.2%) reported a drastic worsening in relation to an interpersonal relationship. The same author also reported that according to the 5-month follow-up questionnaire, 4 individuals (5.0%) reported a deterioriation in the subjective, overall evaluation of their condition.

11

² The original sample consisted of 81 individuals. In Table 3 and 4, 70 is reported as sample size as this is the number of participants about whom the author displayed enough data to allow the calculation of effect sizes.

Finally, in a study of 48 participants, 4 individuals (8.3%) reported negative outcomes or side effects such as short-term, negative physical symptoms (n=1), intimate relationship break-up (n=1), increased loneliness in the family (n=1), and workplace bullying (n=1), which respondents linked to their participation in the intervention (Rieger & Stückemann, 1999).

Quantitative synthesis of data on psychopathology

The meta-analytic investigations resulted in a Hedges' g of 0.53 (Table 4) indicating that on average, psychopathological symptom scores of those who participated in the intervention decreased 0.53 standard deviation (moderately strong effect) compared to their pre-intervention scores or the control group – depending on study design. The confidence interval for the effect size ranged from 0.39 to 0.68 and the Z-value was 7.20 with a corresponding p-value of <0.001 suggesting that the null hypothesis – i.e. that the effect size would be zero – is to be rejected.

The Q-value was 2.79 with 4 degrees of freedom and a corresponding p-value of 0.595. Thus, the observed dispersion was actually less than what would be expected by chance suggesting that there is no evidence that the true effect size varies from study to study. The I^2 statistic was 0% indicating that none of the variance in observed effects reflects variance in true effects but that all of it is due to sampling error. The variance of true effects in log units (τ^2) was <0.001 and the standard deviation of true effects in log units (τ) was <0.001 indicating that the between-study variance is estimated as zero. The 95% prediction interval for the overall effect size was 0.296 to 0.753 indicating that in the universe of populations represented by the studies included in the current analyses, the true effect size in 95% of cases would fall somewhere in this range.

The subgroup analysis indicated that the pooled effect size of studies with a controlled design (g=0.50, CI=0.23-0.76) was not statistically different (Q=0.105, p=0.746) from that of studies with an uncontrolled design (g=0.55, CI=0.37-0.72), indicating that the mean effect size is in the moderate range both for studies with controlled and non-controlled designs.

Discussion

Intervention effectiveness

The goal of this study was to systematically gather and synthesize the quantitative evidence regarding the effectiveness of family constellation therapy in terms of mental health outcomes.

Based on the results of this review, we can conclude that the quantity and overall quality of the evidence is low, the latter mainly due to the frequently lacking control group and the typically short follow-up period. Most likely, not independently from these facts, the majority of the evidence has been published in outlets other than peer-reviewed journals. Importantly though, the explorative analysis examining the relationship between methodological quality and study outcomes indicated that those studies that reported treatment benefits were of higher methodological quality suggesting that the evidence may be more convincing if additional higher quality studies will be published.

Out of the 12 studies included in the present review, 9 showed significant treatment benefits post-intervention. The outcome variables selected by study authors were quite diverse, which is not surprising considering the major role current (Weissman, Markowitz, & Klerman, 2008) or the internal representation of early (Young, Klosko, & Weishaar, 2003) interpersonal relationships play in our bio-psycho-social health. The studies showing no significant treatment benefit reported comparable effect sizes to those reported in the studies showing statistically significant treatment benefit, suggesting that the former studies were underpowered (had too low sample sizes to detect existing treatment effects).

The results of the meta-analysis on indices of general psychopathology from 5 studies with comparable outcomes indicated a moderately strong treatment effect – independently of the controlled (n=2) or uncontrolled (n=3) nature of the study design. The variance of effect sizes across studies was estimated to be zero, which is most likely an underestimate due to the low number of studies included; most likely, the impact of this treatment also varies by population. However, the present findings suggest that the variation in effects is minor; indicating that the impact of the treatment for all comparable populations (self-selected participants from the general population) falls close to the mean effect size reported here.

In summary, data from the included quantitative, prospective studies suggest that family constellation therapy is a consistently and moderately effective intervention in the general population to decrease psychopathological symptoms. These findings are parallel to the results of retrospective effectiveness studies, which also indicated treatment benefit. For instance, in a study of 57 Austrian respondents, approximately $2/3^{\rm rd}$ of participants reported increased happiness, courage, optimism, and coping abilities as a result of the intervention (Jost, 2007), while in a study of participants from Germany, 92% of the respondents reported that the intervention was helpful

for them (Mraz, 2006). In a retrospective study of English-, French- and Russian speaking participants, 87% of those who sought treatment for interpersonal difficulties (n=119) reported that their problems resolved as the outcome of the intervention, while the same value in the case of mental health issues (n=31) was 90% (Thomas, 2010). A study of 209 Hungarian participants reported that out of 26 quality of life domains covered in the evaluation, participants experienced statistically significant improvement in 23 areas after the intervention (Zseni et al., 2011). An interesting aspect of this study was the investigation of the effect of problem severity, with analyses indicating that the intervention was more effective among individuals with less severe mental health or interpersonal relationship challenges (the same was reported by Höppner, 2006). Finally, authors of a study – examining a sample of 139 inpatient substance use treatment participants from Germany – reported that intervention participants completed the entire treatment regime with a significantly higher likelihood (81%) than those who did not participate (50%) in family constellation therapy (Mahr & Brömer, 2008).

Tolerability / safety

Considering theoretical concerns (Deutsche Gesellschaft für Systemische Therapie und Familientherapie, 2003; Talarczyk, 2011) and anecdotic data on the risks of family constellation therapy (Langlotz, 1998b, 2001), a major focus of our work was to summarize data on tolerability. Altogether, authors of four studies reported minor or moderate negative effects in a small proportion (5-8%) of participants that theoretically could have been linked to participation in the intervention [Jost (2007) reported similar proportions (3.4%) in their retrospective study].

The non-intended effects / correlating events reported included ruptures in interpersonal relationships, short-term somatic or mental health symptoms, or unfavorable change in other problem areas the participants worked on during the intervention. The studies reviewed here suggest that the often strong emotional responses family constellation therapy can generate in a very condensed time frame can facilitate improvement but can also temporarily destabilize individuals with less stable mental health status. This aspect of the results points towards the importance of post-intervention screening and providing intervention participants with the opportunity to receive professional mental health support to process their experience if needed (Langlotz, 2005). It is also worthy of mentioning that in all of the studies where iatrogenic effects were studied, the intervention provider was a psychologist or psychiatrist and also an expert in

family constellation therapy (for details, see original data extraction tool published as online supplementary material to this article). This leaves the question open, whether iatrogenic effects are more prevalent or severe if the intervention is provided by less experienced / trained professionals, an issue which deserves attention in future studies.

Strengths and limitations

A major strength of the present systematic review is the comprehensive search process including a large number of data bases and six languages. In addition, this is the first study using meta-analytic techniques in an attempt to quantitatively summarize outcome data in relation to family constellation therapy. Further, two researchers independently assessed each included study contributing to a higher reliability of the data extraction process. Finally, the review is based on an a priori developed and publicly registered research protocol.

Despite these strengths, a number of limitations should be acknowledged as well. First, both the electronic searches, the screening process and checking for eligibility criteria was completed by one researcher only decreasing the reliability of these processes. Most importantly, due to the often lacking controlled design, the possibility cannot be ruled out that the reported beneficial changes are the results of external factors and not the intervention itself. However, as studies with controlled designs indicated effect sizes similar to the combined effect of the meta-analysis in relation to psychopathology, and as studies with numerous assessment points indicated improvement right after the intervention (Langlotz, 2006) but not between the pre-intervention assessment points (Höppner, 2006), it is plausible to assume that the results are truly indicative of the effectiveness of family constellation therapy.

The overall low number of studies included (and the even lower number of peer-reviewed publications among them) also limits the reliability of the findings. This is especially true for the meta-analytic investigations, where it was not feasible to conduct an important aspect of meta-analyses due to the low number of studies: the estimation of publication bias. It is worthy of mentioning though that even with the Cochrane collaboration, the median number of studies included into a systematic review is 3 (Davey, Turner, Clarke, & Higgins, 2011). Therefore, while we are aware that the reliability of the findings is suboptimal due to the low amount of evidence available, we argue that synthetizing these data is helpful and necessary to at least preliminarily inform clinical practice and inspire further research.

A further limitation of the present review is the lack of distinction made among subtypes of family/systemic constellations: interventions provided in the studies reported on in this review were considered as a homogenous, single type of intervention as empirical studies most often do not specify the subtype of the intervention they investigated. However, there are numerous subtypes (Langlotz, 2010; Nelles, 2007) and formats of delivery (e.g., 2-day single workshop vs. repeated, few-hour sessions over the course of several weeks) for this form of group therapy and their effectiveness might vary.

Future directions

Future authors interested in studying the effectiveness of family constellation therapy are encouraged to replicate the previous findings in adequately powered investigations employing controlled (preferably randomized controlled) designs and several intervention providers simultaneously to allow the explicit examination of therapist effects. Studies with longer follow-up time (6 months or more) could significantly contribute to our knowledge regarding the stability of treatment benefits. In view of the ongoing debate on the safety of the intervention, further studies with an explicit focus on tolerability could help us better understand in which populations and under which conditions (e.g., therapist's training background, length of debriefing, accessibility of support post-intervention) can the intervention be delivered in a safe manner.

Considering the ongoing diversification within family constellation therapy, authors of future studies are also encouraged to specify the mode of delivery and subtype of family/systemic constellations they employ when reporting on the effectiveness of the intervention. In addition, synthetizing the relatively large number of qualitative studies we have identified through our searches (Fig 1; for detailed bibliographic data of these records, see Supplementary file 'Excluded items') could also contribute to a better understanding of the effectiveness and treatment mechanisms of family constellation therapy.

Finally, there is a huge gap between the theory and anecdotic evidence versus the solid research data related to the application of family constellation therapy for a large variety of specific mental disorders. Authors have described the use of this form of brief group therapy with clients struggling with psychosomatic- (Baitinger, 1999; Elsner & Kölle, 2010; Hausner, 2015), eating-(Bourquin, 2011), mood- (Asztalos, Angster, & Pusztai, 2011; Brink, 1998; Ramos & Ramos, 2019), anxiety- (Essen, 1998; Franke, 1996), substance use- (Döring-Meijer & Hellinger, 2000;

Gemeinhardt, 2006; Ingwersen, 2000; Mahr & Brömer, 2008), trauma-related (Assel, 2009; Nazarkiewicz & Bourquin, 2017; Ruppert, 2006) and even psychotic disorders (Hellinger, 2001; Langlotz, 1998a; Ruppert, 2004; Weber & Drexler, 2002), while quantitative empirical research to date has almost exclusively focused on samples from the general population. Therefore, there is a clear need to formally investigate the efficacy / effectiveness and safety / tolerability of the intervention in specific client / patient populations to better understand to whom family constellation therapy can be beneficial on their journey toward recovery or simply a happier and more fulfilling life.

References

- APA Publications and Communications Board Working Group on Journal Article Reporting Standards. (2008). Reporting standards for research in psychology: Why do we need them? What might they be? *American Psychologist*, 63(9), 839-851. http://dx.doi.org/10.1037/0003-066X.63.9.839
- Assel, B. (2009). Von der Familienaufstellung zur Traumaaufstellung [From family constellation to trauma constellation]. *Praxis der Systemaufstellung*, *12*(1), 35-42.
- Asztalos, M., Angster, M., & Pusztai, I. (2011). Family constellations in therapy-resistant cases of patients suffering from depression and a wish to die. *International Journal of Psychiatry in Clinical Practice*, 2(Sup2), 15. http://dx.doi.org/10.3109/13651501.2011.626553
- Baitinger, H. (1999). Die Wirkungen des Familien-Stellens auf den psychosomatischen Symptomkomplex und die Bedeutung einzeltherpeutischer Nacharbeit [Effects of family constellation on psychosomatic symptoms and the relevance of aftercare in individual therapy]. *Praxis der Systemaufstellung*, 1(2), 38-43.
- Bornhäuser, A., & Wolff, J.-E. (2014). RCT-Studie: Ziele der Studienteilnehmer und Zielerreichung [RCT study: Goals and goal attainment of study participants]. In J. Weinhold, A. Bornhäuser, C. Hunger, & J. Schweitzer (Eds.), *Dreierlei Wirksamkeit. Die Heidelberger Studie zu Systemaufstellungen* (pp. 134-147). Heidelberg: Carl-Auer Verlag.
- Bourquin, P. (2011). Familienstellen, Magersucht und Bulimie [Family constellation, anorexia, and bulimia]. *Praxis der Systemaufstellung, 14*(1), 34-39.

- Brink, O. (1998). Diagnose und Therapie bei depressiven und suizidalen Klienten mit dem Familien-Stellen [Diagnosis and therapy of clients with depression and suicidality using family constellation]. In G. Weber (Ed.), *Praxis des Familien-Stellens* (pp. 299-304). Heidelberg: Carl-Auer-Systeme.
- Broughton, V. (2006). Constellations in an individual setting. *Self & Society*, *33*(4), 20-26. http://dx.doi.org/10.1080/03060497.2006.11086256
- Butollo, W., Franke, U., & Hellinger, B. (2017). The river never looks back: Historical and practical foundations of Bert Hellinger's family constellations. Heidelberg: Carl-Auer Verlag.
- Choi, K.-O., & Oh, K.-Y. (2018). [A study on married women's experiences in family constellation against induced abortion]. *Journal of the Korea Contents Association*, 18(9), 294–307. http://dx.doi.org/10.5392/JKCA.2018.18.09.294
- Chu, V. (2008). Neugeburt einer Familie. Familienstellen in der Gestalttherapie [Rebirth of a family. Family constellation in Gestalt therapy]. Wuppertal: Peter Hammer Verlag.
- Cohen, D. B. (2006). Family constellations: An innovative systemic phenomenological group process from Germany. *The Family Journal*, 14(3), 226-233. http://dx.doi.org/10.1177/1066480706287279
- Cohen, D. B. (2009). *I carry your heart in my heart: Family constellations in prison*. Heidelberg: Carl-Auer-Systeme.
- Cohen, J. (1988). Statistical power analysis for the behavioral sciences. New York, NY: Routledge Academic.
- Davey, J., Turner, R. M., Clarke, M. J., & Higgins, J. P. T. (2011). Characteristics of meta-analyses and their component studies in the Cochrane Database of Systematic Reviews: a cross-sectional, descriptive analysis. *BMC Medical Research Methodology*, 11(1), 160. http://dx.doi.org/10.1186/1471-2288-11-160
- Deutsche Gesellschaft für Systemische Therapie und Familientherapie. (2003). Stellungnahme der DGSF zum Thema "Familienaufstellungen" [Statement of the German Society for Systemic Therapy and Family Therapy on the topic of family constellations]. Retrieved from https://www.dgsf.org/themen/berufspolitik/hellinger.htm
- Döring-Meijer, H., & Hellinger, B. (2000). Leiden ist leichter als lösen. Familienaufstellungen mit Suchtkranken. Ein Praxisbuch mit Bert Hellinger [Suffering is easier than resolving.

- Family constellation with people suffering from addictions. A course book with Bert Hellinger]. Paderborn: Junfermann Verlag.
- Elsner, H., & Kölle, H. (2010). Aufstellen im Krankenhaus für Psychosomatische Medizin [Using constellations in clinics for psychosomatic medicine]. *Praxis der Systemaufstellung, 13*(1), 30-36.
- Essen, C. (1998). Aufstellungen bei Angstsymptomatik und Panikattacken [Constellations for anxiety and panic disorders]. In G. Weber (Ed.), *Praxis des Familien-Stellens* (pp. 305-312). Heidelberg: Carl-Auer-Systeme.
- Franco de Sá, R., Nogueira, J., & De Almeida Guerra, V. (2019). Traditional and complementary medicine as health promotion technology in Brazil. *Health Promotion International*, 34(Suppl1), 74-81. http://dx.doi.org/10.1093/heapro/day087
- Franke, U. (1996). Systemische Familienaufstellung. Eine Studie zu systemischer Verstrickung und unterbrochener Hinwendung unter besonderer Berücksichtigung von Angstpatienten [Systemic family constellation. A study into systemic entanglements and the interrupted movement with a special attention on individuals with anxiety disorders]. München: Profil.
- Geils, C., & Edwards, S. D. (2018). Extended family constellations workshop efficacy on intuition measure and experience. *Journal of Psychology in Africa*, 28(3), 224-228. http://dx.doi.org/10.1080/14330237.2018.1475527
- Gemeinhardt, B. (2006). Systemisch-lösungsfokussierte Gruppentherapie im Suchtbereich [Systemic solution-focused group therapy in drug rehabilitation]. In R. Basdekis Jozsa & M. Krausz (Eds.), *Gruppentherapie in der Suchtbehandlung. Konzepte und praktisches Vorgehen* (pp. 215-235). Stuttgart: Klett Cotta.
- Georgiadou, S. (2012). Participants' experiences in Hellinger's family constellation work: A grounded theory study. Dissertation. University of Louisiana at Monroe, Ann Arbor.
- Goode, K. P. (2015). Enhancing the affective domain in order to reduce fear of death in first-year student nurses. Dissertation. University of Hertfordshire, Hatfield.
- Häuser, W., Klein, R., & Schmidt-Keller, B. (1998). Familienaufstellen mit Bert Hellinger aus der Sicht teilnehmender Klientinnen und ihrer Therapeutinnen [Family constellation with Bert Hellinger from the clients' and their therapists' perspective]. In G. Weber (Ed.), *Praxis des Familien-Stellens* (pp. 478-488). Heidelberg: Carl-Auer-Systeme.

- Hausner, S. (2015). Even if it costs me my life: Systemic constellations and serious illness. New York: Gestalt Press.
- Hellinger, B. (1994). Ordnungen der Liebe [Orders of Love]. Heidelberg: Carl Auer.
- Hellinger, B. (2001). Liebe am Abgrund: ein Kurs für Psychose-Patienten [Love in the abyss: A course for patients with psychosis]. Heidelberg: Carl-Auer-Systeme.
- Hong, Q. N., Pluye, P., Fàbregues, S., Bartlett, G., Boardman, F., Cargo, M., . . . Vedel, I. (2018).
 Mixed Methods Appraisal Tool (MMAT). Version 2018. Retrieved from http://mixedmethodsappraisaltoolpublic.pbworks.com/w/file/fetch/127425851/MMAT_2
 018_criteria-manual_2018-04-04.pdf
- Höppner, G. (2006). "Heilt Demut wo Schicksal wirkt?" Evaluationsstudie zu Effekten des Familien-Stellens nach Bert Hellinger ['Does humility heal where fate is in effect?' An effectiveness study on Bert Hellinger's family constellations]. Heidelberg: Carl-Auer Verlag.
- Hunger, C., Bornhäuser, A., Link, L., Schweitzer, J., & Weinhold, J. (2014). Improving experience in personal social systems through family constellation seminars: Results of a randomized controlled trial. *Family Process*, *53*(2), 288-306. http://dx.doi.org/10.1111/famp.12051
- Hunger, C., Weinhold, J., Bornhäuser, A., Link, L., & Schweitzer, J. (2015). Mid- and long-term effects of family constellation seminars in a general population sample: 8- and 12-month follow-up. *Family Process*, *54*(2), 344-358. http://dx.doi.org/10.1111/famp.12102
- Hurley, J., Koenning, M., & Bray, A. (2018). Responding to intergenerational psychological trauma: a literature review paper on the place of family constellation therapy. *Psychotherapy and Counselling Journal of Australia*, 6(1), 1-14.
- Ingwersen, F. (2000). Kinder in der Todesnähe das Geheimnis der Drogensucht. Suchtbehandlung mit Familienaufstellungen in einer Psychosomatischen Klinik [Children in the near of death the secret of drug addiction. Addiciton therapy with family constellation in a clinic for psychosomatic disorders]. In H. Döring-Meijer (Ed.), *Die entdeckte Wirklichkeit* (pp. 107-124). Paderborn Junfermann.
- Jafferany, M., Capec, S., Yaremkevych, R., Andrashko, Y., Capec, G., & Petrek, M. (2019).
 Effects of family constellation seminars on itch in patients with atopic dermatitis and psoriasis: A patient preference controlled trial. *Dermatologic Therapy*, 32(6), e13100. http://dx.doi.org/10.1111/dth.13100

- Jost, R. (2007). Familienaufstellungen im Urteil der Klienten. Eine retrospektive Befragung [Family constellations according to clients' judgement. A retrospective inquiry]. *Blickpunkt EFL-Beratung*, *4*(18), 56-59.
- Junge, G. (1998). Familienaufstellung nach Bert Hellinger. Eine qualitative Untersuchung anhand einer Nachbefragung von Klienten [Family constellations of Bert Hellinger. A qualitative study based on client interviews]. Universität Hamburg, Hamburg.
- Krüger, M., & Schmidt-Michel, P.-O. (2003). Familienaufstellungen in der Psychiatrischen Tagesklinik [Family constellations in the psychiatric day clinic]. *Krankenhauspsychiatrie*, 14(3), 90-94. http://dx.doi.org/10.1055/s-2003-42672
- Laireiter, A.-R., & Mitterhuemer, J. (2011). Die Beurteilung von Familienaufstellungen Eine Konsumenten-Studie [Evaluation of family constellations A consumer study]. *Psychologie in Österreich*, *31*(2-3), 136-147.
- Langlotz, E. R. (1998a). Familien-Stellen mit Psychosekranken: ein Kurs mit Bert Hellinger [Family constellation with psychotic patients: A course with Bert Hellinger]. Heidelberg: Carl-Auer-Systeme.
- Langlotz, E. R. (1998b). Wann kann systemische Familientherapie schaden? [When can systemic family therapy cuase harm?]. *Praxis der Systemaufstellung, 1*(1), 35-39.
- Langlotz, E. R. (2001). Kann Familien-Stellen schaden? Erfahrungen eines Psychiaters [Can family constellation cause harm? Experiences of a psychiatrist]. In G. Weber (Ed.), Derselbe Wind lässt viele Drachen steigen - Systemische Lösungen im Einklang (pp. 402-407). Heidelberg: Carl-Auer-Systeme.
- Langlotz, E. R. (2005). Zur Effizienz des Familienstellens [On the effectiveness of family constellations]. Retrieved from http://www.e-r-langlotz.de/systemische_familientherapie/public_effizienzFam.php
- Langlotz, E. R. (2006). Effizienzforschung "prozessorientiertes" Familienstellen. Neue Ergebnisse mit dem SCL 90 R [Effectiveness of process-oriented family constellations. New results with the SCL-90-R]. Retrieved from http://www.e-r-langlotz.de/systemische_familientherapie/public_effizienzforschung.php
- Langlotz, E. R. (2010). Systemische Selbst-Integration. Systemaufstellung in der psychiatrischen Praxis [Systemic self-integration. System constellation in the psychiatric praxis]. *Praxis der Systemaufstellung*, 13(2), 74-76.

- Mahr, A., & Brömer, H. (2008). Aufstellungen in der Suchtrehabilitation Erfahrungen und Ergebnisse [Constellations in addiction rehabilitation. Experiences and results]. *Praxis der Systemaufstellung*, 11(2), 66-74.
- McQuillin, J., & Welford, E. (2013). How many people are gathered here? Group work and family constellation theory. *Transactional Analysis Journal*, 43(4), 352-365. http://dx.doi.org/10.1177/0362153713519743
- Mraz, R. (2006). Nachgeprüft. Ergebnisse einer 10–Jahres–Katamnese aus über 850 Aufstellungen [Proven. Results of a 10-year retrospective study based on more than 850 constellations]. Praxis der Systemaufstellung, 9(2), 94-101.
- Nazarkiewicz, K., & Bourquin, P. (Eds.). (2017). Trauma und Begegnung. Praxis der Systemaufstellung [Trauma and encounter. The practice of system constellations] Göttingen: Vandenhoeck & Ruprecht.
- Nelles, W. (2005). Die Hellinger-Kontroverse. Fakten, Hintergruende, Klarstellungen [The Hellinger controversy. Facts, background, clarifications]. Freiburg im Breisgau; Basel; Wien: Herder.
- Nelles, W. (2007). Klassisches Familien-Stellen, Bewegungen der Seele, Bewegungen des Geistes-Wohin bewegt sich die Aufstellungsarbeit? [Classic family constellation, movements of the soul, movements of the spirit mind Where does constellation work go?]. *Praxis der Systemaufstellung*, 10(1), 32-45.
- North American Systemic Constellations. (2019a). Medical Professionals & Constellations. Retrieved from http://www.nasconnect.org/medical-professionals.html
- North American Systemic Constellations. (2019b). Psychotherapists & Constellations. Retrieved from http://www.nasconnect.org/psychotherapists.html
- Orban, P. (2008). Kursbuch Aufstellungsarbeit: Grundlagen-Methoden-Beispiele [Text book of family constellation therapy: Theory, methods, examples]. München: Kösel.
- Pritzker, S. E., & Duncan, W. L. (2019). Technologies of the social: Family constellation therapy and the remodeling of relational selfhood in China and Mexico. *Culture, Medicine, and Psychiatry*, 43(3), 468-495. http://dx.doi.org/10.1007/s11013-019-09632-x
- Ramos, S., & Ramos, J. A. (2019). Process of change and effectiveness of family constellations:

 A mixed methods single case study on depression. *The Family Journal*, 27(4), 418-428. http://dx.doi.org/10.1177/1066480719868706

- Reuter, E. (2005). *Gehirn-Wäsche: Macht und Willkür in der" systemischen Psychotherapie" nach Bert Hellinger [Brainwash: Power and despotism in the 'systemic psychotherapy' of Bert Hellinger]*. Berlin: Peter Lehmann Antipsychiatrieverlag.
- Rieger, D., & Stückemann, L. (1999). "Finden was wirkt" (Hellinger 1993). Eine explorative Untersuchung zur Wirkungsweise der systemischen Familienaufstellung nach Bert Hellinger (Diplomarbeit). ['Finding what works' (Hellinger, 1993). An explorative study on the effectiveness of Bert Hellinger's systemic family constellations. Thesis]. Universität Freiburg, Freiburg.
- Ruppert, F. (2004). Verwirrte Seelen Psychosen aus Sicht einer systemischen Psychotraumalogie [Confused souls Psychoses from the viewpoint of a systemic psychotraumatology]. Karlsruhe: Verlag Angelika Steinhardt.
- Ruppert, F. (2006). Constellations under the sign of multi-generational systemic psychotraumatology. *Self & Society*, *33*(4), 10-19. http://dx.doi.org/10.1080/03060497.2006.11086255
- Schneider, J. R. (2010). Zur neuerlichen Kritik an der Aufstellungsarbeit und der DGfS [On the renewed criticism of constellation work and the German Association for Systemic Constellations]. *Praxis der Systemaufstellung*, *13*(1), 105-110.
- Schumacher, T. (2000). Systematische Strukturen in Familie und Organisation: eine Studie zu Auswirkungen von Familienaufstellungen auf subjektive Beziehungsbilder [Systematic structures in families and organizations: A study into the effects of family constellations on the perception of interpersonal relationships]. Bonn: Rheintal Institut Verlag.
- Sethi, Y. (2009). *Does the process of family constellations improve relationships and wellbeing?*Thesis. Australian College of Applied Psychology, Sydney.
- Stiefel, I., Harris, P., & Zollmann, A. W. F. (2002). Family constellation A therapy beyond words. *Australian and New Zealand Journal of Family Therapy*, 23(1), 38-44. http://dx.doi.org/10.1002/j.1467-8438.2002.tb00484.x
- Stones, B. (2006). A brief history of Bert Hellinger's family constellations. *Self & Society*, *33*(4), 5-9. http://dx.doi.org/10.1080/03060497.2006.11086254
- Talarczyk, M. (2011). Family Constellation Method of Bert Hellinger in the context of the Code of Ethics for Psychotherapists. *Archives of Psychiatry and Psychotherapy*, *13*(3), 65-74.

- Thomas, G. K. (2010). Therapy in the new millennium: New sciences and their application to therapy. Effectiveness of systemic family constellations. Thesis. California State University, Northridge.
- Weber, G. (1993). Zweierlei Glück. Die systemische Psychotherapie Bert Hellingers [Capricious good fortune. Bert Hellinger's systemic psychotherapy]. Heidelberg: Carl Auer.
- Weber, G., & Drexler, D. (2002). Familien-Stellen bei Psychosen [Family constellation in psychoses]. *Psychotherapie im Dialog*, *3*(3), 243-247. http://dx.doi.org/10.1055/s-2002-34539
- Weinhold, J., Hunger, C., Bornhäuser, A., Link, L., Rochon, J., Wild, B., & Schweitzer, J. (2013). Family constellation seminars improve psychological functioning in a general population sample: results of a randomized controlled trial. *Journal of Counseling Psychology*, 60(4), 601-609. http://dx.doi.org/10.1037/a0033539
- Weinhold, J., & Reinhard, A. (2014). Der Forschungsstand zur Wirksamkeit von Systemaufstellungen [The current state of the evidence regarding the effectiveness of systemic constellations]. In J. Weinhold, A. Bornhäuser, C. Hunger, & J. Schweitzer (Eds.), *Dreierlei Wirksamkeit. Die Heidelberger Studie zu Systemaufstellungen* (pp. 36–63). Heidelberg: Carl-Auer.
- Weissman, M. M., Markowitz, J. C., & Klerman, G. (2008). *Comprehensive guide to interpersonal psychotherapy*. New York: Basic Books.
- Young, J. E., Klosko, J. S., & Weishaar, M. E. (2003). Schema therapy: A practitioner's guide. New York: Guilford Press.
- Zseni, A., Varga, K. S., Angster, M., Béleczki, N., Füzér, G., Néveri, E., . . . Lőrik, D. (2011). Első lépés a családállítás hatásainak nyomában [First steps of tracing the effects of family constellations A follow-up study]. *Magyar Pszichológiai Szemle*, 66(2), 269-298. http://dx.doi.org/10.1556/MPSzle.66.2011.2.2

Table 1. Methodological characteristics of the studies included

Study	Design	MMAT	MMAT	MMAT	MMAT	MMAT	MMAT	MMAT	Statistics	Total
		S 1	S2	1	2	3	4	5		score
Geils & Edwards,	single group,	1	1	0	1	1	0	1	0	5
2018	pre-post									
Goode, 2015	randomized-	1	1	1	1	0	0	0	0	4
	controlled trial									
Höppner, 2006	two-group, non-	1	1	1	1	1	0	1	0	6
	randomized									
	crossover									
Hunger et al., 2014†	randomized-	1	1	1	1	1	0	1	1	7
	controlled trial									
Hunger et al., 2015†	single group,	1	1	1	1	1	0	1	1	7
	pre-post									
Krüger & Schmidt-	two-group,	1	1	1	1	1	1	1	0	7
Michel, 2003	matched control									
	group design									
Langlotz, 2005	single group,	1	1	0	1	1	0	1	0	5
	pre-post									
Langlotz, 2006	single group,	1	1	0	1	0	0	1	0	4
	pre-post									

Rieger &	single group,	1	1	1	1	1	0	1	1	7
Stückemann, 1999	pre-post									
Schumacher, 2000	single group,	1	1	1	1	0	0	1	0	5
	pre-post									
Sethi, 2009	single group,	1	1	0	0	1	0	1	0	4
	pre-post									
Weinhold et al., 2013†	randomized-	1	1	1	1	1	0	1	1	7
	controlled trial									

[†] Marked studies are based on the same/partially overlapping sample

Table 2. Sample characteristics, outcome variables, and main results of the studies included

Study	Sample	Outcome variables	Results	Treat-
		(assessment tools)		ment
				benefit
Geils &	General	Self-perceived intuition	No change in intuition scores (p =0.06; d =0.30) [considering	No
Edwards,	population	(Types of Intuition	the magnitude of the effect size and the tendency toward	
2018	(n=8), South	Scale)	significance, this is most likely the result of insufficient	
	Africa		statistical power]	
Goode, 2015	Nursing	Fear of death	No change in fear of death in either the intervention	No
	students	(Multidimensional Fear	(p=0.189, d=0.42) or the no-intervention control group	
	(<i>n</i> =75), United	of Death Scale)	(p=0.810, d=0.08), improvement in the control group with	
	Kingdom		fear-of-death-specific intervention ($p=0.002$, $d=0.96$)	
Höppner,	General	Psychopathology (SCL-	Improvement in all but one target areas in the intervention	Yes
2006	population	90-R's global wellness	group: Psychopathology: p <0.001, d =0.46; positive self-	
	(n=70),	index); positive self-	image: $p < 0.001$, $d = 0.37$; general mental health: $p < 0.001$,	
	Germany	image (Frankfurter	d=0.24; self-acceptance: $p<0.001$, $d=35$; self-doubt:	
		Selbstkonzeptskalen);	p<0.001, d = 0.38 ; general self-efficacy: p =0.005, d = 0.23 ,	
		self-acceptance (Skala	external control beliefs: $p=0.268$, $d=0.09$, sense of	
		zur Erfassung der	coherence: <i>p</i> <0.001, <i>d</i> = 0.26 .	
		Selbstakzeptiurung);	No change in the control group (T1 vs. T3) in any areas.	
		general mental health	Psychopathology: $p=0.313$, $r=0.30$; positive self-image:	
		(Skalen zur Psychischen	p=0.213, $r=0.38$; general mental health: $p=0.625$, $r=0.15$;	

Study	Sample Outcome variables		Results	Treat-
		(assessment tools)		ment
				benefit
		Gesundheit); self-doubt	self-acceptance: $p=0.092$, $r=0.51$; self-doubt: $p=0.202$,	
		(Unsicherheits-	r=0.39; general self-efficacy: $p=0.590$, $r=0.16$; external	
		fragebogen); general	control beliefs: $p=0.798$, $r=0.08$; sense of coherence:	
		self-efficacy and	p=0.878, $r=0.05$. Improvement in 3 areas in the control	
		external control beliefs	group after receiving intervention (<i>n</i> =11). Psychopathology:	
		(Fragebogen zu	p=0.008, $r=0.81$; positive self-image: $p=0.022$, $r=0.69$;	
		Komptenz- und	general mental health: $p=0.074$, $r=0.54$; self-acceptance:	
		Kontrollüberzeugungen)	p=0.028, $r=0.66$; self-doubt: $p=0.173$, $r=0.41$; general self-	
		, sense of coherence	efficacy: $p=0.358$, $r=0.28$; external control beliefs: $p=0.444$,	
		(Fragebogen zur	r=0.23; sense of coherence: $p=0.721$, $r=0.11$	
		Lebensorientierung)		
Hunger et	General	Perceived quality of	Larger improvement in the intervention group than in the	Yes
al., 2014†	population	personal social system	control group in all but one assessed target areas: belonging	
	(n=208),	(Experience in Social	(T2: p =0.021, d =0.32; T3: p =0.075, d =0.27), autonomy (T2:	
	Germany	Systems Questionnaire,	p<0.001, d =0.62; T3: p <0.001, d =0.61), accord (T2:	
		the Interpersonal	p<0.001, d =0.59; T3: p =0.001, d =0.50), confidence (T2:	
		Problematic Relations	p<0.001, d =0.54; T3: p =0.031, d =0.38), experience in social	
		scale of the Outcome	systems altogether (T2: p <0.001, d =0.61; T3: p <0.001,	
		Questionnaire and the	d=0.53); interpersonal problematic relations (T2: p =0.035,	

Study	Sample	Outcome variables	Results	Treat-
		(assessment tools)		ment
				benefit
		Interpersonal Problems	d=0.32; T3: p =0.021, d =0.36); interpersonal problems (T2:	
		scale of the Tool for the	p=0.003, d=0.45; T3: p<.001, d=0.53)	
		Evaluation of the		
		Psychotherapeutic		
		Progress)		
Hunger et	General	Overall psychological	Improvement in all four areas: overall psychological	Yes
al., 2015†	population	functioning (Outcome	functioning (T2: p<0.001, d=0.41; T3: p<0.001, d=0.49);	
	(n=104),	Questionnaire); overall	overall psychological distress (T2: $p < 0.001$, $d = 0.39$; T3:	
	Germany	psychological distress	p=0.001, d=0.50); goal attainment (T2: $p<0.001, d=0.35$; T3:	
		(Questionnaire for the	p<0.001, d =0.44); perceived quality of personal social	
		Evaluation of Treatment	system (T2: <i>p</i> <0.001, <i>d</i> =0.61; T3: <i>p</i> <0.001, <i>d</i> =0.57). Clinical	
		Progress); goal	significance: depending on the indicator, reliable positive	
		attainment	change in 33-35% of participants at 8-month follow-up (T2)	
		(Incongruence	and 33-40% at 12-month follow-up (T3)	
		Questionnaire);		
		perceived quality of		
		personal social system		
		(personal domain of the		

Study	Sample	Outcome variables	Results	Treat-
		(assessment tools)		ment
				benefit
		Experience In Social		
		Systems Questionnaire)		
Krüger &	Psychiatric	Psychopathology	Decrease in psychopathology scores in the intervention	Yes
Schmidt-	outpatients with	(Positive and Negative	group ($p=0.016$, $d=0.71$), while no significant decrease in	
Michel,	a history of	Syndrome Scale;	symptom scores in the control group ($p=0.083$, $d=0.43$). No	
2003	psychiatric	PANSS)	significant time x group interaction ($p=0.239$, $\eta^2=0.076$)	
	hospitalization		most likely due to low statistical power (cf. width of	
	(n=20),		confidence interval for this study in the meta-analysis,	
	Germany		supplementary file 2)	
Langlotz,	General	Psychopathology	Reduction in all 6 psychopathology domain scores: anxiety	Yes
2005	population	(Personality Assessment	(p<.001, d=0.82), depression $(p<.001, d=0.69)$, paranoia	
	(n=35),	Inventory)	(p<.001, d=0.34), schizophrenia $(p<.001, d=0.60)$,	
	Germany		borderline features (p <.001, d =0.84), suicidal ideation	
			(<i>p</i> <.001, <i>d</i> = 0.49)	
Langlotz,	Individuals	Psychopathology (SCL-	Decrease on psychopathology indicators at all 3 follow up	Yes
2006	with at least	90-R: somatization,	assessment points with the exception of phobic anxiety at T4:	
	two SCL-90R	obsessive-compulsive,	somatization (T2 : $p=0.013$, $d=0.57$; T3 : $p=0.001$, $d=0.54$;	
	scales showing	interpersonal sensitivity,	T4: $p=0.004$, $d=0.59$), obsessive-compulsive (T2: $p<0.001$,	
	elevated values	depression, anxiety,	d=0.95; T3: p =0.001, d =0.84; T4: p =0.001, d =0.87),	

Study	Sample	Outcome variables	Results	Treat-
		(assessment tools)		ment
				benefit
	(n=21),	hostility, phobic	interpersonal sensitivity (T2 : $p=0.001$, $d=0.84$; T3 : $p<0.001$,	
	Germany	anxiety, paranoid	d=0.93; T4: p =0.001, d =0.86), depression (T2: p <0.001,	
		ideation, psychoticism,	d=1.02; T3: p =0.004, d =0.77; T4: p =0.002, d =0.80), anxiety	
		global wellness index,	(T2: p=0.010, d=0.74; T3: p=0.006, d=0.63; T4: p=0.001,	
		positive symptom	d=0.89), hostility (T2 : p =0.004, d =0.98; T3 : p =0.005,	
		distress index, and	d=0.79; T4: p =0.001, d =0.89), phobic anxiety (T2: p =0.014,	
		positive symptom total)	d=0.43; T3: p =0.038, d =0.43; T4: p =0.139, d =0.33),	
			paranoid ideation (T2 : $p=0.004$, $d=0.76$; T3 : $p=0.002$,	
			d=0.66; T4: p =0.001, d =0.73), psychoticism (T2: p =0.001,	
			d=0.88; T3: p =0.001, d =0.87; T4: p =0.001, d =0.83), global	
			wellness index (T2 : $p=0.001$, $d=1.00$; T3 : $p<0.001$, $d=0.97$;	
			T4: $p=0.001$, $d=0.96$), positive symptom distress index (T2:	
			<i>p</i> <0.001, <i>d</i> =1.14; T3: <i>p</i> =0.001, <i>d</i> =0.73; T4: <i>p</i> <0.001,	
			d=0.95), and positive symptom total (T2 : $p=0.004$, $d=0.71$;	
			T3: $p=0.001$, $d=0.81$; T4: $p=0.001$, $d=0.84$)	
Rieger &	General	Fatigue, depression,	Improvement in perceived relationship with mother	Yes
Stückema	population	anger, vigor (Profile of	(autonomy: p <0.001, η =.64; attachment: p =0.003, η =.52), no	
nn, 1999	(n=48),	Mood States), perceived	change in perceived relationship with father (autonomy:	
	Germany	quality of family	p=.335, $d=0.21$; attachment: $p=0.101$, $d=0.29$), mixed results	

Study	Sample	Outcome variables	Results	Treat-	
		(assessment tools)		ment	
				benefit	
		relationships in terms of	regarding perceived relationship with spouse (autonomy:		
		attachment and	$p=0.043, \eta=.49$; attachment: $p=0.416, \eta=.52$), mixed results		
		autonomy (Subjektives	regarding perceived relationship with first child (autonomy:		
		Familienbild	$p=.008$, $\eta=.78$; attachment: $p>0.999$, $d<0.01$); no change in		
		Fragebogen)	perceived relationship with 2^{nd} child (autonomy: $p=.455$,		
			d=0.14; attachment: $p=0.881$, $d=0.02$). Improvement in		
			relation to fatigue (p =0.004, d =0.91), depression (p =0.002,		
			d=0.99), vigor ($p=0.008$; $d=0.55$); no change in anger		
			(<i>p</i> =0.224, <i>r</i> = 0.18)		
Schumacher,	General	Perceived quality of	Improvement in perceived family relationships in terms of	Yes	
2000	population	family relationships in	autonomy (T2: $p=0.006$, $d=.28$; T3: $p<0.001$, $d=.51$) and		
	(n=53),	terms of attachment and	attachment (T2: p =0.001, d =.44; T3: p =0.020, d =.32)		
	Germany	autonomy (Subjektives			
		Familienbild			
		Fragebogen)			
Sethi, 200	General	Perceived quality of	Improvement on both target areas between pre- and post-	No	
	population	family relationships and	intervention; however, formal statistical analyses of these		
	(n=30),	wellbeing; both assessed	changes were not conducted and not enough descriptive data		
	Australia		were provided to allow the calculation of effect size		

Study	Sample	Outcome variables	Results	Treat-	
		(assessment tools)		ment	
				benefit	
		through ad hoc			
		questions			
Weinhold et	General	Psychological	Larger improvement in the intervention group than in the	Yes	
al., 2013†	population	functioning (Outcome	control group on all three target areas: psychological		
	(n=208),	Questionnaire 45.2),	functioning (T2: $p=0.003$, $d=0.45$; T3: $p=0.003$, $d=0.46$),		
	Germany	distress (Questionnaire	distress (T2: p <0.001, d =0.51; T3: p =0.001, d =0.51),		
		for the Evaluation of	motivational incongruence (T3: p <0.001, d =0.55; T3:		
		Treatment Progress),	p<0.001, d =0.52)		
		motivational			
		incongruence			
		(Incongruence			
		Questionnaire)			

[†] Marked studies are based on the same/partially overlapping sample

Values in bold are calculated by the review authors based on raw data reported in the original article or provided by the authors of the original studies (Dr. Krüger and Dr. Langlotz).

Table 4. Results of the meta-analysis investigating the effectiveness of family constellation therapy on overall psychopathology

Group by	Study name	Time poin	t	Statistics for each study					
Design		Hedges's g		Standard error	Variance	Lower limit	Upper limit	Z-Value	p-Value
Controlled	Krüger, 2003 (PANSS)	Blank	0.531	0.445	0.198	-0.341	1.404	1.194	0.233
Controlled	Weinhold, 2013 (OQ-45.2)	Combined	0.491	0.140	0.020	0.216	0.766	3.502	0.000
Controlled			0.495	0.134	0.018	0.233	0.757	3.699	0.000
Non-controlled	Höppner, 2006 (SCL90-R / C	9 33)ank	0.451	0.114	0.013	0.227	0.675	3.946	0.000
Non-controlled	Langlotz, 2005 (PAI)	Blank	0.602	0.167	0.028	0.275	0.929	3.605	0.000
Non-controlled	Langlotz, 2006 (SCL90-R / G	SC)ombined	0.893	0.254	0.064	0.395	1.390	3.519	0.000
Non-controlled			0.547	0.088	0.008	0.374	0.720	6.187	0.000
Overall			0.531	0.074	0.005	0.387	0.676	7.202	0.000

PANSS: Positive and Negative Syndrome Scale; PAI: Personality Assessment Inventory; SCL90-R / GSI: Global Severity Index of the Symptom Checklist-90-R; OQ-45.2: Outcome Questionnaire 45.2

Figure 1. Overview (flow chart) of the study selection process

