

RESEARCH ARTICLE

Integrative systematic review of psychodrama psychotherapy research: Trends and methodological implications

Hod Orkibi *, Rinat Feniger-Schaal

School of Creative Arts Therapies, Faculty of Social Welfare and Health Sciences, University of Haifa, Haifa, Israel

* horkibi@univ.haifa.ac.il



Abstract

Background

Psychodrama is an experiential psychotherapy in which guided role-play is used to gain insights and work on personal and interpersonal problems and possible solutions. Despite the wealth of literature describing clinical work, psychodrama intervention research is relatively scarce compared to other psychotherapies and psychological interventions.

Objective

For this reason we implemented the integrative approach to systematic review that authorizes the combination of publications with diverse methodologies and all types of participants, interventions, comparisons, and outcomes. Our aim was to produce a comprehensive summary of psychodrama intervention research in the last decade that critically evaluates methodological issues to inform future studies.

Methods

We searched four major electronic databases (PsycINFO, PubMed, Scopus by Elsevier, and Web of Science) for peer-reviewed articles on psychodrama interventions published in English between 1 January 2007 and 31 December 2017. The quality of qualitative and mixed methods studies was assessed on the basis of pre-established guidelines, and the risk of bias was assessed for all quantitative randomized control studies, consistent with the PRISMA protocol.

Findings

The database search and a hand search resulted in 31 psychodrama intervention publications. Overall, these studies examined the effects of psychodrama on more than 20 different outcomes and most studies had adult clients. The next largest group was adolescents, whereas only two studies involved children. Thus psychodrama intervention research in the last decade suggests there are promising results in all methodologies, and highlights the need to enhance methodological as well as reporting quality and to theorize and examine

OPEN ACCESS

Citation: Orkibi H, Feniger-Schaal R (2019) Integrative systematic review of psychodrama psychotherapy research: Trends and methodological implications. PLoS ONE 14(2): e0212575. <https://doi.org/10.1371/journal.pone.0212575>

Editor: Cheng-Shi Shiu, University of California Los Angeles, UNITED STATES

Received: October 11, 2018

Accepted: February 5, 2019

Published: February 19, 2019

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

Data Availability Statement: All relevant data are within the manuscript and its Supporting Information files.

Funding: The authors received no specific funding for this work.

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

modality-specific mechanisms that lead to therapeutic change. Recommendations to improve methodology, transparency, and specificity in reporting future psychodrama and other psychotherapy research are discussed.

Introduction

Psychodrama (PD), originated in 1921 by J. L. Moreno, is an experiential psychotherapy in which clients use guided role-play to work on their personal and interpersonal problems and possible solutions through actions rather than words alone [1, 2]. PD offers clients a “fail-safe” reality where feelings, thoughts, and behaviors can be explored and insights can be gained into past issues, present challenges, and future possibilities. The specificity of PD is that clients are encouraged to express their feelings directly and in the first person, to talk *to* rather than *about* their significant others (who are normally played by a group member or represented by an empty chair), which facilitates high levels of in-session experiencing. Although Moreno developed PD based on his group work and interactions with children using storytelling and role-play [1], today PD is used with clients of all ages, with individuals, couples, and families [3].

Despite the wealth of literature describing clinical work with clients of all ages and different problems, PD intervention research is relatively scarce compared to research on other psychotherapies. The focus in the PD literature has mostly been on describing and explaining processes through anecdotal experiences, clinical vignettes, and case illustration reports. Some of the key obstacles to conducting research are the predominantly clinical orientation of most PD practitioners and the related fact that almost all training programs take place in private rather than research institutions and are thus more experiential and clinical than research-oriented.

In the PD literature, three major systematic reviews of research have been published. In Kellermann’s review of 23 outcome studies (articles published between 1952 and 1985) he concluded that PD “is a valid alternative to other therapeutic approaches, primarily in promoting behavior change with adjustment, antisocial, and related disorders” [4]. Kipper and Ritchie [5] conducted the first meta-analytic study that focused on the effectiveness of using specific PD techniques in 25 experimentally designed studies (articles published between 1965 and 1999). The analysis revealed “an overall effect size that points to a large size improvement effect similar to or better than that commonly reported for group psychotherapy in general [of 0.50, and] the techniques of role reversal and doubling emerged as the most effective interventions” [5]. Wieser published the latest descriptive review of 52 studies (articles published between 1948 and 2006) on the treatment effects of PD, and concluded that “there is still a need for basic research into the effectiveness of psychodrama therapy” [6].

Types of review methods

There are several kinds of research review methods. The most basic is the simple *literature review* that only summarizes the literature on a given topic in narrative form and does not necessarily follow rigorous and explicit strategies for search, selection, or analysis of the literature [7]. A more in-depth research review method is the *systematic review* that implements an explicit method for conducting a comprehensive search, and for selecting and evaluating previous quantitative studies on a well-defined clinical question (e.g., does PD reduce youth depression?). Systematic reviews include narrative analysis of the primary studies and, sometimes, descriptive statistics and quantification of qualitative information. They provide reliable evidence on gaps in current research, solid grounds for determining evidence-based practice,

and the first stage of meta-analysis studies [8]. While systematic reviews have traditionally focused predominantly on questions relating to the effectiveness of interventions, there are at least 10 different types of systematic reviews that address diverse questions and the “information needs of healthcare professionals and policy makers” [9]. One type, which is particularly pertinent to this study, is a methodological systematic review that examines methodological issues relating, for example, to design and conduct. A *meta-analysis* is considered the highest level of evidence and the gold standard of literature-based research, where carefully selected data from primary quantitative studies are combined with rigorous statistical procedures to calculate an overall effect size [10].

Nevertheless, given the relatively small number of PD studies, we implemented an *integrative review* approach that allows for the combination of publications with diverse methodologies including quantitative, qualitative, and mixed methods [11]. This review approach can produce a more comprehensive portrayal of research trends in a given field. As such, integrative reviews can help generate new knowledge, as well as inform research, practice, and policy initiatives [12].

The present study

Drawing on a recent typology of systematic reviews in the medical and health sciences, the present study is an integrative methodological systematic review [9]. Hence the overarching question in this review was to determine which trends and methodological issues could be identified in psychodrama psychotherapy research. Accordingly, this systematic review had two aims: (a) to summarize 10 years of PD intervention research with all types of participants, interventions, comparisons, and outcomes according to the eligibility criteria described below; and (b) to critically evaluate methodological issues to help determine the requirements for future studies.

The present study is part of a wider project reviewing both PD and drama therapy research [13]. But because these two treatment modalities are considered different healthcare professions, two distinct reports were generated. Briefly, in psychodrama, clients typically use role-play to enact themselves, parts of themselves, or significant others in their real lives, and hence work more directly on reality-based issues. In contrast, drama therapy is more fantasy-based and clients typically use role-play to enact fictional and symbolic roles, use storytelling, puppetry, masks, and miniature objects to work more indirectly and with greater dramatic distance from their issues. However, some contemporary variants of psychodrama practice involve working with metaphors and imagination, thus blurring the boundaries between classical PD and drama therapy [14–16].

Methods

We systematically searched for studies using drama as a therapeutic modality, both PD and drama therapy, in four major electronic databases: PsycINFO, PubMed, Scopus by Elsevier, and Web of Science. Search areas were title, abstract, and keywords. The search terms for all the databases were: *Psychodrama* OR *dramatherapy* OR *drama therapy* AND *intervention* OR *program* OR *effect** OR *evaluat** OR “*case study*.” In addition, we conducted a hand search in the following three journals because they are specific to the treatment modalities: *Drama Therapy Review*, *Dramatherapy* (published by the British Association of Dramatherapists), and *The Journal of Psychodrama Sociometry and Group Psychotherapy* (published by the American Society of Group Psychotherapy and Psychodrama). We also searched a special issue on Psychodrama published in the *International Journal of Psychotherapy*. Note that the Australian and Aotearoa New Zealand Psychodrama Association Journal was excluded because the

journal's Guidelines for Contributors do not indicate a formal peer review process. In addition, the *Zeitschrift für Psychodrama und Soziometrie* (Journal for Psychodrama and Sociometry) was excluded because most of its papers are in German. The *British Journal of Psychodrama and Sociodrama* was also excluded because it is only available in print and is not indexed.

Data extraction

A spreadsheet for data extraction was piloted and then used to record information of interest and the PICOS framework specifying participants, interventions, comparisons, outcomes, and study designs [17]. Information was extracted from each study on: (1) study method: quantitative, qualitative, mixed methods; (2) study design (e.g., randomized / nonrandomized group comparison pre-post, single group pre-post, single subject design, and follow up); (3) characteristics of participants: *N*, age, gender, diagnosis, setting, country; (4) intervention type (duration, frequency, structure, protocol) versus control or placebo; (5) primary outcome measures; (6) results (significance, effect size). The authors of five studies were contacted with questions to verify information about their study [18–23].

Eligibility criteria

Eligibility criteria were for the publication to be an intervention research with quantitative and/or qualitative data, written in English, and published in a peer reviewed journal. No restrictions were imposed on publication status. Publication date was restricted to between January 2007 and December 2017 (the date last searched was 10 January 2018). Eligibility assessment was performed independently by the two authors and cases of disagreement were discussed until all issues were resolved and there was consensus.

Assessment of qualitative and mixed methods studies

The assessment of the qualitative and mixed methods studies drew on Hawker and colleagues' tool for critically and systematically reviewing research conducted using different paradigms [24]. Consistent with the purpose of this review, we focused on the seven methodological dimensions defined by Hawker et al.: method and data, sampling, data analysis, ethics and bias, findings, and transferability. Three dimensions that do not address methodological issues are not reported in this review (i.e., "abstract and title," "introduction and aims," "implications and usefulness").

Risk of bias assessment

Consistent with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) [25], we used the Risk of Bias (RoB) assessment tool that was developed by the Cochrane Collaboration for randomized controlled trials [26]. Generally, RoB assessment focuses on the internal validity of a study, and in particular on methodological flaws that can lead to bias. Because the results of a study may in fact be unbiased despite a methodological flaw, the tool assesses the *risk* of bias. Bias is assessed in several domains for which there is empirical evidence that biasing can influence the estimates of an intervention's effectiveness. A bias can lead to underestimation or overestimation of the true intervention effect [25].

Recently, it has been noted that because the RoB tool was developed within the field of medicine, its application to the context of psychotherapy outcome research requires several adjustments [27]. Drawing on the adjusted criteria suggested in Munder and Barth, we assessed the following six potential sources of bias: (1) random sequence generation, (2) concealment of allocation to conditions, (3) blinding of participants and personnel to condition assignment,

(4) handling of incomplete outcome data, (5) selective outcome reporting, and (6) treatment implementation. For each source of bias, two raters judged whether the risk of bias was *low*, *unclear*, or *high* according to definitions and examples in Munder and Barth (2018). Note that because all the primary outcomes in the studies were self-reported and because in psychotherapy it is impossible to blind patients to intervention aims and content we did not code RoB given the absence of blinding of the outcome assessors [28]. Hence, the RoB tool was only used for the assessment of studies with a randomized group comparison design.

Results

This article follows the reporting guidelines of the PRISMA framework where possible [25]. Fig 1 presents a PRISMA flow diagram indicating the studies retrieved for the review, screening, and assessment of eligibility (S1 Table).

We identified 162 articles through Scopus and 113 through PsycINFO. The Web of Science and PubMed searches yielded 66 and 25 articles, respectively. Additionally, 16 articles were found through hand search and a reference list search (total initial $N = 382$). Within these results, we searched for duplications and excluded 129 articles, yielding 253 articles after duplicates had been removed. Next, 201 articles were excluded based on full-text examination of publication type and the language of main text. Specifically, we excluded articles that were review articles, book reviews, media reviews, editorials, obituaries, and case studies/ examples/ illustrations/ vignettes. Because we focused on intervention research, articles that only described techniques or treatment programs without testing their effect on participants were also excluded. Similarly, articles concerning therapists, caregivers, teachers, or students in training rather than therapy clients were also excluded. We excluded articles where the word “drama” was used vaguely without specific indication of PD or drama therapy, and articles with eclectic treatments where PD or drama therapy were not indicated as the primary approaches. We also excluded several articles that were automatically detected because their Abstract was in English but the main text was in a different language (Persian, Croatian, Turkish, Polish, Portuguese, German, and French). One study was excluded because it could not be obtained. In total, 55 articles met the eligibility criteria, of which 31 were PD articles (60%) and 24 were drama therapy articles. This review focuses on the 31 PD articles presented in Tables 1 to 4 (the drama therapy articles are reviewed in-masked for the review, 2018).

General characteristics of psychodrama studies

As can be seen in Fig 2, the majority of the PD research publications were from Turkey ($n = 12$, 39%), with the next largest groups from Italy ($n = 4$, 13%) and Israel ($n = 3$, 10%), and USA ($n = 3$, 10%).

The frequency of PD publications per year ranged from one (in 2008) to eight (26% in 2017), with a total average of three publications per year. Of the 31 PD publications, 20 were quantitative, four were qualitative, and seven had a mixed methods design.

Regarding client populations and primary outcomes, as can be seen in Fig 3, the majority of the PD research publications dealt with youth at risk ($n = 10$, 32%), followed by students ($n = 4$, 13%), and adults with different mental health conditions ($n = 4$, 13%).

The most frequent outcomes, in descending order, were adolescents behavioral problems ($n = 6$, 19%), followed by anxiety ($n = 5$, 16%) and depression ($n = 4$, 13%). Next was quality of life ($n = 3$, 10%) followed by symptoms and global functioning ($n = 3$, 10%). Two studies measured attachment (6%) and two studies measured in-session process variables (6%), of which one measured dramatic engagement [29]. Only one study measured spontaneity, which is a core construct in PD theory [30]. As also seen in Fig 3, other primary outcomes included

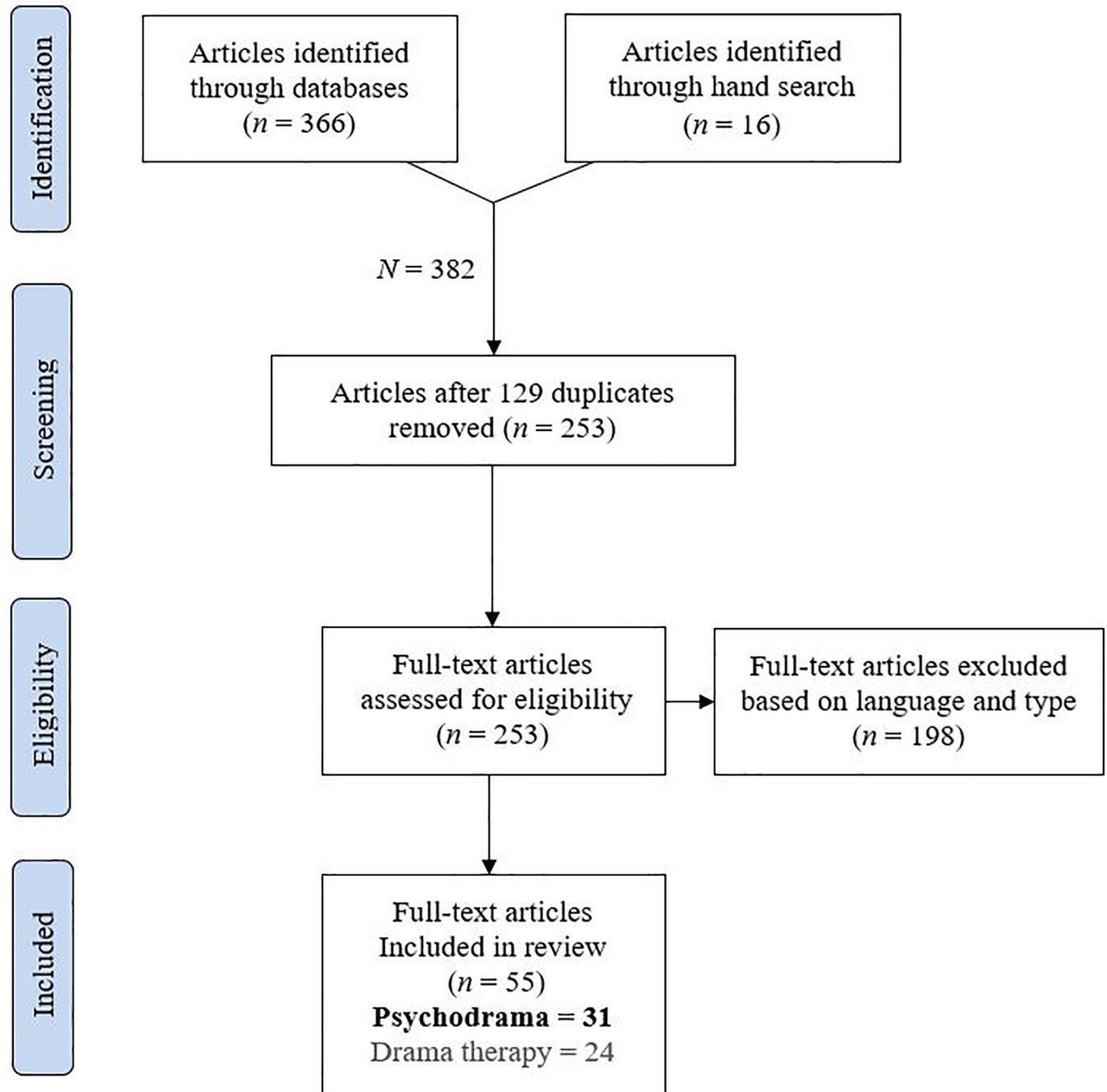


Fig 1. Systematic search according to the PRISMA statement methodology.

<https://doi.org/10.1371/journal.pone.0212575.g001>

subjective well-being, smoking cessation, burnout, hopelessness, loneliness, sense of coherence, public stigma, self-stigma, self-esteem, self-efficacy, and self-concept. The next sections present narrative reviews of the characteristics of these studies as a function of methodology.

Qualitative studies

All four qualitative studies had a single group design. As seen in [Table 1](#), the number of participants in the qualitative studies ranged from 7 to 13 ($M = 10$, $SD = 2.6$). All the studies involved

Table 1. Characteristics of the qualitative studies.

First Author (year) country	N	Clients	Setting	Intervention (protocol)	Data collection and analysis	Primary Findings
Alby et al. (2017) Italy	10	Adults w. cancer & family members ($M_{age} = 57$)	University (recruited from a hospital)	Group analytic PD, 1.5h semimonthly, 2 years (analytic three phases)	Interviews & thematic analysis	Increased self-awareness, expresses and shares feelings, reduces fears, aids decision making
Konopik (2013) USA	13	Adults w. MHC ($M_{age} = 40$)	Hospital	Group PD, 2h weekly, 4–6 sessions (structured five stages)	Observational, interviews & communication content analysis	Changes in emotions, family of origin issues, self-awareness and self-worth, shift in views
Menichetti (2016) Italy	8	Adults w. cancer ($M_{age} = 59$)	Hospital	Group PD, 1.5h weekly, 1–80 sessions (three PD phases)	Interviews & Interpretative phenomenological analysis	Self-expression, ability to give and receive help, sense of agency, cope with grieving
Vural (2014) Turkey	7	Mothers of children with ADHD ($M_{age} = NR$)	Special education center	Group PD, 2h weekly, 12 sessions (semi structured)	Observational & analysis of participants in-session statements	Instilled hope, self-confidence, better coping, reduced anger and punishment

PD = psychodrama. MHC = mental health conditions. ADHD = Attention deficit hyperactivity disorder. NR = not reported.

<https://doi.org/10.1371/journal.pone.0212575.t001>

adults: two studies examined Italian adults with cancer [20, 31], one study dealt with adults with mental health conditions in the USA [32], and one study was about Turkish mothers of children with Attention Deficit Hyperactivity Disorder (ADHD) [33]. Two of the studies dealt with hospital settings [31, 32], one study was conducted in a special education center [33], and one at a university [20].

PD treatment was delivered in a group format in all four studies: one study delivered analytic PD group therapy [20] and the three others followed the “classical” PD approach. In three studies, session duration ranged from 1.5 to 3 hours per week, and the number of sessions ranged from 1 to 80. The exception is the study with the analytic PD group that delivered two 1.5 hours sessions per month, for two years [20]. In terms of session structure, two studies had a structured [32] or a semi-structured format [33], one followed the standard three PD phases of warm-up, action, and sharing [31], and one followed an analytic 3-phase structure that included free associations as the group warm-up phase [20]. Information concerning the qualifications of the therapists was only provided in two studies: “a drama therapist educated in a national psychodrama institution and an assistant therapist” [33] and a “certified psychodrama therapist” [32]. Only one study indicated that the PD therapist received clinical supervision during treatment implementation [33].

Regarding data collection, two studies conducted post-treatment interviews [20, 31], and one study included both post-treatment interviews and direct observation during the sessions [32]. In one study, data were only collected by notes taken during the sessions [33]. Data analysis involved content/ thematic analysis [20, 32], interpretative phenomenological analysis [31], and analysis and quantification of participants’ in-session statements [33]. Detailed and clear descriptions of the data analysis procedures were presented in two studies [31, 32], whereas in one study the description was unclear [33] and in another study minimal details were provided about the data analysis [20].

In most studies, some context and setting were described. The recruitment procedure was reported in all studies, but information on sample demographics was relatively sparse across studies, with one study lacking even the a minimal necessary information such as participating mothers’ age and marital status [33]. Informed consent from participants was only reported in two studies [31, 33] and the issue of protecting participants’ confidentiality and anonymity

Table 2. Characteristics of mixed methods studies.

First author (year) country	N	Clients	Setting	Intervention (protocol)	Design	Primary outcomes	Qualitative findings
Chae (2017) South Korea	17	College students w. difficulties (M _{age} = NR)	University	1 Group PD, 2h, 10 sessions in 2 weeks (three PD phases)	Single group, pre-post Qual: post treatment feedback and interviews	Reduced insecure attachment: avoidance & anxiety	More positive view of self, others, and relationships
Dogan (2010) Turkey	21	Masters' students w. interpersonal difficulties (M _{age} = 25)	University	2 Groups: PD vs. placebo on effective learning, 2h weekly, 12 weeks (three PD phases)	Randomized, pre-post Qual: observation during sessions and last session feedback	No group differences in insecure attachment: avoidance & anxiety	More self-understanding and confidence, insight, hope, relationships
Gatta (2010) Italy	6	Adolescents w. psychiatric disorders (M _{age} = 17)	Semi-residential public service	2 Groups: PD vs. matched coping skills TAU, 1.25h weekly, 12 sessions (semi-structured)	Nonrandomized group comparison, pre-post Qual: case studies	Group difference in some symptoms	Participants' feedback on the treatment
Harkins (2011) UK	76	Adult inmates (M _{age} = 35)	Prisons	1 Group: PD + CBT, 2–3 days (protocol)	Single group, pre-post & observations Qual: post treatment interviews	Improved self-efficacy, motivation to change, confidence in skills	More confidence and better prepared to cope with the future after release
Karabilgin (2012) Turkey	7	Adults w. HIV/AIDS (M _{age} = 33)	University center	1 Group PD, 10 sessions of 6h (three PD phases)	Single group, pre-post Qual: pre and post treatment focus groups with descriptive analysis	Improved mental health only, out of eight quality of life dimensions	More acceptance and confidence to talk about HIV/AIDS feelings, cope with fear. Less depression
McVea (2011) Australia	17	Adults w. unresolved emotions (aged 27–66)	Private practice	1 Group, 2.5 days workshop (three PD phases)	Single group, pre-post, follow-up Qual: post workshop feedback, BSR, CCI	Gain in in-session resolution, reduction in interpersonal distress, but not in symptoms	Improvements in interpersonal functioning and sense of self
Terzioğlu (2017) Turkey	30	Women w. infertility (M _{age} = 32)	University hospital	2 Groups: PD vs. control NR, 3h weekly, 8 sessions (three PD phases)	Nonrandomized group comparison, pre-post Qual: observation, interview-unclear	Improved self-esteem, less depression, hopelessness. Result reported for anxiety are unclear	Increased awareness, competence, self-worth

PD = psychodrama. Qual = qualitative data collection and analysis. NR = not reported. TAU = treatment as usual. BSR = brief structured recall method. CCI = Client change interview protocol.

<https://doi.org/10.1371/journal.pone.0212575.t002>

was mentioned in one study alone [32]. Only one study mentioned the potential influence of the therapists or researchers on the findings, “since the patients who were interviewed were selected in consultation with group facilitators” [31]. There were no other indications of the researchers’ reflexivity regarding their potential influence on the findings in terms of their own bias, perspective, role, and interactions with participants.

Finally, in all studies sufficient data were presented to support the findings, ranging from relatively long quotations [31] to short quotations [20, 32], and paraphrases of the participants’ statements [33]. A procedure to enhance the credibility of the findings was only reported in one study [32], where the themes were “confirmed in consultation with a psychodrama expert.” It is not clear, however, if this expert was blind to the study aims.

Mixed methods studies

As shown in Table 2, of the seven mixed methods studies, four had a single group design [23, 34–36]. Two studies had nonrandomized group comparison designs: one with a treatment-as-

Table 3. Characteristics of quantitative single group studies.

First author (year) Country	N	Clients	Setting	Intervention (protocol)	Design	Primary Outcomes
Akinsola (2013) Nigeria	25	Children with social anxiety (aged 7–16)*	School	Group PD, three days (three PD phases)	Single group pre-post	Decreased social and performance anxieties
Bilge (2017) Turkey	28	University students with anger problems ($M_{age} = 21$)	University	Group psychoeducation & PD, 2h bimonthly, 3 sessions (3-stage protocol)	Single group pre-post	Decreased anger
Biolcati (2017) Italy	30	University students with MHC ($M_{age} = 22$)	University center	Group analytical PD, 1.5h weekly 40 sessions (unclear)	Single group pre-post	Improved well-being, decreased symptoms and risk
Högberg (2008) Sweden	14	Suicidal children and adolescents ($M_{age} = 15$)	Outpatient clinic	Individual DP** 1.5-2h bimonthly M sessions = 17 (semi-structured)	Single group pre-post + follow-up	Improved global functioning pretest-posttest & pretest-follow-up
Orkibi (2017) ¹ Israel	16	Adolescents at-risk ($M_{age} = 15$)	School	Group PD, 1.5h weekly 16–20 sessions (three PD phases)	Single case design change process	Increased in-session dramatic engagement predicted productive behaviors
Orkibi (2014) Israel	12	Adults with MHC and students (aged 22–60)*	University center	Group PD & theater 2h weekly, 20 sessions (three PD phases)	Single case design	Pretest-posttest & pretest-follow-up decreased public stigma and self-stigma, increase in self-esteem

*Mean age not reported.

** Some sessions included parents. PD = psychodrama. MHC = mental health conditions. NR = not reported. 1 = Orkibi, Azoulay, Regev, et al. (2017).

<https://doi.org/10.1371/journal.pone.0212575.t003>

usual control group matched for age, gender and diagnosis to minimize selection bias [37], and one with an unmatched and unspecified control group [38]. One mixed methods study had a randomized group comparison design and a 4-week control group on “effective learning and efficient study methods” which was shorter than the 12-week PD group [39].

The number of participants in the mixed methods studies ranged from six to 76 ($M = 25$, $SD = 24$). Most of the studies involved adult participants. One study involved adult inmates in the UK who were approaching their release date [36], one study involved adults in Turkey with HIV or AIDS [23], one study involved adults in Australia with unresolved painful emotional experiences [34], and one study involved women in Turkey with a diagnosis of infertility [38]. One study involved college students in South Korea struggling with life adjustment and interpersonal relationships [35], one involved Master’s students in Turkey with interpersonal difficulties [39], and one study involved adolescents in Italy with psychiatric disorders [37]. The setting of two studies was universities [35, 39], one study was conducted at a university care center [23], one was conducted at a fertilization unit of a university hospital [38], one in prisons [36], one in private practice [34], and one was conducted in a semi-residential public service setting [37].

In all seven studies PD treatment was delivered in a group format, of which one study used a semi-structured analytic PD group [37] and one used a protocol for a group PD combined with cognitive-behavioral techniques [36]. All the other studies generally followed the classical

Table 4. Characteristics of quantitative non-randomized and randomized group comparison studies.

First author (year) Country	N	Clients	Setting	Intervention (protocol)	Design	Primary Outcomes
Non-Randomized Group Comparison Studies						
Oguzhanoglu (2014) Turkey	28	Women with major depression ($M_{age} = 35$)	University polyclinic	3 Groups: PD & med, vs. med only vs. control, 3h weekly, 16 sessions (loose structure)	NRGC pre-post	Similar decrease in depression in PD & med group and in med only group
Orkibi (2017) ¹ Israel	40	Adolescents at-risk ($M_{age} = 15$)	School	2 Groups: PD vs. waiting list control, 1.5h weekly, 16–20 sessions (three PD phases)	NRGC pre-post & process-outcome analysis	Only in PD increases in 3 out of 4 self-concepts, decrease in loneliness. Some process-outcome correlations
Randomized Group Comparison Studies						
Aytemur (2012) Turkey	113	Adult smokers ($M_{age} = 47$)	University clinic	2 Groups: PD & ST vs. ST only 2h weekly, 8 sessions (semi structured)	RGC 3-wave measures	PD increased the success rate of smoking cessation in the early period
Dehnavi (2016) Iran	30	Men with opiate dependence ($M_{age} = 31-29$)	Addiction treatment clinic	2 Groups: PD vs. control 2h, 12 sessions in 6 weeks (three PD phases)	RGC pre-post	Increase in general quality of life
Kähönen (2012) Finland	77	Adults working in public service with burnout ($M_{age} 47.5$)	Private occupational healthcare services	3 Groups: PD vs. analytic vs. control, 6h, 17 days (unclear)	RGP * pre-post + 6-month follow-up	PD group showed a higher increase in sense of coherence than the analytic group during the intervention
Karataş (2011) Turkey	36	High school students manifesting aggression	School	3 Groups: PD vs. placebo of interaction vs. control, 1.5-2h weekly, 10 sessions (three PD phases)	RGC pre-post + 12-week follow-up	PD group higher increase in problem solving and decrease in aggressions, but follow-up only for aggression
Karataş (2014) Turkey	45	University students low on SWB and high on hopelessness ($M_{age} = NR$)	University	3 Groups: PD vs. control vs. placebo of reading, 1.5-2h weekly, 12 sessions (three PD phases)	RGC pre-post + 10-week follow-up	PD group increased SWB and decreased hopelessness compared to control and placebo. Follow-up effect lasted for hopelessness but not SWB
Karataş (2009a) Turkey	23	High school students manifesting aggression (9th grade)	School	2 Groups: PD vs. control 1.5-2h weekly, 14 sessions (three PD phases)	RGC pre-post + 16-week follow-up	PD decreased total aggression, anger, hostility, but not physical and verbal aggression
Karataş (2009b) Turkey	36	High school students with aggression (9th grade)	School	3 Groups: PD vs. CBT vs. control, 1.5-2h weekly, 10–14 sessions (three PD phases)	RGC pre-post + 16-week follow-up	PD and CBT decreased aggression compared to control, but CBT more than PD
Özbaş (2016) Turkey	82	Oncology nurses with burnout (aged 28–37)	Hospital inpatient oncological clinics	2 Groups: PD vs. control 2h weekly, 10 sessions (semi structured)	RGC pre-post + 3-month follow-up	PD group decrease burnout and increase psychological and workplace empowerment
Smokowski (2009a,b) ² USA	81	Latino immigrant adolescents** ($M_{age} = 14$)	Community	2 Groups: PD vs. support group, 3h weekly, 8 sessions (protocol)	RGP pre-post + 1-year follow-up	(2009a) Pre-post: group differences were not significant (2009b) Pre-follow-up: PD group maintained superior effects than support group
Sproesser (2010) Brazil	16	Adults with Parkinson's disease ($M_{age} = 58-60$)	Hospital outpatient clinic	2 Groups: PD vs. waiting list control 1.5h bimonthly, 12 sessions (unclear)	RGC pre-post	PD group had stronger improvement in depression, anxiety, and quality of life

(Continued)

Table 4. (Continued)

First author (year) Country	N	Clients	Setting	Intervention (protocol)	Design	Primary Outcomes
Tarashoeva (2017) Bulgaria	40	Adults with panic disorders (M_{age} 42)	Private mental health center	2 Groups: PD & med vs. & med only control 3h weekly, 25 sessions (unclear)	RGC pre-post + 6-month follow-up	PD group greater reduction in anxiety symptoms, improved spontaneity, quality of life and social functioning

PD = psychodrama. Med = medication. NRG = Non-Randomized Group Comparison. RGC = Randomized Group Comparison. ST = standard treatment.

SWB = subjective well-being. NR = not reported.

* Randomization was only for treatment groups.

**Some sessions included parents.

1 = Orkibi, Azoulay, Snir, et al. (2017). 2 = the results of the same study were reported in two separate articles.

<https://doi.org/10.1371/journal.pone.0212575.t004>

PD approach with a standard structure of three PD phases. In three studies, session duration ranged from 1.25 to 3 hours per week, and the number of sessions ranged from 8 to 12 [37–39]. One study delivered six weekly hours for 10 weeks [23]. Two studies were delivered in a concentrated format of a 2- or 3-day intervention [36] and a 2.5-day workshop [34]. Two of the seven mixed methods studies did not provide information concerning the qualifications of the therapists [35, 36]. Only one study indicated that adherence to PD treatment was monitored for treatment fidelity [34] and another study indicated that the PD therapist received clinical supervision during treatment implementation [23].

Qualitative data were collected using post-treatment individual interviews and/or feedback forms in three studies [34–36]. Two studies included direct observation during the sessions in addition to interviews [38, 39], one had focus group interviews both before and after treatment [23], and one presented case study descriptions of clients [37]. Detailed and clear descriptions of qualitative data analysis procedures were absent in all studies, and only one study specified the procedure used for qualitative data analysis as “descriptive analysis” [23]. Quantitative data were collected in all seven studies before and after treatment (pretest-posttest), and one study also collected data at 2-week and 3-month follow-ups [34]. Regarding quantitative data analysis, because of the small sample sizes and/or non-normally distributed variables, four studies

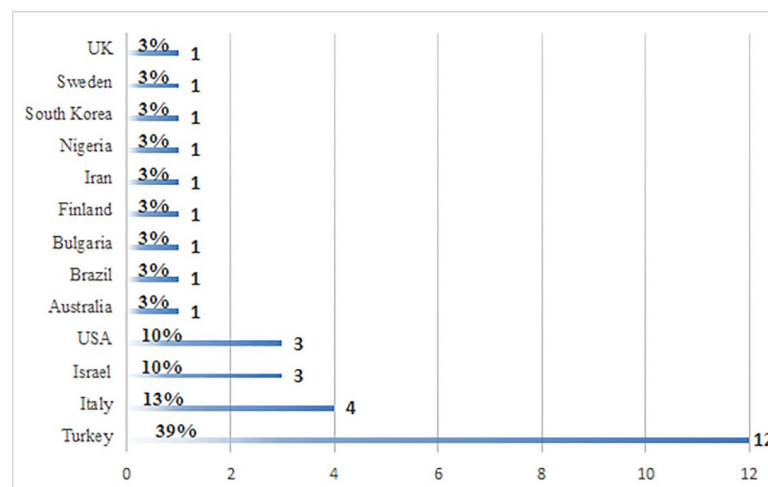


Fig 2. Frequency of included publications by country.

<https://doi.org/10.1371/journal.pone.0212575.g002>

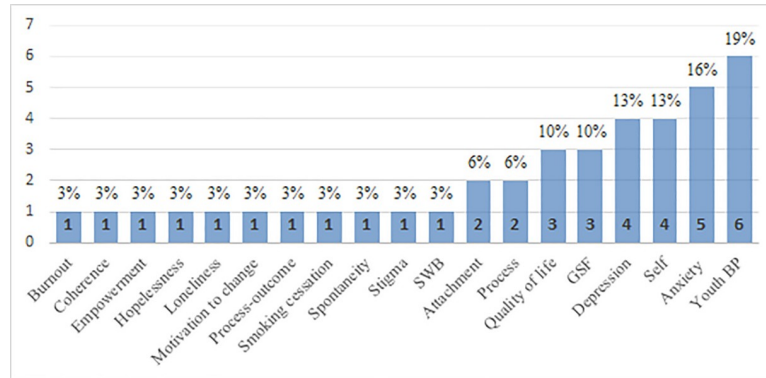


Fig 3. Primary outcomes. Frequency of primary outcomes. SWB = subjective wellbeing, GSF = general symptoms and functioning, BP = behavioral problems.

<https://doi.org/10.1371/journal.pone.0212575.g003>

computed non-parametric tests: a Mann-Whitney U test for independent samples and/or a Wilcoxon matched-pairs signed-rank test for paired samples [23, 37–39]. Two studies computed *t*-tests [35, 36], of which one also computed a repeated measures analysis of variance [36], and one study computed linear mixed models analysis but without a detailed description of the analysis [34]. Only two studies reported effect sizes. The study with adult inmates reported the highest (medium) effects for improved self-efficacy ($d = 0.67$) and social and friendship skills ($d = 0.54$) [36]. The study on adults with HIV or AIDS reported a large effect size on mental health improvement ($d = -.82$) [23].

In all studies, some context and setting were described but more details could have enhanced the findings’ transferability. The recruitment procedure was reported in all studies and information on the sample demographics varied across studies but was satisfactory. Obtaining informed consent from participants was indicated in five studies, with the exception of two [39] of which one only stated that participation was voluntary [35]. The issue of protecting participants’ confidentiality and anonymity was mentioned in one study alone [34] but two other studies referred to the group as a “confidential environment” [23] or to the “confidentiality of the sessions” [35].

Three studies mentioned the potential influence of the therapists or the researchers on the qualitative findings; e.g., through their interactions with clients/ interviewees [23, 34, 36], but only one implemented a procedure: “To reduce potential bias arising from dual roles, protagonists who had been directed by the first author were interviewed by other team members” [34]. In all studies sufficient qualitative data were presented to support the findings, ranging from relatively long quotations [34–36], to short quotations [23, 37, 38], or a combination of both [39]. Procedures to enhance the credibility of the qualitative findings were not reported in any of the studies. Finally, while only two studies were defined as “mixed methods” [38, 39], the collection of both qualitative and quantitative data was appropriate for the purpose of all the studies. In all studies, the integration of qualitative and quantitative data was presented in the Discussion section to a varying degree. No study included an integrated mixed method question or hypothesis.

Quantitative single group studies

As seen in Table 3, the number of participants in the six single group studies ranged from 12 to 30 ($M = 21$, $SD = 7.76$). Three studies involved children and/or adolescents [29, 40, 41] and two studies involved university students: one on anger management problems in Turkey [42] and one on mental health conditions in Italy [43]. One study consisted of both adults with

mental health conditions and students in Israel [44]. The setting of two studies was schools [29, 41], three studies were conducted at universities [42–44], and one study was conducted at an outpatient clinic [40].

In five studies, PD treatment was delivered in a group format, of which one study used a 3-stage protocol for psychoeducation sessions based on PD and some cognitive-behavioral techniques [42], in one group analytical PD was used with a protocol that was not reported [43], and the other four group PD studies followed the standard three PD phase structure. One study was delivered in an individual format, but some sessions were also attended by the adolescent's or child's parents [40]. In three studies, session duration ranged from 1.25 to 2 hours per week, and the number of sessions ranged from 16 to 40 [29, 43, 44]. In two studies PD was delivered twice a month [40, 42] and in one study in a 3-day format with unspecified hours [41]. Information concerning the qualification of the therapists was provided in only three of the six studies [29, 43, 44]. One study indicated that the PD director was “one of the researchers” [41] and another study indicated that participants were treated “by the author” [40]. Only one study described a procedure for ensuring implementation fidelity [44] and one study indicated that the PD therapist received clinical supervision during treatment implementation [29].

Data were collected in four studies before and after treatment [40–43], and one of these studies also collected data at a 22-month follow-up [40]. Two studies had a single case design (also termed “single subject design”) with repeated measures over the course of the treatment, of which one also had three post-treatment measures [44] and one measured in-session changes [29].

Data analysis included a non-parametric Mann-Whitney U test for independent samples and/or a Wilcoxon matched-pairs signed-rank test for paired samples [40, 42]. Analysis of variance and paired *t*-tests were computed in two studies [41, 43]. One study with a single case design computed a hierarchical linear model that took the nested data structure into account and compared aggregated measures at three time points [44]. Another study with a single case design computed a hierarchical linear model and odds ratio, which is an effect size statistic reflecting the odds of clients' dramatic engagement in predicting their productive in-session behaviors, with effects ranging from small (0.70) to medium (4.30) [29]. No other studies reported effect sizes.

The recruitment procedure was reported in all the studies. Information on sample demographics varied across studies but was generally satisfactory, with the exception of two studies that only reported age and gender [40, 41]. In three of the six studies it was indicated that participants provided informed consent [40, 41, 43].

Non-randomized group comparison

The number of participants in the two non-randomized group comparison studies ranged from 23 to 40 ($M = 32$, $SD = 12$), as seen in Table 4.

One study involved women coping with the first episode of major depression was conducted at a university polyclinic in Turkey [19]. This study included three groups and compared the effects of group PD with anti-depressant medication to the effects of a group with antidepressant medication alone and to a healthy control group of volunteers with similar ages and education levels. The second study involved adolescents at risk that was conducted at a middle school in Israel [45]. This study included two groups and compared the effects of group PD to a waiting list control group.

Regarding the treatment, the study at the university polyclinic had a loosely structured treatment with a session duration of three hours per week, for 16 sessions [19]. The study at

the school did not follow a treatment protocol, and session duration was 1.5 hours per week for 16 to 20 sessions [45]. In both studies, information concerning the qualification of the therapists was provided and both studies indicated that the therapist was given clinical supervision during treatment implementation.

Data were collected in both studies before and after treatment (i.e., pretest-posttest). In one study, due to small sample size and non-normally distributed variables, the data analysis included a non-parametric Kruskal Wallis test for the comparisons of the three groups and a Wilcoxon matched-pairs signed-rank test for paired samples [19]. The second study included a mixed-design repeated measures analysis of variance and process-outcome analyses based on correlations between the change scores for the process and outcome variables [45]. Only the latter study reported effect sizes, with effects ranging from medium (.11) to large (.28) partial eta-squared. In both studies, the recruitment procedure was reported but information on sample demographics was more detailed in one study [45] than the other [19]. In both studies participants provided informed consent.

Randomized group comparison

As seen in Table 4, the number of participants in the 11 randomized group comparison studies ranged from 16 to 113 ($M = 53$, $SD = 30$). Four studies were on adolescents, of which three were on adolescents with aggression problems in Turkey (reported as distinct studies [46–48]) and one on Latino immigrant adolescents in the US (the same study was reported in two separate articles [21, 22]). The remaining seven studies were on adult smokers in Turkey [49], men with opiate dependence in Iran [50], adults with burnout who work in public services in Finland [18], oncology nurses coping with burnout in Turkey [51], adults with Parkinson's disease in Brazil [52], and adults with panic disorders in Bulgaria [30]. One study was on university students with low subjective well-being and high hopelessness in Turkey [53].

The settings of three studies were schools [46–48], three studies were conducted at health service centers [18, 30, 50], two studies at hospital clinics [51, 52], and two at universities [49, 53]. One study was conducted in community settings [21, 22].

In all studies PD treatment was delivered in a group format, of which five followed the standard three PD phase structure [46–48, 50, 53]. Two treatments were semi-structured [49, 51], one treatment had a structured protocol [21, 22], whereas in three studies the content of the PD treatment was unclear [18, 30, 52].

In most studies ($n = 8$ of 11, 72%), session duration ranged from 1.5 to 3 hours per week, and the number of sessions ranged from eight to 25 [21, 22, 30, 46–49, 51, 53]. One study delivered a 2-hour group PD for 12 sessions over six weeks [50], another delivered 6-hour group PD sessions over 17 days [18], and in one study a 1.5-hour group PD was delivered twice a month for 12 sessions [52].

Therapists' qualifications were reported in eight studies, of which only six reported that the PD groups were led by psychodramatists [22, 30, 46, 49, 51, 53] and two studies reported that the PD groups were led by a psychologist and a physiotherapist [18], and a psychologist [52]. The remaining three studies did not report any information about the therapists' qualifications [47, 48, 50]. Only two studies indicated the PD therapist received clinical supervision during treatment implementation [22, 46].

Data collection in most (9 of 11, 81%) of the randomized group comparison studies included before and after treatment measures and a follow-up measure that ranged from 10 weeks to 1 year after the treatment, with the exception of two studies that only included pretest-post measures [50, 52]. Data analysis in most studies (7 of 11, 64%) used parametric tests such as analysis of variance and/or *t*-tests, [18, 30, 46, 47, 50–52], as well as regression analysis

[21, 22]. Three studies used non-parametric tests, including a Mann-Whitney U test for independent samples, a Wilcoxon matched-pairs signed-rank test for paired samples, and/or a Kruskal-Wallis test for independent samples, a chi-square test or a Fisher exact test for categorical data [48, 49, 53]. Three studies reported moderate to large effect sizes [18, 21, 22, 50], and one study reported an odds ratio indicating a small effect size [49].

The recruitment procedure was reported in all studies and information on sample demographics varied across studies but was generally satisfactory, with the exception of five studies that only reported age (or grade), gender, and/or scores on a screening instrument [18, 46–48, 53]. In two studies it was indicated that participants provided informed consent [51, 52] and two studies indicated that participants volunteered [46, 47, 53].

Risk of bias assessment

This section reports the risk of bias analysis results for 12 randomized group comparison studies (11 in Table 4, and the Dogan et al. 2010 study in Table 2). A summary of the risk of bias analysis is presented in Table 5 and the proportion of studies for each of the six potential sources of bias is presented in Fig 4.

As can be seen, none of the studies was identified as having a low risk of bias across all six sources of bias. Evidence for potential *selection biases* (both “random sequence generation” and “concealment of allocation to conditions”) were unclear in the majority of studies (75%, 9 of 12). One study was at low risk of these two selection biases because it had a restricted block

Table 5. Risk of bias assessment results per study.

First author (year)	Random sequence generation (selection bias)	Allocation concealment (selection bias)	Blinding (performance bias)	Incomplete outcome data (attrition bias)	Selective reporting (reporting bias)	Treatment implementation (performance bias)
Aytemur (2012)	Low	Low	Unclear	Unclear	Unclear	Low
Dehnavi (2016)	Unclear	Unclear	High	Low	Unclear	Unclear
Dogan (2010)	Unclear	Unclear	Unclear	High	Low	Low
Kähönen (2012)	High	High	High	High	High	Unclear
Karataş (2011)	Unclear	Unclear	Unclear	Low	High	Unclear
Karataş (2014)	Unclear	Unclear	Unclear	Low	High	Low
Karataş (2009a)	Unclear	Unclear	Unclear	Unclear	Unclear	Unclear
Karataş (2009b)	Unclear	Unclear	Unclear	Unclear	Unclear	Low
Özbaş (2016)	Low	Unclear	High	Unclear	High	Low
Smokowski (2009a,b)*	Unclear	Unclear	Unclear	High	Unclear	Low
Sproesser (2010)	Unclear	Unclear	High	Low	High	Unclear
Tarashoeva, (2017)	Unclear	High	Unclear	Low	Unclear	Unclear

Risk of bias assessment was performed only for the 11 studies with randomized group comparison, corresponding to Table 4 and Dogan (2010) in Table 2.

*The results of the same study were reported in two separate articles.

<https://doi.org/10.1371/journal.pone.0212575.t005>

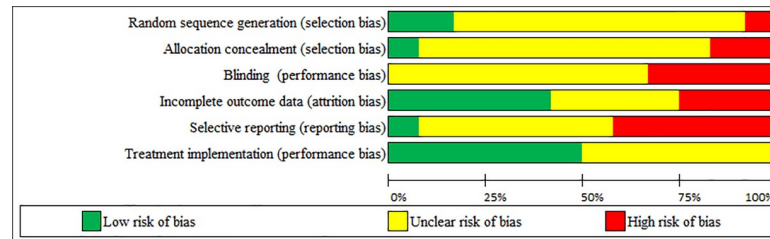


Fig 4. Proportion of studies for each of the six potential sources of bias.

<https://doi.org/10.1371/journal.pone.0212575.g004>

randomization method and used sealed envelopes “prepared before enrollment by the head nurse” [49], suggesting that participants and investigators enrolling participants were blind to the allocation to conditions. A high risk of these two sources of selection biases was evident in a study where clients were randomized “into the two different intervention groups but not into the control group” [18]. Thus, randomization was incomplete and clients or investigators could have possibly foreseen the allocation to conditions. Another study was at high risk for inadequate allocation concealment because it was explicitly defined as “an open” clinical study in which the allocation to conditions was not concealed [30].

The blinding of participants and personnel (i.e., clients and therapists) is intended to prevent the direct and indirect influence of their expectations or motivations on the study outcomes. However, this type of *performance bias* is impossible to avoid in psychotherapy: therapists know what treatment they are delivering, and clients not only are aware of receiving that treatment but also have expectations concerning the treatment [27]. For this reason, psychotherapy reviews have considered there is a low risk of bias on the basis of two adjusted criteria: (a) the use of an active control group with a credible treatment involving comparable treatment expectancies, and (b) the assessment of client-perceived “treatment expectancies or credibility prior to or early (i.e., after the first session) in treatment and in which the treatment group had equal (zero difference) or lower expectancies or credibility than the control group received” (see Appendix C, p. 4, in [28]). In the present review, none of the studies was identified as having a low risk of performance bias according to these criteria. Evidence for potential blinding performance bias was unclear in the majority of studies (67%, 8 of 12) that had an active control group but did not measure client-perceived treatment expectancies or credibility. Four studies (33%) were at high risk of performance bias because they had an inactive control group and no measures of treatment expectancies or credibility [18, 50–52].

Attrition bias refers to the mishandling of incomplete outcome data. There was a low risk of attrition bias in five studies (42%) with no reported dropouts, appropriate imputation of missing data, and/ or unlikeliness that the missing data were related to the outcome [30, 48, 50, 52, 53]. The risk of attrition bias was unclear in four studies (33%) due to insufficient reporting of attrition [46, 47, 49, 51]. There was a high risk of attrition bias in three studies that had a dropout rate exceeding 20% and no imputation of missing data [18, 21, 22, 39].

Selective outcome *reporting bias* refers to cases where outcomes that were collected according to the methods section (or to a pre-study published protocol or a clinical trial’s registry) are not reported. There was a low risk of reporting bias in only one study [39], and an unclear risk in five studies [21, 22, 30, 46, 47, 49, 50]. There was a high risk of reporting bias in five studies where not all of the pre-specified primary outcomes or not all of the results related to the outcome were reported [18, 48, 51–53].

Treatment implementation is a type of performance bias that refers to deviations from intended treatment that are likely to have affected the outcome and to the qualifications of the

therapist delivering the treatment [27]. There was a low risk of *treatment implementation bias* in six studies because both the psychodramatists' qualifications and the treatment content were reported [22, 39, 46, 49, 51, 53]. The remaining six studies were at unclear risk of treatment implementation bias because the qualifications of the therapist were not reported [47, 48, 50], the therapist was not identified as a psychodramatist [18, 52], and/or a psychodramatist delivered the therapy but no content was described [30]. None of the studies was identified as having a high risk of treatment implementation bias.

Discussion

This is the first integrative review to systematically examine methodological issues in 31 intervention studies on PD psychotherapy. Overall, the studies examined the effect of PD on more than 20 different outcomes and most studies had adult clients. The next largest group was adolescents, and only two studies involved children.

In reviewing PD intervention studies, it is important to consider its core theoretical concepts of spontaneity and creativity. Spontaneity refers to the pro-creative, catalyzing state of readiness by which creativity emerges, whereas creativity refers to the creative act itself; namely the creative, cognitive, emotional, behavioral response or a tangible creative product [54]. Moreno considered creativity essential to adapting to life changes and to coping with unexpected challenges [55]. This notion corresponds to contemporary views of creativity as an important factor for adaptation [56, 57]. Relatedly, Moreno associated mental health with the ability to enact a wide repertoire of roles that enable the individual to act flexibly and adequately, in the right way at the right time [58]. Thus, it is striking that only one study included in this review measured spontaneity [30] and that none measured creativity or other PD constructs. To substantiate PD as a distinct treatment modality, future intervention studies would do well to measure such modality-specific constructs that might account for treatment effects based on psycho-dramatic theoretical reasoning.

In light of these findings, in the next sections we recommend specific guidelines to improve the methodology, transparency, and specificity in reporting PD, or other intervention research. The proposed guidelines build on several resources in psychology [59, 60], psychotherapy [27], and music therapy [61].

Qualitative studies

The four qualitative studies presented sufficient excerpts from the data which showed that overall, PD can contribute to clients' self-worth and promote self-awareness, self-expression, and better perceived coping with difficulties. The descriptions of participants' characteristics, the contexts, and settings could have been more detailed to allow for comparison with other contexts and settings and to enhance the transferability of the findings. According to the American Psychological Association's new Journal Article Reporting Standards ("APA JARS") for Qualitative Research, "transferability of findings in qualitative research to other contexts is based on developing deep and contextualized understandings that can be applied by readers" [60]. The rationale for choosing a specific data analysis method varied across studies and, generally, descriptions could have been clearer and more detailed. In addition, the researchers did not provide a reflexive self-description of their views and backgrounds (e.g., demographic/cultural characteristics, values, experience with the phenomena, training, etc.), or whether and how reflexivity was used to mitigate their potential influences on the participants and/or the interpretation of the findings. It is worth noting that one study applied a procedure to enhance the credibility of the findings by confirming the emergent themes with a psychodrama expert [32]. Future studies could enhance the credibility of the findings by implementing procedures

such as data triangulation, disinterested peer debriefing, external auditing, or member checking (also called “informant feedback” or “respondent validation”).

Mixed methods studies

Mixing qualitative and quantitative methods strengthens the internal and external validity of the findings, in that the advantages of each method complement one another, and their shortcomings are considerably offset [62]. In the seven mixed methods studies, data integration was presented in the Discussion section, which was useful for interpreting the findings based on the two datasets as well as for giving a (qualitative) voice to participants and ensuring that the quantitative results reflected their experiences. An “integrated” mixed method question or hypothesis which directly addresses the mixing of the quantitative and qualitative strands of the study were not put forward in any of the studies. An example of an integrated question that conveys the *methods* of the study is “Do the qualitative data help explain the results from the initial quantitative phase of the study?” and an example that conveys the *content* of the study is “Do clients identify helpful aspects in therapy that shed light on why or how therapy was effective?” In addition, most studies did not contain a statement about the rationale for choosing a mixed methods design such as a more comprehensive and nuanced understanding of the research problems. Researchers should also be encouraged to state the type of mixed methods design used, and specify the timing of the data collection (concurrent or sequential) as well as the emphasis (equal or unequal) on each data type. Examples of mixed method designs include Convergent Parallel design, Exploratory Sequential design, and Explanatory Sequential design [63]. One mixed methods approach that has been adopted by some PD researchers in Europe [64] is the Hermeneutic Single-Case Efficacy Design which integrates quantitative outcome and weekly change data with clients’ qualitative accounts of change over therapy, while examining alternative non-therapy explanations for changes in therapy [65].

Quantitative studies

Single group. Quantitative studies with a single group pretest-posttest design are sometimes the only viable option in applied research that occurs in natural field settings. This design is quasi-experimental and some consider it uninterpretable given the multiple threats to internal validity associated with the lack of a control group [66]. Although this design allows the researcher to determine whether a change occurred between pretest and posttest, this change may be attributed to factors other than the treatment, such as history (i.e., an event that can impact the outcome) or maturation (i.e., clients’ natural growth or development). Thus, results of studies with a single group pretest-posttest design have less certainty.

When a control group is unavailable, studies with a single case design can provide more certainty on whether the intervention is responsible for change because this design involves multiple measures over the course of the treatment, with each client serving as his or her own control for purposes of comparison [67]. Single case design is underused in PD research although it is widely employed in applied field studies when a control group is unavailable. Experts have stated that single case designs can “provide a rigorous experimental evaluation of intervention effects. . . [and] a strong basis for establishing causal inference” [68]. Established standards for single case designs require a minimum of five data points in a phase (and at least three phases) to meet evidence standards without reservations [68]. For example, five systematic measures before (i.e., at baseline), during, and after the treatment may demonstrate an effect when the data pattern in the treatment phase differs from the data pattern observed in the baseline phase.

Nonequivalent groups. The four studies (of which two were mixed methods) with a non-randomized group comparison design represent another type of quasi-experimental study that is considered more interpretable than the single group pretest-posttest design [66]. Studies with a non-randomized group comparison design typically include nonequivalent groups that can differ on various characteristics due to the lack of randomization. In particular, two of the studies included in this review had a control group of participants who were matched on several characteristics to control for pre-treatment differences that could introduce selection bias [19, 37]. In studies with a non-randomized group comparison design, matching or examination of pretest equivalence can increase confidence in attributing any observed posttest differences between groups to the treatment rather than to some pretest differences. In addition, researchers are encouraged to employ statistical procedures that account for pre-treatment differences such as using the pretest score as a covariate in analyzing the posttest scores [69].

Randomization. Of the studies included in this review, 48% ($n = 12$) had a randomized group comparison design. To reduce selection bias, researchers may choose to allocate participants randomly to conditions with a computerized random number generator or a table of random numbers, etc. rather than a systematic but non-random approach (e.g., birth or admission dates, record number). The allocation to conditions should be concealed so that study personnel who enroll participants remain blind to the condition assignment, for example by using sequentially numbered, opaque, sealed envelopes. This can prevent study personnel from excluding or changing allocations of particular clients that can undermine randomization and “might cause baseline differences in characteristics relevant to outcomes” [27].

Control groups. Five group comparison studies included an active control group (“active comparator”) that was used to examine the specific effect of PD treatment compared to other acknowledged psychotherapies or a placebo activity. Most group studies, however, compared PD to standard care or passive control with no treatment. Therefore, the specificity of the PD effects remained equivocal in most studies. As the current evidence suggests that PD has an effect, an important next step would be to examine whether this effect is due to specific “active ingredients” in PD or more general common factors that are shared across therapeutic modalities [70]. The use of a placebo or an active control group with a credible treatment can also minimize performance bias related to the impossibility of blinding therapists and clients in psychotherapy research by controlling for treatment expectancies that may influence the outcome [27].

Mediation and moderation. Mediation and moderation are being increasingly explored in psychotherapy research that aims to go beyond basic questions about effectiveness [71, 72], but is still rare in PD research. Hypothesizing mediators (i.e., a variable that accounts for the indirect association between an intervention and outcome) can shed light on “why” or “how” therapy leads to change, which ultimately can facilitate the optimization of change [73, 74]. To date, there are no evidence-based explanations of precisely why psychodrama works and how it leads to changes. Future studies could pinpoint modality-specific mediators that might account for treatment effects based on theoretical reasoning and a scientific rationale for why the specified PD component (e.g., role-reversal) might influence the outcome of interest (e.g., empathy). Similarly, hypothesizing moderators (i.e., a construct that influences the direction or magnitude of the relationship between the intervention and outcome) can lead to a better understanding of “when” or “for whom” therapy leads to change [75]. Moderators are often client or therapist characteristics (e.g., sex, ethnicity, temperament) or the treatment delivery format (e.g., individual vs. group treatment) [73, 74]. Mediators and moderators should be clearly articulated in the theoretical model upon which the treatment is based to establish “treatment differentiation” that clarifies how the PD treatment differs from the comparison or control group.

The PD treatment. Clear descriptions of treatment duration, frequency and settings varied across studies. It is particularly important to provide justifications when examining PD treatments that are unique in delivery format, such as when PD is delivered in an intensive format or with a considerable gap between sessions (i.e., weeks or months). Of the studies included here, 54% ($n = 16$) employed the three standard phases of a PD session (warm-up, action, and sharing). The remaining studies had semi-structured treatments ($n = 6$, 20%), a structured protocol ($n = 4$, 13%), or an unclear /unspecified structure ($n = 4$, 13%). Replication and transparency can be promoted by clearly reporting the objectives and the content of the sessions, appending the treatment protocol where applicable, and by specifying the essential active components and how they are incorporated into the PD treatment. In any case, it is important to define PD-specific terminology for international cross-disciplinary readership.

Implementation fidelity. Treatment implementation fidelity was monitored in three studies that used video recordings [22, 34, 44]. Implementation fidelity is a concern for rigorous research and includes two main components. One is *treatment integrity* that refers to (a) the extent to which a therapist delivers the intervention with adequate *adherence* to the manual and/or intended treatment modality and its theory-specified techniques or methods, and (b) the *competence*, or skillfulness, with which these techniques or methods are implemented. The second component of implementation fidelity is *treatment differentiation* that refers to the extent to which a treatment differs from a comparison or control condition [76]. Because clinical research is designed to compare the treatment modality rather than the therapist's ability, variations in therapists' competence and adherence must be minimized and treatment differentiation must be maximized. Thus, observational measures of video/audiotaped sessions or self-report implementation checklists can be used for the ongoing review of implementation fidelity [77].

Relatedly, because PD is not a manualized treatment, it is important to establish an operational definition of PD treatment during the research design phase. Then, during treatment, competent adherence to this definition should be monitored. For example, (a) enactment of at least one scene that approximated a real-life situation or was an externalization of the client's inner experience; (b) application of at least one of the four fundamental PD techniques (soliloquy, doubling, mirroring, role reversal, and concretization), and (c) the fact that group members played roles in the client's enactment [34, 78]. In applied research, as long as there is competent adherence to the essential active components of a treatment, the treatment can be somewhat flexible and adaptable according to clients' needs or external conditions. Overall, the need for monitoring implementation fidelity calls for the development of measures that focus on PD therapists' competence and adherence, similar to other therapies [79].

Therapist training. Another important component of implementation fidelity is therapist training, selection, and clinical supervision to monitor delivery quality. "Psychodrama" is a profession-specific term that should not be used to describe treatment delivered by healthcare professionals who do not have PD credentials. Because this issue influences the quality of delivery directly, the PD practitioner should be adequately trained to deliver the treatment and subjected to assessment and ongoing evaluation and supervision. Information on specific theoretical orientation and the therapists' pre-implementation training should also be provided. Seven (23%) of the studies here indicated that the therapists received clinical supervision during treatment implementation. Overall, more information on the qualification level of the therapist who applied the PD is needed, since only 58% of the studies ($n = 18$) reported that the treatment was delivered by a trained psychodramatist.

Handling missing data. Almost half of the studies had a low risk of attrition bias. Non-respondents and clients' early dropout from treatment leads to incomplete data. To minimize attrition bias due to incomplete data, researchers can choose from several methods for

handling missing data [80]. Researchers are encouraged to use both *intent-to-treat analysis* (ITT), in which all participants initially enrolled in the treatment are included in the group to which they were initially assigned by statistically imputing (replacing) missing data with substituted values, as well as *completer analysis* in which only clients who completed the treatment are included in the analysis and no data imputation is performed. It is important to note any difference in results for the effect of treatment “as assigned” (ITT) versus treatment “as treated” [81, 82].

Effect size and statistical power. Statistical significance measures how likely it is that differences in outcome between treatment and control groups are real and not due to chance. The most commonly used measures of statistical significance are $p < .05$ and 95% confidence intervals (CI) that “estimate the range within which the real results would fall if the trial is conducted many times” [83]. But while measures of statistical significance can inform the reader as to whether a difference or an effect exists, they cannot reveal the size of the effect. In other words, the effect size indicates not only whether a treatment affects clients, but how much it affects them [84]. Unlike a p value, effect size is independent of sample size. Common effect size indices are Cohen’s d for the comparison of group means (e.g., t -test) and the partial eta-squared for the analysis of variance [85]. Most of the quantitative studies (67%) reported the statistical significance of the findings but did not report effect size, which is important in determining the magnitude of the difference between groups.

Another procedure that is underused in PD research is the calculation of statistical power, which is defined as the probability of avoiding a Type II error; i.e., rejecting the null hypothesis of no effect when in fact there is an effect. A-priori calculation of statistical power can help researchers determine the optimal sample size for testing their hypotheses during the design phase of their study. A reasonable minimum level of power to aim for is .8 [86].

A related concept to consider is the *clinical significance* of the findings (also termed clinical or practical importance) which refers to “the extent to which therapy moves someone outside the range of the dysfunctional population or within the range of the functional population” [87]. Thus, clinical significance refers to the practical importance of treatment in terms of actually reflecting a clinically meaningful change. In some cases, a large sample size may yield a statistically significant difference between groups that has little clinical significance (i.e., clinically meaningful change) and in other cases a non-statistically significant result may fail to detect an important and meaningful difference between groups; e.g., due to an underpowered study with a small sample. Researchers are encouraged to employ statistical methods of analysis that estimate clinical significance in psychotherapy research [88, 89]. Finally, it is imperative to report all of the study’s pre-specified (primary and secondary) outcomes, including non-significant results, to avoid selective reporting bias. A rationale should be provided if researchers only report the total score of a scale that includes distinct subscales scores mentioned in the Introduction or Method sections.

Conclusions

This integrative systematic review sought to produce a comprehensive summary of the state of PD intervention research in the last decade as well as to highlight and critically evaluate methodological issues that can inform future quantitative, qualitative, and mixed methods studies. The limitations of this review include the exclusion of non-English studies and articles published in non-indexed journals. Overall, the review suggests that PD intervention research in the last decade has followed an upward trajectory, and reports promising results across methodologies. This review constitutes a call to improve methodological and reporting quality. Advances are also needed in theorizing and examining PD modality-specific mechanisms of

action or change—which distinguish PD from traditional talk psychotherapy—and their impact on desirable outcomes. The current findings should not be viewed as discouraging because although PD is a long-standing experiential psychotherapy that originated in the beginning of the 20th century, it is an emerging sector in the psychotherapy research arena. As PD research becomes more focused and methodologically rigorous, meta-analytic reviews will provide empirical indices of documented effects. Meanwhile, we encourage authors, journal editors, and reviewers to consider our recommendations for PD intervention research. Further improvement in the methodology, transparency, and specificity of reporting PD intervention research is important not only for scientific purposes but also for the professional status of PD and quality of care.

Supporting information

S1 Table. PRISMA checklist.

(PDF)

Author Contributions

Conceptualization: Hod Orkibi, Rinat Feniger-Schaal.

Formal analysis: Hod Orkibi.

Methodology: Hod Orkibi, Rinat Feniger-Schaal.

Validation: Hod Orkibi, Rinat Feniger-Schaal.

Visualization: Hod Orkibi.

Writing – original draft: Hod Orkibi.

Writing – review & editing: Hod Orkibi, Rinat Feniger-Schaal.

References

1. Moreno JL. Psychodrama—first volume. Beacon, NY: Beacon House; 1946.
2. Moreno JL. Who shall survive? Foundations of sociometry, group psychotherapy and sociodrama. 3 ed. Beacon, NY: Beacon House; 1978.
3. Holmes P, Karp M, Watson M. Psychodrama since Moreno: Innovations in theory and practice: Routledge; 2005.
4. Kellermann PF. Outcome research in classical psychodrama. *Small Group Research*. 1987; 18(4):459–69. <https://doi.org/10.1177/104649648701800402>
5. Kipper DA, Ritchie TD. The effectiveness of psychodramatic techniques: A meta-analysis. *Group Dynamics: Theory, Research, and Practice*. 2003; 7(1):13–25. <https://doi.org/10.1037/1089-2699.7.1.13>
6. Wieser M. Studies on treatment effects of psychodrama therapy. In: Baim C, Burmeister J, Maciel M, editors. *Psychodrama: Advances in theory and practice*. London: Routledge; 2007. p. 271–92.
7. Galvan JL. *Writing literature reviews: A guide for students of the social and behavioral sciences*. 3rd ed. Glendale, CA: Pyrczak; 2006.
8. Zhao K, Bai ZG, Bo A, Chi I. A systematic review and meta-analysis of music therapy for the older adults with depression. *International Journal of Geriatric Psychiatry*. 2016; 31(11):1188–98. <https://doi.org/10.1002/gps.4494> PMID: 27094452
9. Munn Z, Stern C, Aromataris E, Lockwood C, Jordan Z. What kind of systematic review should I conduct? A proposed typology and guidance for systematic reviewers in the medical and health sciences. *BMC Medical Research Methodology*. 2018; 18(1):5. <https://doi.org/10.1186/s12874-017-0468-4> PMID: 29316881
10. Lee JH. The effects of music on pain: A meta-analysis. *Journal of Music Therapy*. 2016; 53(4):430–77. <https://doi.org/10.1093/jmt/thw012> PMID: 27760797

11. Whitemore R, Knafk K. The integrative review: Updated methodology. *Journal of Advanced Nursing*. 2005; 52(5):546–53. <https://doi.org/10.1111/j.1365-2648.2005.03621.x> PMID: 16268861
12. Hopia H, Latvala E, Liimatainen L. Reviewing the methodology of an integrative review. *Scandinavian Journal of Caring Sciences*. 2016; 30(4):662–9. <https://doi.org/10.1111/scs.12327> PMID: 27074869
13. Feniger-Schaal R, Orkibi H. Integrative systematic review of drama therapy intervention research. Submitted for publication. 2019.
14. Orkibi H. The user-friendliness of drama: Implications for drama therapy and psychodrama admission and training. *The Arts in Psychotherapy*. 2018; 59:101–8. <https://doi.org/10.1016/j.aip.2018.04.004>.
15. Azoulay B, Orkibi H. The four-phase CBN Psychodrama model: A manualized approach for practice and research. *The Arts in Psychotherapy*. 2015; 42:10–8. <https://doi.org/10.1016/j.aip.2014.12.012>
16. Hug E. Current trends in psychodrama: Eclectic and analytic dimensions. *The Arts in Psychotherapy*. 1997; 24(1):31–5. [https://doi.org/10.1016/S0197-4556\(96\)00046-9](https://doi.org/10.1016/S0197-4556(96)00046-9).
17. Methley AM, Campbell S, Chew-Graham C, McNally R, Cheraghi-Sohi S. PICO, PICOS and SPIDER: a comparison study of specificity and sensitivity in three search tools for qualitative systematic reviews. *BMC Health Services Research*. 2014; 14:579. <https://doi.org/10.1186/s12913-014-0579-0> PMC4310146. PMID: 25413154
18. Kähönen K, Nääätänen P, Tolvanen A, Salmela-Aro K. Development of sense of coherence during two group interventions. *Scandinavian Journal of Psychology*. 2012; 53(6):523–7. <https://doi.org/10.1111/sjop.12020> PMID: 23170866
19. Oguzhanoglu NK, Sözeri-Varma G, Karadag F, Tümkaya S, Muharrem E, Kiroglu Y. Prefrontal cortex neurochemical metabolite levels in major depression and the effects of treatment: An 1HMRS study. *Türk Psikiyatri Dergisi*. 2014; 25(2):75. PMID: 24936754
20. Alby F, Angelici G, Picinotti S, Zuccheromaglio C. A pilot study on an analytic psychodrama group for cancer patients and family members. *Rassegna di Psicologia* [Internet]. 2017; 34(1):[67–77 pp.]. Available from: <http://rassegnadipsicologia.it/index.php/rdp/article/view/66>.
21. Smokowski PR, Bacallao M. Entre dos mundos/between two worlds: Youth violence prevention for acculturating latino families. *Research on Social Work Practice*. 2009; 19(2):165–78. <https://doi.org/10.1177/1049731508315989>
22. Smokowski PR, Bacallao M. Entre dos mundos/between two worlds youth violence prevention: Comparing psychodramatic and support group delivery formats. *Small Group Research*. 2009; 40(1):3–27. <https://doi.org/10.1177/1046496408326771>
23. Karabilgin ÖS, Gökengin GB, Doğaner İ, Gökengin D. The effect of psychodrama on people living with HIV/AIDS. *European Journal of Psychotherapy & Counselling*. 2012; 14(4):317–33. <https://doi.org/10.1080/13642537.2012.734529>
24. Hawker S, Payne S, Kerr C, Hardey M, Powell J. Appraising the evidence: Reviewing disparate data systematically. *Qualitative Health Research*. 2002; 12(9):1284–99. <https://doi.org/10.1177/1049732302238251> PMID: 12448672.
25. Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gøtzsche PC, Ioannidis JP, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: Explanation and elaboration. *PLoS medicine*. 2009; 6(7):e1000100. <https://doi.org/10.1371/journal.pmed.1000100> PMID: 19621070
26. Higgins JPT, Altman DG, Gøtzsche PC, Jüni P, Moher D, Oxman AD, et al. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ*. 2011; 343. <https://doi.org/10.1136/bmj.d5928> PMID: 22008217
27. Munder T, Barth J. Cochrane's risk of bias tool in the context of psychotherapy outcome research. *Psychotherapy Research*. 2018; 28(3):347–55. <https://doi.org/10.1080/10503307.2017.1411628> PMID: 29224503
28. Laird KT, Tanner-Smith EE, Russell AC, Hollon SD, Walker LS. Comparative efficacy of psychological therapies for improving mental health and daily functioning in irritable bowel syndrome: A systematic review and meta-analysis. *Clinical Psychology Review*. 2017; 51:142–52. <https://doi.org/10.1016/j.cpr.2016.11.001>. PMID: 27870997
29. Orkibi H, Azoulay B, Regev D, Snir S. Adolescents' dramatic engagement predicts their in-session productive behaviors: A psychodrama change process study. *The Arts in Psychotherapy*. 2017; 55:46–53. <http://dx.doi.org/10.1016/j.aip.2017.04.001>.
30. Tarashoeva G, Marinova-Djambazova P, Kojuharov H. Effectiveness of psychodrama therapy in patients with panic disorders: Final results. *International Journal of Psychotherapy*. 2017; 21(2):55–66.
31. Menichetti J, Giusti L, Fossati I, Vegni E. Adjustment to cancer: Exploring patients' experiences of participating in a psychodramatic group intervention. *European Journal of Cancer Care*. 2016; 25(5):903–15. <https://doi.org/10.1111/ecc.12412> PMID: 26515989

32. Konopik DA, Cheung M. Psychodrama as a social work modality. *Social Work*. 2013; 58(1):9–20. <https://doi.org/10.1093/sw/sws054> PMID: 23409336
33. Vural P, Akkaya C, Küçükparlak I, Ercan I, Eracar N. Psychodramatic group psychotherapy as a parental intervention in attention deficit hyperactivity disorder: A preliminary study. *The Arts in Psychotherapy*. 2014; 41(3):233–9. <https://doi.org/10.1016/j.aip.2014.02.004>.
34. McVea CS, Gow K, Lowe R. Corrective interpersonal experience in psychodrama group therapy: A comprehensive process analysis of significant therapeutic events. *Psychotherapy Research*. 2011; 21(4):416–29. <https://doi.org/10.1080/10503307.2011.577823> PMID: 21623548
35. Chae SE, Kim SJ. Group psychodrama for Korean college students. *Journal of College Student Psychotherapy*. 2017; 31(1):59–70. <https://doi.org/10.1080/87568225.2016.1239512>
36. Harkins L, Pritchard C, Haskayne D, Watson A, Beech AR. Evaluation of geese theatre's re-connect program: Addressing resettlement issues in prison. *International Journal of Offender Therapy and Comparative Criminology*. 2011; 55(4):546–66. <https://doi.org/10.1177/0306624X10370452> PMID: 20472705
37. Gatta M, Lara DZ, Lara DC, Andrea S, Paolo TC, Giovanni C, et al. Analytical psychodrama with adolescents suffering from psycho-behavioral disorder: Short-term effects on psychiatric symptoms. *The Arts in Psychotherapy*. 2010; 37(3):240–7. <https://doi.org/10.1016/j.aip.2010.04.010>.
38. Terzioğlu C, Özkan B. Psychodrama and the emotional state of women dealing with infertility. *Sexuality and Disability*. 2017. <https://doi.org/10.1007/s11195-017-9514-8>
39. Dogan T. The effects of psychodrama on young adults' attachment styles. *The Arts in Psychotherapy*. 2010; 37(2):112–9. <https://doi.org/10.1016/j.aip.2010.02.001>.
40. Högberg G, Hällström T. Active multimodal psychotherapy in children and adolescents with suicidality: Description, evaluation and clinical profile. *Clinical Child Psychology and Psychiatry*. 2008; 13(3):435–48. <https://doi.org/10.1177/1359104507088348> PMID: 18783125
41. Akinsola EF, Udoka PA. Parental influence on social anxiety in children and adolescents: Its assessment and management using psychodrama. *Psychology*. 2013; 4(03):246–53. <http://dx.doi.org/10.4236/psych.2013.43A037>.
42. Bilge A, Keskin G. An evaluation of the effectiveness of anger management education enriched by psychodrama. *Journal of Psychiatric Nursing/Psikiyatri Hemsireleri Dernegi*. 2017; 8(2):59–65.
43. Biolcati R, Agostini F, Mancini G. Analytical psychodrama with college students suffering from mental health problems: Preliminary outcomes. *Research in Psychotherapy: Psychopathology, Process and Outcome*. 2017; 20(3):201–9. <https://doi.org/10.4081/ripppo.2017.272>
44. Orkibi H, Bar N, Eliakim I. The effect of drama-based group therapy on aspects of mental illness stigma. *The Arts in Psychotherapy*. 2014; 41(5):458–66. <http://dx.doi.org/10.1016/j.aip.2014.08.006>.
45. Orkibi H, Azoulay B, Snir S, Regev D. In-session behaviours and adolescents' self-concept and loneliness: A psychodrama process–outcome study. *Clinical Psychology and Psychotherapy*. 2017; 24: O1455–O63. <https://doi.org/10.1002/cpp.2103>. PMID: 28653318
46. Karatas Z, Gokcakan Z. The effect of group-based psychodrama therapy on decreasing the level of aggression in adolescents. *Turkish Journal of Psychiatry*. 2009; 20(4):1–9.
47. Karatas Z, Gokcakan Z. A comparative investigation of the effects of cognitive-behavioral group practices and psychodrama on adolescent aggression. *Educational sciences: Theory and practice*. 2009; 9(3):1441–52.
48. Karatas Z. Investigating the effects of group practice performed using psychodrama techniques on adolescents' conflict resolution skills. *Educational Sciences: Theory and Practice*. 2011; 11(2):609–14.
49. Aytemur ZA, Pı̇şmişođlu B, Kilinç O, Pı̇şmişođlu E, Hacıevliyaglı SS, Karaman C. Intensive clinic intervention plus psychodrama in smoking cessation and effects on cessation outcome. *Türkiye Klinikleri Journal of Medical Sciences*. 2012; 32(3):630–7. <https://doi.org/10.5336/medsci.2011-23965>
50. Dehnavi S, Bajelan M, Pardeh SJ, Khodaviren H, Dehnavi Z. The effectiveness of psychodrama in improving quality of life among opiate-dependent male patients. *International Journal Of Medical Research & Health Sciences*. 2016; 5(5):243–7.
51. Özbaş AA, Tel H. The effect of a psychological empowerment program based on psychodrama on empowerment perception and burnout levels in oncology nurses: Psychological empowerment in oncology nurses. *Palliative and Supportive Care*. 2016; 14(4):393–401. <https://doi.org/10.1017/S1478951515001121> PMID: 26466981
52. Sproesser E, Viana MA, Quagliato EMAB, de Souza EAP. The effect of psychotherapy in patients with PD: A controlled study. *Parkinsonism & Related Disorders*. 2010; 16(4):298–300. <https://doi.org/10.1016/j.parkreldis.2009.08.008>.
53. Karatas Z. Effects of psychodrama practice on university students' subjective well-being and hopelessness. *Eğitim ve Bilim*. 2014; 39(173):117–28.

54. Moreno JL. Theory of spontaneity-creativity. *Sociometry*. 1955; 18(4):105–18.
55. Moreno JL. The third psychiatric revolution and the scope of psychodrama. *Group psychotherapy*. 1964; 17(2–3):149–71.
56. Cohen LM. Adaptation, adaptiveness, and creativity. In: Runco MA, Pritzker SR, editors. *Encyclopedia of creativity*. 2nd ed. San Diego, CA: Academic Press; 2011. p. 9–23.
57. Orkibi H, Ram-Vlasov N. Linking trauma to posttraumatic growth and mental health through emotional and cognitive creativity. *Psychology of Aesthetics, Creativity, and the Arts*. 2018. Advance online publication. <https://doi.org/10.1037/aca0000193>
58. Fox J, editor. *The essential Moreno: Writings on psychodrama, group method, and spontaneity*. New York, NY: Springer publications.; 1987.
59. Appelbaum M, Cooper H, Kline RB, Mayo-Wilson E, Nezu AM, Rao SM. Journal article reporting standards for quantitative research in psychology: The APA Publications and Communications Board task force report. *American Psychologist*. 2018; 73(1):3–25. <http://dx.doi.org/10.1037/amp0000191>. PMID: 29345484
60. Levitt HM, Bamberg M, Creswell JW, Frost DM, Josselson R, Suárez-Orozco C. Journal article reporting standards for qualitative primary, qualitative meta-analytic, and mixed methods research in psychology: The APA Publications and Communications Board task force report. *American Psychologist*. 2018; 73(1):26–46. <http://dx.doi.org/10.1037/amp0000151>. PMID: 29345485
61. Robb SL, Carpenter JS, Burns DS. Reporting guidelines for music-based interventions. *Journal of Health Psychology*. 2011; 16(2):342–52. <https://doi.org/10.1177/1359105310374781> PMID: 20709884
62. Creswell JW, Plano Clark VL. *Designing and conducting mixed methods research*. 3rd ed. Los Angeles, CA: SAGE; 2018.
63. Creswell JW. *Research design: Qualitative, quantitative, and mixed methods approaches*. 4th ed. Thousand Oaks, CA: SAGE Publications; 2014.
64. Gonzalez A-J, Martins P, Lima MPd. Studying the Efficacy of Psychodrama With the Hermeneutic Single Case Efficacy Design: Results From a Longitudinal Study. *Frontiers in Psychology*. 2018; 9(1662). <https://doi.org/10.3389/fpsyg.2018.01662> PMID: 30250442
65. Elliott R. Hermeneutic single-case efficacy design. *Psychotherapy Research*. 2002; 12(1):1–21. <https://doi.org/10.1080/713869614> PMID: 22471329
66. Heppner PP, Wampold BE, Owen J, Wang KT, Thompson MN. *Research design in counseling*. 4th ed. Boston, MA: Cengage Learning; 2015.
67. Kazdin AE. *Single-case research designs: Methods for clinical and applied settings*. 2nd ed. New York: Oxford University Press; 2011.
68. Kratochwill TR, Hitchcock J, Horner R, Levin JR, Odom S, Rindskopf D, et al. Single-case designs technical documentation. What works clearinghouse [Internet]. 2010. Available from: <https://ies.ed.gov/ncee/wwc/Document/229>.
69. Kisbu-Sakarya Y, MacKinnon DP, Aiken LS. A Monte Carlo comparison study of the power of the analysis of covariance, simple difference, and residual change scores in testing two-wave data. *Educational and Psychological Measurement*. 2013; 73(1):47–62. <https://doi.org/10.1177/0013164412450574> PMID: 26412869
70. Wampold BE. How important are the common factors in psychotherapy? An update. *World Psychiatry*. 2015; 14(3):270–7. <https://doi.org/10.1002/wps.20238> PMID: 26407772
71. Lutz W, Knox S. *Quantitative and qualitative methods in psychotherapy research*. New York, NY: Routledge; 2014. <https://doi.org/10.1080/10503307.2014.989290>
72. Lambert MJ. *Bergin and Garfield's handbook of psychotherapy and behavior change*. 6th ed. Hoboken, N.J.: John Wiley 2013.
73. Kazdin AE. Understanding how and why psychotherapy leads to change. *Psychotherapy Research*. 2009; 19(4–5):418–28. <https://doi.org/10.1080/10503300802448899> PMID: 19034715
74. Kazdin AE. Moderators, mediators and mechanisms of change in psychotherapy. In: Lutz W, Knox S, editors. *Quantitative and qualitative methods in psychotherapy research. Explorations in mental health*. New York, NY: Routledge; 2014. p. 87–101.
75. Frazier PA, Tix AP, Barron KE. Testing moderator and mediator effects in counseling psychology research. *Journal of counseling psychology*. 2004; 51:115–34. <https://doi.org/10.1037/0022-0167.51.1.115>
76. Hildebrand MW, Host HH, Binder EF, Carpenter B, Freedland KE, Morrow-Howell N, et al. Measuring treatment fidelity in a rehabilitation intervention study. *American Journal of Physical Medicine & Rehabilitation*. 2012; 91(8):715–24. <https://doi.org/10.1097/PHM.0b013e31824ad462> PMID: 22377824

77. Ang K, Heggul N, Gao W, Higginson IJ. Strategies used in improving and assessing the level of reporting of implementation fidelity in randomised controlled trials of palliative care complex interventions: A systematic review. *Palliative Medicine*. 2018; 32(2):500–16. <https://doi.org/10.1177/0269216317717369> PMID: 28691583
78. Kellermann PF. *Focus on psychodrama: The therapeutic aspects of psychodrama*. Philadelphia: Jessica Kingsley Publishers; 2000.
79. Denton WH, Johnson SM, Burlison BR. Emotion focused therapy-therapist fidelity scale (EFT-TFS): Conceptual development and content validity. *Journal of Couple & Relationship Therapy*. 2009; 8(3):226–46. <https://doi.org/10.1080/15332690903048820> PMID: 20052311
80. Sv Buuren. *Flexible imputation of missing data*. Boca Raton, FL: CRC Press; 2012.
81. Ten Have TR, Normand S-LT, Marcus SM, Brown CH, Lavori P, Duan N. Intent-to-treat vs. Non-intent-to-treat analyses under treatment non-adherence in mental health randomized trials. *Psychiatric annals*. 2008; 38(12):772–83. <https://doi.org/10.3928/00485713-20081201-10> PMID: 20717484
82. Gupta SK. Intention-to-treat concept: A review. *Perspectives in Clinical Research*. 2011; 2(3):109–12. <https://doi.org/10.4103/2229-3485.83221> PMID: 21897887
83. Leung W. Balancing statistical and clinical significance in evaluating treatment effects. *Postgraduate medical journal*. 2001; 77(905):201–4. <https://doi.org/10.1136/pmj.77.905.201> PMID: 11222834
84. Sullivan GM, Feinn R. Using effect size—or why the p value is not enough. *Journal of Graduate Medical Education*. 2012; 4(3):279–82. <https://doi.org/10.4300/JGME-D-12-00156.1> PMID: 23997866
85. Ferguson CJ. An effect size primer: A guide for clinicians and researchers. *Professional Psychology: Research and Practice*. 2009; 40(5):532–8. <https://doi.org/10.1037/a0015808>
86. Cohen J. A power primer. *Psychological Bulletin*. 1992; 112(1):155–9. <https://doi.org/10.1037/0033-2909.112.1.155> PMID: 19565683
87. Jacobson NS, Truax P. Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology*. 1991; 59(1):12–9. <https://doi.org/10.1037/0022-006x.59.1.12> PMID: 2002127
88. Lambert MJ, Ogles BM. Using clinical significance in psychotherapy outcome research: The need for a common procedure and validity data. *Psychotherapy Research*. 2009; 19(4–5):493–501. <https://doi.org/10.1080/10503300902849483> PMID: 20183403
89. Ronk FR, Hooke GR, Page AC. Validity of clinically significant change classifications yielded by Jacobson-Truax and Hageman-Arrindell methods. *BMC Psychiatry*. 2016; 16(1):187. <https://doi.org/10.1186/s12888-016-0895-5> PMID: 27267986