

Franziska Kühne | Peter Eric Heinze | Florian Weck

Standardized patients in psychotherapy training and clinical supervision

study protocol for a randomized controlled trial

Suggested citation referring to the original publication:

Trials 21 (2020) 276

DOI <https://doi.org/10.1186/s13063-020-4172-z>

ISSN (online) 1745-6215

Postprint archived at the Institutional Repository of the Potsdam University in:

Postprints der Universität Potsdam

Humanwissenschaftliche Reihe ; 628

ISSN 1866-8364

<http://nbn-resolving.de/urn:nbn:de:kobv:517-opus4-469295>

DOI <https://doi.org/10.25932/publishup-46929>

STUDY PROTOCOL

Open Access



Standardized patients in psychotherapy training and clinical supervision: study protocol for a randomized controlled trial

Franziska Kühne* , Peter Eric Heinze and Florian Weck

Abstract

Background: Psychotherapy is highly effective and widely acknowledged for treating various mental disorders. Nevertheless, in terms of methods for teaching effective psychotherapeutic approaches and competencies, there has been a lack of investigation. Training and supervision are the main strategies for teaching therapist competencies, and standardized role-plays with simulated patients (i.e., trained individuals playing someone with a mental disorder) seem useful for evaluating training approaches. In medical education, this procedure is now internationally established. However, so far, little use has been made of standardized role-playing to evaluate training and supervision in the area of clinical psychology and psychotherapy.

Methods: In this study, standardized role-plays are used to evaluate methods for training and supervision. Central cognitive behavioral approaches for treating depression are taught in the training. The first experiment compares an active training approach (i.e., model learning) with a passive one (i.e., reading manual-based instructions). The second experiment compares a direct supervision technique (i.e., supervision based on video analysis) with an indirect one (i.e., supervision based on verbal reporting). In each experiment, 68 bachelor's and master's students of psychology will be randomly assigned to the experimental and control groups. Each student takes part in three role-plays (baseline, post and 3-month follow-up), which are all videotaped. Two independent raters assess therapist competencies in each role-play on the basis of a standardized competence scale.

Discussion: The research project aims to contribute to the development of specific training and supervision methods in order to improve psychotherapy training and patient care.

Trial registration: ISRCTN Registry, [ISRCTN19173895](https://www.isrctn.com/ISRCTN19173895). Registered on 10 December 2019.

Keywords: Clinical psychology, Education, Psychotherapeutic competencies, Psychotherapy research, Role-playing, Simulated patients, Standardized patients, Randomized controlled trial

Background

Training and supervision of psychotherapists are central to the dissemination of evidence-based psychotherapy [1]. Results on the efficacy of psychotherapy from randomized controlled trials (RCTs) may be integrated adequately into clinical practice only with effective

training. To date, psychotherapy has more often been part of research than psychotherapy training; thus, current training approaches cannot be characterized as evidence-based [2, 3].

Assessment of psychotherapeutic competencies

The appropriate assessment of psychotherapeutic competencies is central to the evaluation of psychotherapy training and clinical supervision methods. Waltz et al.

* Correspondence: dr.franziska.kuehne@uni-potsdam.de
Department of Psychology, Clinical Psychology and Psychotherapy,
University of Potsdam, Karl-Liebknecht-Str. 24-25, 14476 Potsdam, Germany



© The Author(s). 2020 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

[4] consider therapist competence as the level of skill in treatment delivery, covering relevant contextual variables (e.g., client variables and stage in therapy). For the measurement of competencies, Muse and McManus [5] refer to Miller's pyramid model [6] and define four levels. At the most basic level of knowledge, measurement may be realized by using multiple-choice questions (MCQs) or essays. For the second level of practical understanding, the former two as well as clinical vignettes and case reports may be used. Assessments of practical understanding may be accomplished via standardized role-plays, and clinical practice assessments via ratings of treatment sessions, supervisory assessments, or patient outcomes. Whereas higher levels of the model are associated more with external validity, lower levels provide more internal validity and opportunities for standardization. Despite their higher external validity, competence assessments within real patient encounters imply a number of difficulties. First of all, competence assessments are dependent on patient difficulty and behavior, and owing to their variability, several measurements are necessary for reliable assessments [7]. By contrast, measuring competence via simulated patients (i.e., at the third level of the pyramid model) reduces the aforementioned problems and facilitates standardization for learning purposes.

Use of simulated patients

In medical training, the use of standardized role-plays is now widespread [8, 9]. "Simulated patients" are trained persons (laypersons, patients, or actors) portraying a person with a diagnosed disease, whereas "standardized patients" are described as providing their portrayal in a consistent and constant manner [10, 11]. Often, the terms are used interchangeably and are abbreviated as "SP" [9]. In contrast to improvised interactions, SPs are provided with intensive role training in advance of their play, facilitating standardization for training and assessment purposes [11]. Furthermore, SPs may be used for the portrayal of rare clinical situations [12] and to foster experiential learning, modeling, skills training, or learning through feedback [13]. Thus, SPs are commonly used for diagnostics or communication training but also for medical examinations, so-called objective structured clinical examinations [8, 9, 14].

Simulated patients in clinical psychology

In contrast, within clinical psychology and psychotherapy, the advantages of SPs are not fully exploited and have been recognized only in recent years [15–19]. Although using SPs in clinical psychology is perceived as important for learning, self-efficacy, self-reflection, and therapeutic alliance [15–18], current surveys concentrate on subjective assessments, small samples, or convenience samples or are limited to one-group designs [15–18].

Thus, there is a need for RCTs using SPs in clinical psychology [8].

Active versus passive strategies

All in all, the evidence indicates that active training strategies (e.g., modeling, feedback, and practical exercises) are more effective than passive ones (e.g., reading manuals) [20]. A meta-analysis by Hill and Lent [21] revealed that training conversational techniques were significantly more effective than no training at all ($d = 0.89$) but that modeling was the most important intervention. But again, most studies were uncontrolled and did not use a training manual, and the meta-analysis did not refer to specific psychotherapeutic interventions (such as cognitive restructuring) but focused on counselling or exploration skills (e.g., showing empathy). Beyond training, competence-based supervision is described as fundamental to the development of psychotherapeutic competencies [22]. Referring to audio or video tapes enables the supervisee to receive specific competence-based feedback, which is considered relevant for skills development [23]. Comparing supervision based on verbal reports (i.e., supervision as usual) with those based on video recordings of therapy sessions, Martino et al. [24] demonstrated that the latter contributed more to competence development. Even so, regarding clinical supervision, only a limited number of systematic and controlled studies have been published to date [12, 25].

Objectives and hypotheses

Therefore, the aim of this study is to use SPs to evaluate specific training and supervision conditions. Psychology students will be taught distinct interventions relevant for the cognitive behavioral therapy of patients with diagnosed depression (i.e., behavioral activation and cognitive interventions based on an established treatment manual) [26]. The focus on depression was chosen because, on the one hand, it is one of the most prevalent mental disorders [27] and, on the other hand, psychotherapeutic competencies are highly relevant for depression treatment [28, 29].

The study combines the development of SP methods with two consecutive experiments. During the developmental phase, scenarios for the SP portrayals will be developed and validated, instruments for the assessment of therapist behavior and SP authenticity will be developed, SPs will be recruited and trained thoroughly, and their portrayal will be validated on the basis of video recordings.

In the first experiment, two specific training strategies (i.e., reading written instructions with or without modeling) will be compared. In the second experiment, two central supervision techniques (i.e., based on treatment videos versus on verbal reporting) will be evaluated. We expect post-training, psychology students randomly

assigned to the active strategy of modeling to display more therapeutic competencies than students who were only reading the written instructions on behavioral activation or cognitive techniques (hypothesis 1). Furthermore, we expect the competence differences between the two groups to persist until the 3-month follow-up (hypothesis 2). Second, we expect that post-supervision, psychology students randomly assigned to supervision based on treatment videos will display more therapeutic competencies than the students who described their treatment verbally to the supervisor (hypothesis 3). Again, we expect the competence differences between the two groups to persist at the 3-month follow-up (hypothesis 4).

Methods/Design

Study design

The study is a single-center RCT (see Fig. 1) in which participants (i.e., bachelor’s and master’s students included as either trainees or supervisees) will conduct therapy sessions with SPs. All sessions will be video-recorded. Following a preparation period, both experiments will be conducted consecutively within a 2-year period. Data will be gathered before and after training/

supervision and at the 3-month follow-up. Participants fulfilling the inclusion criteria (see below) will be randomly assigned by computer-generated numbers in a 1:1 ratio and allocated to the experimental or control groups and to the order of the training topics (behavioral activation and cognitive strategies or vice versa). The allocation will be implemented by a researcher (who is independent of the role-plays). The researcher will enroll and assign participants and provide sealed envelopes containing the study material. To reduce the possibility for unblinding, the researcher and student assistants will urge all participants not to speak to any other students about the trial.

In this educational low-risk trial, there is no formal implementation of a trial steering committee, data monitoring committee, or stakeholder group. Nevertheless, the funder decides upon the interim project report upon further funding. The principal investigators (PIs) are responsible for local organization, which is implemented by the researcher and his student assistants (e.g., regarding recruitment and ensuring informed consent). The PIs and the researcher meet weekly to oversee conduct and progress of the trial. Statistical counseling will be provided by another department. At any time, students

TIMEPOINT	Enrolment	Allocation	Experiment 1				Experiment 2			
	-t ₁	0	Pre	Training	Post	3-months follow-up	Pre	Supervision	Post	3-months follow-up
ENROLMENT:										
Eligibility screen	X									
Informed consent	X									
Allocation		X								
INTERVENTIONS:										
Experiment 1: Modeling video				X						
Experiment 1: Unspecific video				X						
Experiment 2: Video based supervision								X		
Experiment 2: Verbal supervision								X		
ASSESSMENTS:										
Demographic data			X			X			X	X
Competencies			X		X	X	X		X	X
Alliance			X		X	X	X		X	X
Adherence			X		X	X	X		X	X
Knowledge			X		X	X	X		X	X
Empathy			X		X	X	X		X	X
Personality					X				X	
Anxiety			X		X	X	X		X	X
Negative effects			X		X	X	X		X	X
Allegiance							X		X	X
Authenticity			X		X	X	X		X	X
Counseling skills			X		X	X	X		X	X
Study perception					X	X			X	X

Fig. 1 SPIRIT (Standard Protocol Items: Recommendations for Interventional Trials) 2013 figure

may consult the university's ethics review committee and the data security officer if problems or harms arise. The researcher and student assistants will note down reasons for discontinuation as well as any problems that may arise during the trial (e.g., burdens and stress to the participants). During the weekly consultations with the PIs, these problems will be discussed and decisions will be made on how to proceed (e.g., to provide individual support and to stop the trial).

Participants

Participants will be recruited from bachelor's and master's psychology courses, mainly at the University of Potsdam. They will be included if they have given verbal and written informed consent to the study researchers and therefore agreed to the video recordings of the standardized role-plays. They will be excluded if they are currently in psychotherapeutic treatment or have insufficient German language skills to participate. Participants can withdraw from the study at any time without any disadvantages arising. According to their preference, students will receive either course credit or an expense allowance of €40. Their retention will be promoted by emphasizing the learning benefits of participation. If a reason for discontinuation or unintended effects is given, it will be recorded and, if required, classified into categories. If any impairments through study participation are perceived, those affected will be offered a dialogue with one of the PIs, who are also licensed psychotherapists.

Standardized patients

The SPs will be recruited at the University of Potsdam from courses other than psychology, such as the educational or natural sciences, in order to reduce the probability that participants and SPs know each other. They will be employed as student assistants in order to ensure their engagement, and they will be blinded to the experimental conditions. The authenticity of their portrayals will be validated beforehand by independent external raters.

Training intervention and comparator

The first experiment pursues the goal of evaluating the role of modeling for developing therapeutic competencies. The interventions are based on an evidence-based manual for the cognitive behavioral treatment of depression [26].

In the experimental group, participants (i.e., student trainees) will watch a video of an experienced psychotherapist who skillfully demonstrates behavioral activation (video 1) and cognitive (video 2) strategies with an SP demonstrating a depressive disorder. Participants will answer two control questions regarding the contents of the videos as a manipulation check. Additionally, participants

will receive written instructions on behavioral activation (instruction 1) and cognitive biases (instruction 2).

In the control group, participants will watch two un-specific learning tutorials (i.e., solving a Rubik's Cube and folding an origami animal) without psychotherapeutic contents, answer two control questions regarding the contents of the videos, and receive the same written instructions as described above.

Supervision intervention and comparator

The second experiment will evaluate the role of video-based feedback for developing therapeutic competencies and is based on the concept of competency-based supervision [22]. In the experimental group, participants (i.e., student supervisees) will show their video on behavioral activation (video 1) and cognitive (video 2) techniques of an interaction with an SP demonstrating a depressive disorder to a supervisor. Thus, supervision will be behaviorally based (direct supervision) [30]. In the control group, participants will report their experiences with the two role-plays to a supervisor; thus, the supervision will be verbally based (supervision as usual) [31].

Duration

Each of the role-plays will last 20 min. Including the pre- and post-measurement, the first appointment will continue for 3.5 h, and the follow-up appointment will last about an hour. Training and supervision interventions are currently not standard elements of clinical psychology courses at German universities but extend beyond the established topics and didactics.

Supervisors

The supervisors will be licensed psychotherapists (cognitive behavioral therapy). They will be provided with a supervision manual based on Falender and Shafranske [22] and Hautzinger [26]. Supervisors are intended to be recruited from an internal database of former study supervisors. In advance of the first supervision session, they will receive a 10-h training session in order to standardize supervision. Supervisors will not be part of the study team (i.e., they will be external). One supervision session will last 30 min, and each supervisor will receive €50.

Raters

The two raters will be licensed psychotherapists or therapists-in-training (cognitive behavioral therapy) and are also intended to be recruited from an internal database of former raters. In advance of the first session, they will receive 10 h of training on the use of the instruments for the assessment of the video-recorded interactions. Raters will not be part of the study team and will conduct their ratings independently of each other.

They will be blinded to the experimental conditions. As a manipulation check, they will be asked at the end of each rating whether they can identify the condition (experimental versus control). One assessment will take around 30 min, and each rater will receive €20.

Primary outcome

Psychotherapeutic competencies serve as the primary outcome. They will be assessed by trained raters (see above) via the German adaptation [32, 33] of the Cognitive Therapy Scale [34]. Therapeutic techniques will be assessed by the raters using a self-developed therapeutic skills checklist based on published instruments [35–39] and adapted to the psychotherapy field.

Secondary outcomes

Therapeutic alliance and the participant empathy will be evaluated by the raters, participants, and SPs. We will use the Helping Alliance Questionnaire [40], or its German version [41], and the Empathy Scale [42], which will be adapted to the three perspectives.

For measuring therapeutic adherence, we will use the Cognitive Behavioral Therapy Adherence Scale (CBT-AS) [43]. The CBT-AS will be adapted for the topics of behavioral activation and of cognitive biases based on the depression manual used [26].

The therapeutic knowledge of participants will be evaluated by MCQs and case vignettes [5]. The materials will be self-developed in accordance with the above-mentioned manual [26] and recommendations for the construction of MCQs [43].

Further outcomes taken into consideration will be the participant's personality (Big Five Inventory, or BFI-K) [44], state–trait anxiety (State–Trait Anxiety Inventory) [45], and individual perception of the study adapted from [46, 47].

The raters will evaluate the SP's authenticity (Ay DS, Kühne F, Weck F: Can simulated patient encounters appear authentic? Development and pilot results of an observer-based rating instrument, submitted), and the SPs will evaluate the participant's counseling skills [42, 48].

Negative therapist effects will be assessed by items adapted from published German questionnaires [49, 50]. Furthermore, possible allegiance effects of supervisors will be captured by a questionnaire adapted to supervision [51].

Statistical analysis and power calculation

To avoid attrition bias, we will include all randomly assigned participants in the primary analyses (intention-to-treat). Second, we will perform per-protocol analyses and compare the results with those from the intention-to-treat analyses. We will use chi-squared tests for nominal data and analyses of variance/analyses of covariance (ANOVAs/ANCOVAs) for ordinal and interval data. For

the analysis of training/supervision effects, the ANOVAs will involve a group factor (training/supervision condition) and a time factor (before, after, and follow-up). Possible pre-differences between the groups will be taken into account via the ANCOVAs. The α level will be set at 0.05 and Bonferroni-corrected for multiple testing. Missing data will be addressed according to the outcome and data structure (e.g., complete case analyses and imputation by expectation-maximization or by last observation carried forward).

Considering prior studies [24, 52], we assume moderate competence differences between the groups (Cohens $d = 0.60$) and high correlations between competence ratings ($r = 0.50$). Given an α of 0.05 and a $1-\beta$ of 0.80 (power), a sample of 68 participants will be necessary in each experiment, resulting in an overall sample size of 136 (G*Power Version 3.1.9.2) [53].

Discussion

This RCT combines the investigation of specific training and supervision techniques in clinical psychology by using SPs and a validated observer-based rating scale for the assessment of psychotherapeutic competencies. We thus expect the study to contribute to evidence-based education and training in the treatment of mental disorders [1]. Developing SP scenarios for the field of clinical psychology, gathering knowledge in the use of SPs, and assessing competencies within standardized encounters will contribute to further developments within our field.

SPs enable approximating clinical encounters but may not replace them. Therefore, external validity will be lower in this study than in patient studies. The legal regulations for qualification and licensure of psychotherapists are about to be reformed by the German Federal Ministry of Health. In the draft law, practical education is extended even within the master's program [54], which will require evidence-based knowledge on effective training strategies. Above and beyond that, professional bodies such as the American Psychological Association [55] require obtaining and enhancing therapeutic competencies during profession-long learning. Therefore, the results might contribute to the development of licensed therapists and thus to the effective treatment of mental health patients.

Trial status

This is the first protocol version (10 December 2019). Recruitment of participants started on 4 December 2019 and will end when a complete sample size is attained (around 31 December 2021). If protocol modifications prove necessary, we will communicate them within the study publications and the final report to the funder.

Abbreviations

ANCOVA: Analysis of covariance; ANOVA: Analysis of variance; CBT-AS: Cognitive Behavioral Therapy Adherence Scale; MCQ: Multiple-choice question; PI: Principal investigator; RCT: Randomized controlled trial; SP: Simulated/standardized patient

Acknowledgements

We further acknowledge the support of the Deutsche Forschungsgemeinschaft and the Open Access Publishing Fund of the University of Potsdam. We thank Martin Brunner (University of Potsdam), who agreed to support the study via statistical counseling, and Brian Bloch (University of Münster) for his English language editing of the manuscript.

Consent to participate

All participants will provide informed consent prior to participation.

Authors' contributions

FW and FK conceptualized and designed this study. FK drafted the first manuscript, and PEH and FW critically revised it. All authors read and approved the final manuscript.

Funding

The study is funded by the German Research Foundation (DFG) (postmaster@dfg.de) to FW (PI) (WE 4654/10–1) and FK (PI) (KU 3790/2–1). The DFG is not involved in the study design; data collection, management, analysis, or interpretation; the writing of the report; or the decision to submit the report for publication and has no ultimate authority over any of these activities.

Ethics approval and consent to participate

The study was approved by the University of Potsdam's ethics review committee (number 9/2018) and its data security officer. The application involves detailed descriptions of ethical issues such as confidentiality and informed consent materials as well as of data management, access, and security.

Consent to publication

On the consent form, participants will be asked for permission for the research team to share the anonymized data set upon reasonable request from other researchers outside the university. This trial does not involve collecting biological specimens for storage.

Competing interests

The authors declare that they have no competing interests.

Received: 23 December 2019 Accepted: 13 February 2020

Published online: 18 March 2020

References

- McHugh RK, Barlow DH. The dissemination and implementation of evidence-based psychological treatments: a review of current treatments. *Am Psychol*. 2010;65:73–84.
- Fairburn CG, Cooper Z. Therapist competence, therapy quality, and therapist training. *Behav Res Ther*. 2011;49:373–8.
- Rakovshik SG, McManus F. Establishing evidence-based training in cognitive behavioral therapy: a review of current empirical findings and theoretical guidance. *Clin Psychol Rev*. 2010;30:496–516.
- Waltz J, Addis ME, Koerner K, Jacobson NS. Testing the integrity of a psychotherapy protocol: assessment of adherence and competence. *J Consult Clin Psychol*. 1993;61:620–30.
- Muse K, McManus F. A systematic review of methods for assessing competence in cognitive-behavioral therapy. *Clin Psychol Rev*. 2013;33:484–99.
- Miller G. The assessment of clinical skills/competence/performance. *Acad Med*. 1990;65:563–7.
- Dennhag I, Gibbons MBC, Barber JP, Gallop R, Crits-Christoph P. How many treatment sessions and patients are needed to create a stable score of adherence and competence in the treatment of cocaine dependence? *Psychother Res*. 2012;22:475–88.
- Eckel J, Merod R, Vogel H, Neuderth S. Einsatz von Schauspielpatienten in den „Psych“-Fächern des Medizinstudiums – Verwendungsmöglichkeiten in der Psychotherapieausbildung? [Use of Standardized Patients in the Psycho-Social Subjects of Medical Studies – Applicability of Standardized Patients in Postgraduate Psychotherapy Training Curricula?] *Psychother Psychosom Med Psychol* 2014;64:5–11.
- McNaughton N, Ravitz P, Wadell A, Hodges BD. Psychiatric education and simulation: a review of the literature. *Can J Psychiatr*. 2008;53:85–93.
- Adamo G. Simulated and standardized patients in OSCEs: achievements and challenges 1992–2003. *Med Teach*. 2003;25(3):262–70.
- Barrows HS. An overview of the uses of standardized patients for teaching and evaluating clinical skills. *Acad Med*. 1993;68(6):443–51.
- Hill CE, Knox S. Training and supervision in psychotherapy. In: Lambert MJ, editor. *Bergin and Garfield's handbook of psychotherapy and behavior change*. 6th ed. New York: Wiley; 2013. p. 775–811.
- Weaver M, Erby L. Standardized patients: A promising tool for health education and health promotion. *Health Promot Pract*. 2012;13:169–74.
- Hodges BD, Hollenberg E, McNaughton N, Hanson MD, Regehr G. The psychiatry OSCE: a 20-year retrospective. *Acad Psychiatry*. 2014;38:26–34.
- Melluish S, Crossley J, Tweed A. An evaluation of the use of simulated patient role-plays in the teaching and assessment of clinical consultation skills in clinical psychologists' training. *Psychol Learn Teach*. 2007;6:104–13.
- Sheen J, McGillivray J, Gurtman C, Boyd L. Assessing the clinical competence of psychology students through objective structured clinical examinations (OSCEs): student and staff views. *Aust Psychol*. 2015;50:51–9.
- Yap K, Bearman M, Thomas N, Hay M. Clinical psychology students' experiences of a pilot objective structured clinical examination. *Aust Psychol*. 2012;47:165–73.
- Partschefeld E, Strauß B, Geyer M, Philipp S. Simulationspatienten in der Psychotherapieausbildung. [Simulated patients in psychotherapy training]. *Psychotherapeut*. 2013;58:438–45.
- Sholomskas DE, Syracuse-Siewert G, Rounsaville BJ, Ball SA, Nuro KF, Carroll KM. We don't train in vain: a dissemination trial of three strategies of training clinicians in cognitive-behavioral therapy. *J Consult Clin Psychol*. 2005;73:106–15.
- Beidas RS, Kendall PC. Training therapists in evidence-based practice: a critical review of studies from a systems-contextual perspective. *Clin Psychol Sci Pract*. 2010;17:1–30.
- Hill CE, Lent RW. A narrative and meta-analytic review of helping skills training: time to revive a dormant area of inquiry. *Psychother Theory Res Pract Train*. 2006;43:154–72.
- Falender CA, Shafranske EP. *Supervision essentials for the practice of competency-based supervision*. Washington: American Psychological Association; 2017.
- Bennett-Levy J. Therapist skills: a cognitive model of their acquisition and refinement. *Behav Cogn Psychother*. 2006;34:57–78.
- Martino S, Paris M, Anez L, Nich C, Canning-Ball M, Hunkele K, et al. The effectiveness and cost of clinical supervision for motivational interviewing: a randomized controlled trial. *J Subst Abuse Treat*. 2016;68:11–23.
- Kühne F, Maas J, Wiesenthal S, Weck F. Supervision in der Verhaltenstherapie: Ein Scoping Review zur Identifikation von Forschungszielen [Supervision in behavioural therapy: A scoping review for identification of research aims]. *Z Klin Psychol Psychother*. 2017;46:73–82.
- Hautzinger M. *Kognitive Verhaltenstherapie bei Depression [Cognitive behavioral therapy for depression]*. 7th ed. Weinheim: Beltz; 2013.
- Jacobi F, Höfler M, Strehle J, Mack S, Gerschler A, Scholl L, et al. *Psychische Störungen in der Allgemeinbevölkerung. Studie zur Gesundheit Erwachsener in Deutschland und ihr Zusatzmodul Psychische Gesundheit (DEGS1-MH). [Mental disorders in the general population. Study on the health of adults in Germany and the additional module mental health (DEGS1-MH)]*. *Nervenarzt*. 2014;85:77–87.
- Webb CA, DeRubeis RJ, Barber JP. Therapist adherence/competence and treatment outcome: a meta-analytic review. *J Consult Clin Psychol*. 2010;78: 200–11.
- Zarafonitis-Müller S, Kuhr K, Bechdorf A. Der Zusammenhang der Therapeutenkompetenz und Adhärenz zum Therapieerfolg in der Kognitiven Verhaltenstherapie – metaanalytische Ergebnisse. [The relationship between therapist's competence and adherence to outcome in cognitive-behavioural therapy – results of a metaanalysis]. *Fortschritte der Neurologie Psychiatrie*. 2014;82:502–10.
- Milne D, Reiser R, Aylott H, Dunkerley C, Fitzpatrick H, Wharton S. The systematic review as an empirical approach to improving CBT supervision. *Int J Cogn Ther*. 2010;3:278–94.

31. Weck F, Kaufmann Y, Höfling V. Competence feedback improves CBT competence in trainee therapists: a randomized controlled pilot study. *Psychother Res*. 2017;27:501–9.
32. Weck F, Hautzinger M, Heidenreich T, Stangier U. Erfassung psychotherapeutischer Kompetenzen: Validierung einer deutschsprachigen Version der Cognitive Therapy Scale. [Assessing psychotherapeutic competencies: Validation of a German version of the Cognitive Therapy Scale]. *Z Klin Psychol Psychother*. 2010;39:244–50.
33. Weck F, Grikscheit F, Höfling V, Stangier U. Assessing treatment integrity in cognitive-behavioral therapy: Comparing session segments with entire sessions. *Behav Ther*. 2014;45:541–52.
34. Young JE, Beck AT. *Cognitive Therapy Scale: Rating manual*. Philadelphia: University of Pennsylvania; 1980.
35. Simmenroth-Nayda A, Heinemann S, Nolte C, Fischer T, Himmel W. Psychometric properties of the Calgary Cambridge guides to assess communication skills of undergraduate medical students. *Int J Med Educ*. 2014;5:212.
36. Burt J, Abel G, Elmore N, Campbell J, Roland M, Benson J, et al. Assessing communication quality of consultations in primary care: initial reliability of the Global Consultation Rating Scale, based on the Calgary-Cambridge Guide to the Medical Interview. *BMJ Open*. 2014. <https://doi.org/10.1136/bmjopen-2013-004339>.
37. Scheffer S. Validierung des Berliner Global Rating (BGR) [Validation of the Berlin Global Rating]. Doctoral dissertation. Freie Universität Berlin. 2009. <https://doi.org/10.17169/refubium-9906>. Accessed 2 Feb 2020.
38. Hill C. Supervisor or peer rating form of helper exploration skills. <http://supp.apa.org/books/Helping-Skills-Fifth/student/PDF/WebFormB.pdf>. Accessed 27 Nov 2019.
39. Martin O, Rockenbauch K, Kleinert E, Stöbel-Richter Y. Aktives Zuhören effektiv vermitteln: Zwei Konzepte im Vergleich [Effectively communicate active listening: comparison of two concepts]. *Nervenarzt*. 2017;88:1026–35.
40. Luborsky L. *Principles of psychoanalytic psychotherapy: a manual for supportive- expressive psychotherapy*. New York: Basic Books; 1984.
41. Bassler M, Potratz B, Krauthauser H. Der Helping Alliance Questionnaire (HAQ) von Luborsky. Möglichkeiten zur Evaluation des therapeutischen Prozesses von stationärer Psychotherapie [The "Helping Alliance Questionnaire" (HAQ) by Luborsky]. *Psychotherapeut*. 1995;40:23–32.
42. Partschefeld E. Evaluation des Einsatzes von Simulationspatienten in der psychotherapeutischen Ausbildung: ein Beitrag zur empirischen Fundierung der Psychotherapieausbildung [Evaluation of the use of simulated patients in psychotherapeutic training: a contribution to the empirical foundation of psychotherapy training]. Doctoral dissertation. Jena: Friedrich Schiller University; 2013.
43. Deutsches Ärzteblatt. Regeln zur Konstruktion von MC-Aufgaben [Rules for the construction of multiple choice tasks]. <https://www.aerzteblatt.de/download/files/2008/06/x0000130089.pdf>. Accessed 12 Apr 2017.
44. Rammstedt B, John OP. Kurzversion des Big Five Inventory (BFI-K): Entwicklung und Validierung eines ökonomischen Inventars zur Erfassung der fünf Faktoren der Persönlichkeit. [Short version of the Big Five Inventory (BFI-K): Development and validation of an economic inventory for assessment of the five factors of personality]. *Diagnostica*. 2005;51:195–206.
45. Englert C, Bertrams A, Dickhäuser O. Entwicklung der Fünf-Item-Kurzskala STAI-SKD zur Messung von Zustandsangst [Development of a 5-Item-Short scale STAI-SKD to measure state fear]. *Zeitschrift für Gesundheitspsychologie*. 2011;19:173–80.
46. Wünderlich MS, Nissen C, Feige B, Philipsen AS, Voderholzer U. Portrayal of psychiatric disorders: are simulated patients authentic? *Acad Psychiatry*. 2012;36(6):501–2.
47. Hauschildt M, Schröder J, Moritz S. Randomized-controlled trial on a novel (meta-) cognitive self-help approach for obsessive-compulsive disorder ("myMCT"). *J Obsessive Compuls Relat Disord*. 2016;10:26–34.
48. Hill CE, Kellems IS. Development and use of the Helping Skills Measure to assess client perceptions of the effects of training of helping skills in sessions. *J Couns Psychol*. 2002;49.
49. Linden M. How to define, find and classify side effects in psychotherapy: from unwanted events to adverse treatment reactions. *Clin Psychol Psychother*. 2013;20:286–96.
50. Peth J, Jelinek L, Nestoriuc Y, Moritz S. Unerwünschte Effekte von Psychotherapie bei depressiven Patienten – Erste Anwendung der Positive and Negative Effects of Psychotherapy Scale (PANEPS). [Adverse Effects of Psychotherapy in Depressed Patients - First Application of the Positive and Negative Effects of Psychotherapy Scale (PANEPS)]. *Psychother Psychosom Med Psychol*. 2018;68:391–8.
51. Devilly GJ, Borkovec TD. Psychometric properties of the credibility/ expectancy questionnaire. *J Behav Ther Exp Psychiatry*. 2000;31:73–86.
52. Weck F, Jakob M, Neng JMB, Höfling V, Grikscheit F, Bohus M. The effects of bug-in-the-eye-supervision on therapeutic alliance and therapist competence in cognitive-behavioral therapy: a randomized controlled trial. *Clin Psychol Psychother*. 2016;23:386–96.
53. Faul F, Erdfelder E, Lang AG, Buchner A. G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav Res Methods*. 2007;29:175–91.
54. German Federal Ministry of Health. Arbeitsentwurf eines Gesetzes zur Reform der Psychotherapeutenausbildung [Working draft of a law reforming the training of psychotherapists]. Berlin: Federal Ministry of Health; 2017. https://www.bpm-ev.de/images/Arbeitsentwurf_PTG_Stand_20.07.2017.pdf. Accessed 2 Feb 2020.
55. American Psychological Association. Ethical principles of psychologists and code of conduct. 2017. <https://www.apa.org/ethics/code/>. Accessed 2 Feb 2020.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

