



Humanwissenschaftliche Fakultät

Franziska Kühne | Jana Maas | Sophia Wiesenthal | Florian Weck

# Empirical research in clinical supervision

a systematic review and suggestions for future studies

Suggested citation referring to the original publication:

BMC Psychology 7 (2019) Art. 54

DOI <https://doi.org/10.1186/s40359-019-0327-7>

ISSN (online) 2050-7283

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

Postprints der Universität Potsdam

Humanwissenschaftliche Reihe ; 580

ISSN 1866-8364

<http://nbn-resolving.de/urn:nbn:de:kobv:517-opus4-438721>

DOI <https://doi.org/10.25932/publishup-43872>



RESEARCH ARTICLE

Open Access



# Empirical research in clinical supervision: a systematic review and suggestions for future studies

Franziska Kühne<sup>\*</sup> , Jana Maas, Sophia Wiesenthal and Florian Weck

## Abstract

**Background:** Although clinical supervision is considered to be a major component of the development and maintenance of psychotherapeutic competencies, and despite an increase in supervision research, the empirical evidence on the topic remains sparse.

**Methods:** Because most previous reviews lack methodological rigor, we aimed to review the status and quality of the empirical literature on clinical supervision, and to provide suggestions for future research. MEDLINE, PsycInfo and the Web of Science Core Collection were searched and the review was conducted according to current guidelines. From the review results, we derived suggestions for future research on clinical supervision.

**Results:** The systematic literature search identified 19 publications from 15 empirical studies. Taking into account the review results, the following suggestions for further research emerged: Supervision research would benefit from proper descriptions of how studies are conducted according to current guidelines, more methodologically rigorous empirical studies, the investigation of active supervision interventions, from taking diverse outcome domains into account, and from investigating supervision from a meta-theoretical perspective.

**Conclusions:** In all, the systematic review supported the notion that supervision research often lags behind psychotherapy research in general. Still, the results offer detailed starting points for further supervision research.

**Trial registration:** PROSPERO; [CRD42017072606](https://doi.org/10.1186/1745-6215-4-2606), registered on June 20, 2017.

**Keywords:** Supervision, Clinical supervision, Systematic review, Evidence-based psychotherapy

## Background

Although in psychotherapy training and in profession-long learning, clinical supervision is regarded as one of the major components for change in psychotherapeutic competencies and expertise, its evidence base is still considered weak [1–3]. Clinical supervision is currently considered a distinct competency in need of professional training and systematic evaluation; however, theoretical developments and experience-driven practice still seem to diverge, and “significant gaps in the research base” are evident ([1], p. 88).

Definitions of supervision underline different aspects, whereas a lack of consensus seems to impede research [1]. Falender and Shafranske [4, 5] stress the development of testable psychotherapeutic competencies in the learners, i.e., their knowledge, skills and values/attitudes, through supervision; on the other hand, supervisors need to develop competence to deliver supervision. Milne and Watkins [6] describe clinical supervision as “the formal provision, by approved supervisors, of a relationship-based education and training that is work-focused and which manages, supports, develops and evaluates the work of colleague/s” (p. 4). In contrast, Bernard and Goodyear [7] emphasize supervision’s hierarchical approach, in as much as it is provided by more senior to more junior members of a profession. The goals of

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



supervision may thus range between the poles of being normative (i.e., ensuring quality and case management), restorative (i.e., providing emotional and coping support) and formative (i.e., promoting therapeutic competence), and, thus, may ultimately lead to effective and safe psychotherapy [6]. Hence, it is pivotal for supervisors to reflect upon their own knowledge or skills gaps, and to engage in further qualification [8]. Clinical supervision may involve different therapeutic approaches and thus addresses therapists from varying mental health backgrounds [8], which is the stance taken in the current review.

Besides providing a definition of clinical supervision, it is relevant to delineate related terms. One is *feedback*, a supervision technique that “refers to the ‘timely and specific’ process of explicitly communicating information about performance” ([8], p. 28). Contrary to supervision, *coaching* strives to enhance well-being and performance in personal and work domains [9], and is therefore clearly distinct from supervision and psychotherapy with mental health patients provided by licensed therapists.

In the supervision literature, there is no paucity of narrative reviews, commentaries or concept papers. Previous reviews have revealed positive effects of supervision, for example on supervisee’s satisfaction, autonomy, awareness or self-efficacy [10–13]. Still, results on the impact of supervision on patient outcomes are still considered mixed [10]. Importantly, there is a knowledge gap regarding the active components of supervision, i.e., the effects of supervision or supervisor interventions on supervisees and their patients [10].

Past reviews, however, suffer from several limitations (for details, see [14]). First of all, strategies used for literature search and screening have not always been described or implemented rigorously, that is, implemented in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA [15]) reporting guidelines (e.g. [10–12, 16–19]). Further, several reviews focus specifically on the positive effects of supervision [19] or specifically on learning disabilities [11], emphasize the authors’ point of view [20, 21], or concentrate on the supervisory relationship only [14]. While the majority of the above-mentioned reviews are narrative, Alfonsson and colleagues conducted a systematic review [14], pre-registered and published a review protocol [22] and implemented a thorough literature search and methodological appraisal. However, since they focused exclusively on cognitive behavioral supervision and on experimental designs, only five studies fit their inclusion criteria. Additionally, interrater agreement was only moderate during screening. Likewise, in our previous scoping review [23], we concentrated on cognitive behavioral supervision. Furthermore, like other

supervision reviews [20, 21], it was published in German only, limiting its scope.

Thus, the current systematic review aimed to complement previous reviews by using a comprehensive methodology and concise reporting. First, we aimed to review the current status of supervision interventions (e.g., setting, session frequency, therapeutic background) and of the methodological quality of the empirical literature on clinical supervision. Second, we aimed to provide suggestions for future supervision research.

## Materials and methods

We conducted a systematic review by referring to the PRISMA reporting guidelines [15]. The review protocol was registered and published with the International Prospective Register of Systematic Reviews (PROSPERO; CRD42017072606).

## Inclusion and exclusion criteria

We included studies referring to clinical supervision as defined above by Milne and Watkins [6] above. Both, supervision conducted on its own or as part of a larger intervention (as in psychotherapy training) were included. Treatment studies in which supervision was conducted solely to foster treatment delivery were excluded because they mainly address study adherence and are still covered in other reviews [24, 25]. Furthermore, clinical supervision had to refer to psychotherapy, whereas supportive interventions accompanying other treatments (e.g., clinical management) were excluded. Thus, we included studies referring to mental health patients, and studies with patients with physical diseases were considered only if the reason for treatment was patients’ mental health. Studies with another population (e.g., simulated patients or pseudo-clients) were excluded. In order to focus the review in the heterogeneous field of clinical supervision, we limited it to adult patients. Studies on family therapy were included if they focused on adults. Studies with mixed adult and child/adolescent populations were included if the results were reported for the adult population separately. No prerequisites were predefined for supervisor qualification. Any empirical study published within a peer-reviewed process (i.e., without commentaries or reviews) and any outcome measures were included. As such, any supervision outcome (e.g., supervisees’ satisfaction or competence), including negative or unexpected outcomes (e.g., non-disclosure), were allowed. In line with Hill & Knox [10], we did not focus on studies exclusively examining the supervision process because firstly, it does not provide knowledge on the effectiveness of supervision, and secondly, relationship variables are already covered by other reviews [11]. Thus, the review focused on supervision interventions, and studies exclusively focusing on the

effects of relationship variables or attitudes between the supervisee and supervisor (i.e., as independent variables) were excluded. However, relationship variables were considered if they were considered as dependent variables in the primary studies.

### Study search

The bibliographic database search was conducted during February and March 2017 in key electronic mental health databases (Fig. 1). To include the current evidence, we focused our search on studies published from 1996 onwards. There were no language restrictions. The following search strategy was used: *supervis\* AND (psychotherap\* OR cognitive-behav\* OR behav\* therapy OR CBT OR psychodynamic OR psychoanaly\* OR occupational therapy OR family therapy OR marital therapy) NOT (management OR employ\* OR child\* OR adolesc\*)*. Then, we inspected the reference lists of the included studies (backward search) and conducted a cited reference search (forward search). We finished our search in July 2017.

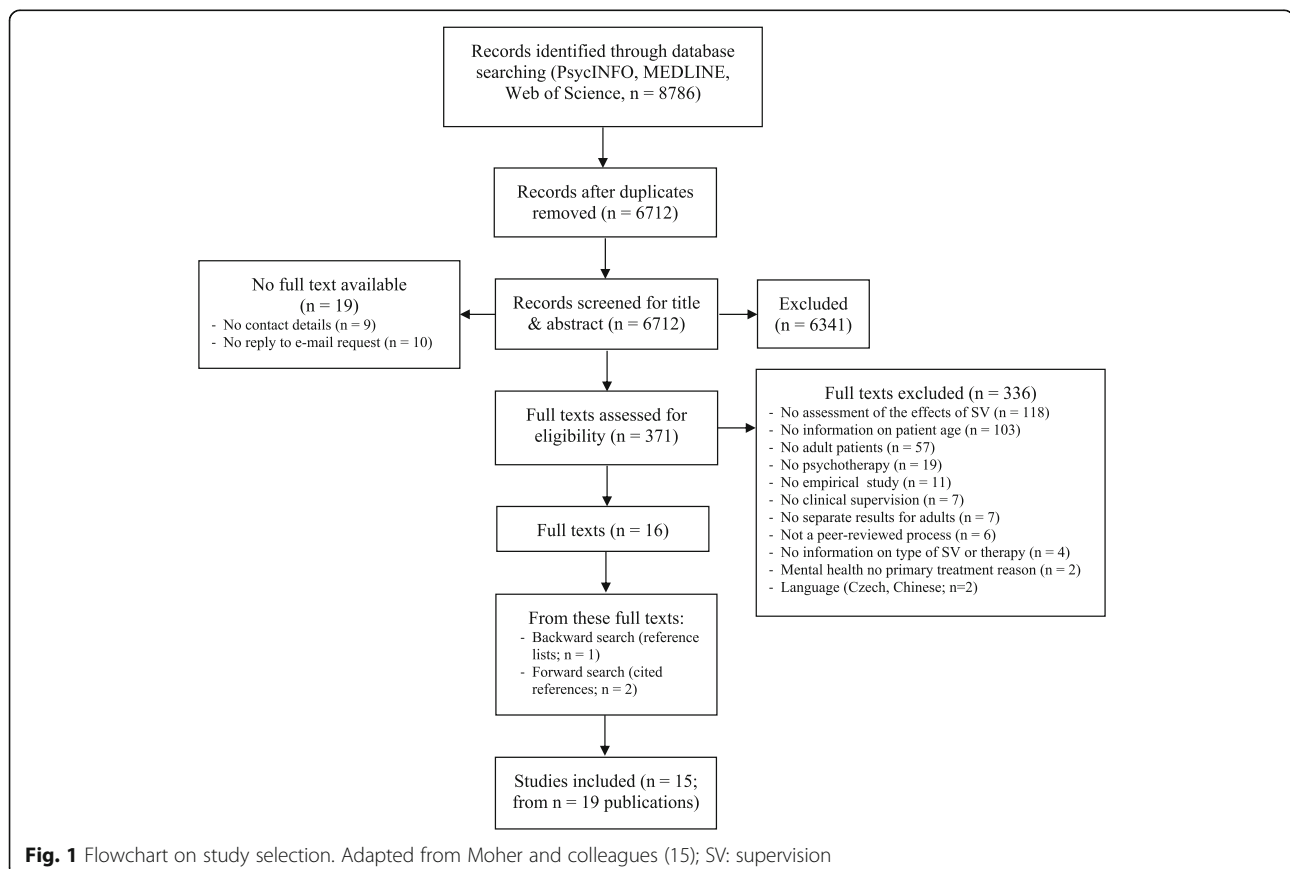
### Screening and extraction

Referring to Perepletchikova, Treat and Kazdin [26], one reviewer (FK) introduced two Master's psychology

students (JM, SW) to the review methods, and the group discussed the review process in weekly one-hour sessions. First, titles and abstracts were screened for inclusion (JM, SW). The first 10% ( $n = 671$ ) of all titles and abstracts were screened by both raters independently. Inter-rater agreement regarding title/abstract screening amounted to  $\kappa = .83$  [CI = .73–.93], which is considered high [27].

Next, full texts of eligible and unclear studies were retrieved and then screened again independently by both raters (JM, SW). Disagreements were resolved through discussion or through the inclusion of a third reviewer (FK). If publications were not available through inter-library loans, a copy was requested from the corresponding author. For nine authors, contact details were not retrievable, and out of the 15 authors that were contacted, five replied. Inter-rater agreement concerning full text screenings for inclusion/exclusion was  $\kappa = .87$  [CI = .77–.97].

For data extraction, we used a structured form that was piloted by three reviewers (FK, JM, SW) on five studies. It comprised information on supervision characteristics (e.g., setting, implementation and competence) and study characteristics (e.g., design, main outcome). Data were extracted independently by two



raters, the results were then compared, and disagreements resolved again by mutual inspection of the original data.

### Methodological quality

Since we included various study designs, we could not refer to one common tool for the assessment of methodological quality. We therefore developed a comprehensive tool applicable to various study designs to allow for comparability between studies. For the development, we followed prominent recommendations [27–29]. The items were as follows: a) an appropriate design regarding the study question; b) the selection of participants; c) measurement of variables/data collection; d) control/consideration of confounding variables; and e) other sources of bias (such as allegiance bias or conflicts of interest). Every item was rated on whether low (1), medium (2) or high (3) threats to the methodological quality were supposed. The resulting sum score ranges from 5 to 15, with higher values indicating the possibility of greater threats to the methodological quality. The methodological quality was rated by two review authors independently (JM or SW and FK). Inter-rater reliability for the sum scores reached ICC<sub>(1, 2)</sub> = .88 [CI = .70–.95], which is considered high [30]. Disagreements in ratings were again resolved through discussion within the review group.

Due to the heterogeneity of the study designs and outcomes, we will present the review results narratively and in clearly arranged evidence tables.

## Results

### Current status of supervision

#### Psychotherapies

Overall, 15 empirical studies allocated to 19 publications were included (Fig. 1). Information on the supervision characteristics is reported on the study level (Table 1). Most of the supervisees used cognitive-behavioral therapy (CBT) as the active intervention [35, 37, 39, 40, 43–45], in four studies, specific interventions such as Motivational Interviewing (MI [38, 42]), Dialectical Behavioral Therapy (DBT [41]); or Problem Solving Treatment (PST [32]) were used, and one study referred to psychodynamic therapy [31] (recommendation to “Conduct supervision from a meta-theoretical perspective”).

#### Supervisions

Only a minority of studies described any form of supervision manual used or any prior training of supervisors [32, 37–39, 42, 43]. In most cases, supervisees were post-graduates or had a PhD degree. Regarding the frequency of supervision sessions, most studies reported weekly sessions [31, 32, 34, 35, 37, 41, 42], and the total number varied considerably from 3 [35] to 78 sessions [31].

Three studies did not describe the supervision frequency [33, 36, 45], and one singled out one supervision session only [44] (recommendation to “Describe how the study is conducted”).

### Interventions

Whereas different forms of feedback or multiple-component supervision interventions were commonly studied, active interventions such as role play were seldom used [37, 39, 40]. Three studies did not describe the interventions used within supervision [35, 44, 45] (recommendation to “Investigate active supervision methods”). Four supervisions used a form of live intervention [36, 41–43], and the remainder conducted supervision face-to-face. All but five studies [32–34, 44, 45] investigated some form of technological support.

### Methodological quality

#### Design

The following sections describe the methodologies used in the studies, which is why all 19 publications are now referred to (Table 2). Five were randomized controlled trials (RCTs [32, 34, 38, 42, 43]), and one was a cluster-RCT [34]. In addition to cohort designs [31, 44], cross-sectional designs were common [35–37, 45, 48, 49]. Only in three publications was follow-up data collected [33, 38, 42]. Most studies covering satisfaction with supervision included one assessment time, usually post-intervention [34, 35, 37, 39, 48, 49].

#### Methodological quality

The assessments of the methodological quality are presented in Table 2. The total methodological quality score was between 9 and 11 in six publications [32, 38, 41–43, 46, 49], between 12 and 13 in eight publications (score of 12–12 [31, 33–36, 45, 49]), and between 14 and 15 in five of the 19 publications [37, 39, 40, 44, 47], with a lower score indicating a lower risk of a threat to the methodological quality. On an item level, most problems referred to the selection of participants, the control of confounders, and other bias such as allegiance bias (Fig. 2; recommendation to “Conduct methodologically stringent empirical studies”).

### Effects of clinical supervision

The most consistent result refers to the high acceptance, satisfaction and the perceived helpfulness of supervision by supervisees [34–37, 39, 41, 44, 48, 49]. Further, the therapeutic relationship [31, 32, 43–45], and therapeutic competence seem to benefit from supervision [37, 38, 40, 42, 43]. On the other hand, non-significant findings [34, 38], small effects [31, 44, 45] and relevant alternative explanations [32, 33, 43, 46] hamper proper conclusions (see Fig. 3).

**Table 1** Supervision characteristics (main studies reported in alphabetical order)

Publication	Therapy	Set- ting	Main mental health problem	SV manual or SVor training	Profession SVor	Therapy manual or SVee training	Profession SVee	Competence level			Supervision	Frequency	Con- tact	For- mat	Tech- nology
								Und	Grad	Post PhD					
Anderson (2012) [31]	PD	O	Depr, anx	-	Manual authors	Manual, instruction, video examples	Licensed PST	-	-	100	CD	Weekly / 1,5yrs	F-t-f	Gr	Audio, video
Bambling (2006) [32]	PST	O	Major depr	Workshop, manual	Graduated in mental health, experienced	Workshop, manual	PS, PST, MHW, SW	-	20	75	5	CD	F-t-f	Ind	-
Davidson (2017) [33]	Psychol- Therapy	O	Depr, anx, stress	-	PS, MHW	-	PS, MHW	-	-	-	-	FBO	F-t-f	Ind	-
Grossl (2014) [34]	Mixed	O	Mental disorders	-	PhD (PS, MFT)	-	Clinical PS, MFT, CS	-	68	32	-	FBO	F-t-f	Ind	-
Hiltunen (2013) [35]	CBT	O	Minor mental health problems	-	PST, experienced	Training	PS	-	100	-	-	-	F-t-f	Gr	Audio, video
Locke (2001) [36]	Mixed	O	-	-	Licensed	Training	-	-	-	-	-	CD, FBP	Live	Gr	Video, phone
Lu (2012) [37]	CBT	O	Comorbid PTSD	Workshop	PhD (PS)	Workshop	PS, CS, NU, SW	-	4	92	4	CD, FBP, RP, expert call	F-t-f	Gr	Audio
Martino (2016) [38]	MI	O	Substance abuse	Workshop, manual, textbook	Certified, licensed CS	Workshop, manual, textbook	Substance abuse CS	9	14	68	1	FBP, coaching	F-t-f	Ind	Audio
Milne (2011) [39]	CBT	O	Depr, anx	Manual, training	Licensed PS	CBT training	PS	-	-	50	-	RP, FBP, others	F-t-f	Ind	Audio
Ng (2007) [40]	CBT	O	Medication-resistant psychosis	-	Certified PT, the CBT trainer	Manual, lectures	NU, SW	-	75	25	-	CD, RP	F-t-f	Gr	Video
Rizvi (2016) [41]	DBT	O	BPD	-	Licensed clinical PS / DBT expert / study author	Training, seminars	Clinical PS	-	-	100	-	FBP	BITE	Ind	PC, webcam
Smith (2012) [42]	MI	O, I	Substance abuse	Workshop	PhD (clinical PS)	Workshop	Substance abuse CS	28	31	40	-	CD, coaching	Live	Ind	Earpiece, phone
Weck (2016) [43]	CBT	O	Depr, anx	Technical instruction	Licensed clinical PS & SVors	CBT training	Clinical PS	-	-	100	-	CD, FBP	F-t-f	Ind	PC, webcam
Willutzki (2005) [44]	CBT	O	Affective disorders, anx	-	-	Training	PS	-	-	100	-	-	Every 4th session	Ind	-
Zarbock (2009) [45]	CBT	O	Affective & phobic disorders, others	-	Experienced	Training	PS	-	-	100	-	-	-	-	-

SV supervisor, SVsor supervisor, SVee supervisee; - not applicable or no information, PD psychodynamic, PST problem-solving therapy, CBT cognitive behavior therapy, DBT dialectical behavior therapy, MI motivational interviewing, Mixed different approaches, Psychol-Therapy psychological therapy, not specified, O outpatient, I inpatient; Depr depression, Anx anxiety, PTSD post-traumatic stress disorder, BPD borderline personality disorder, PST psychotherapist, PS psychologist, MHW mental health worker, SW social worker, MFT marriage and family therapist, CS counselor, PT psychiatrist, NU nurse, Und undergraduate, no degree, student, Grad graduate, Bachelor degree, Post postgraduate, Master's degree, PhD doctoral degree, CD case discussion, FBO feedback on performance, RP role play, coaching provide model behavior, suggest statements, F-t-f face-to-face, BITE bug-in-the-eye, Gr group, Ind individual



**Table 2** Study characteristics (main studies (bold type) and concomitant publications reported together)

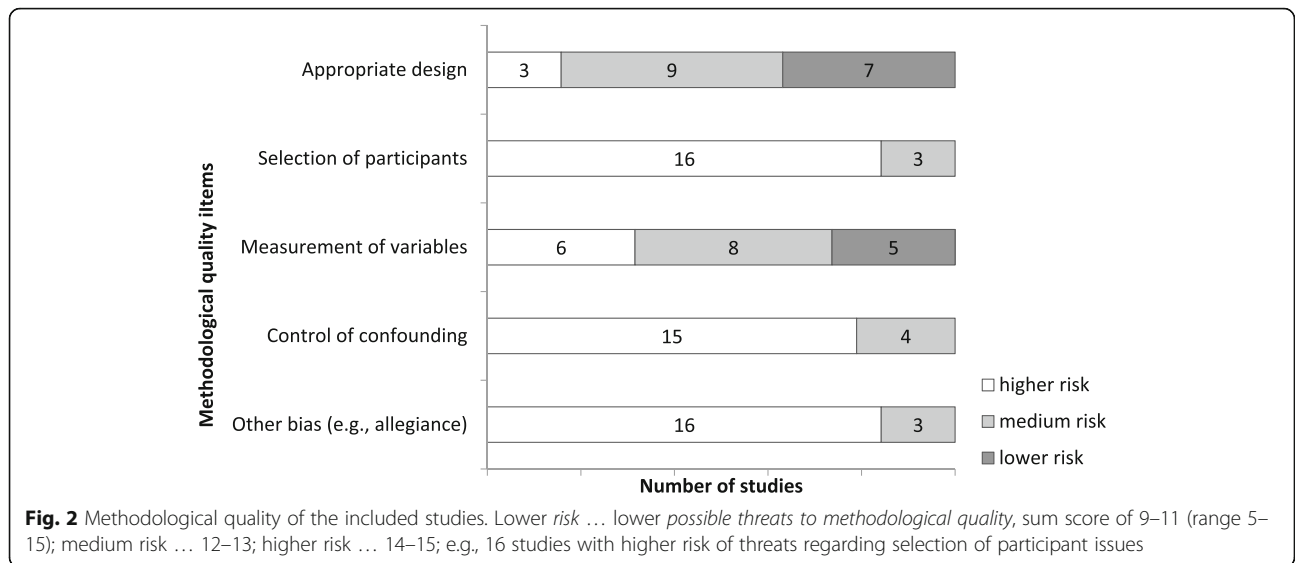
Publication	Design	Intervention group (n patient)	Control group (n patient)	Ass	Q	Main outcomes	Negative effects
Anderson (2012) [31]	Cohort	Cohort year 2: Time-limited PD-SV Cohort year 3: Some early SV	Cohort year 1: No SV (84 all groups)	R	12	Sign. better adherence, therapeutic relationship and advanced PD techniques in PD-SV Small effects that do not seem sustainable	N/S
Anderson (2017) [46]	Cohort	Directive SV	Non-directive SV (40 both groups)	R	11	Sign. greater adherence of SVEes if SVor used directive style Alternative explanations: SVor personality, didactic methods, individual differences of SVEes	-
Bambling (2006) [32]	RCT	Alliance process-focused SV (34) Alliance skill-focused SV (31)	No SV (38)	Q, R	9	Sign. increased therapeutic alliance and decreased depression in all groups, group differences after session 1 Pat. in SV groups sign. More satisfied and less dropout than in control group	-
Davidson (2017) [33]	Cluster RCT	FB to SVEe and SVor on Pat outcome, alerts as to worsening (16)	FB to SVEe on Pat outcome, no alert (25)	Q	13	Pat. in control group sign. Less distressed (post, FU), also in therapists' ratings, but with more sessions Large pat. and therapist drop-out	risk for self-harm evaluated
Grossl (2014) [34]	RCT	FB to SVEe and SVor on Pat outcome	SAU (138 both groups)	Q	12	N.s. differences between groups SVEes in intervention group sign. More satisfied with SV	-
Hiltunen (2013) [35]	CS	CBT-SAU (35)	-	Q	13	Perceived satisfaction with SV	-
Locke (2001) [36]	CS	Live-SV (108)	-	Q	13	Pat. felt comfortable with Live-SV Perceived helpfulness and low intrusiveness of Live-SV predicted therapy satisfaction	-
Lu (2012) [37]	CS	CBT-SV with fidelity FB (26)	-	Q, R	14	SV and E-mail FB perceived as helpful, pat. Symptoms sign. Decreased 91% of SVEes achieved certification with first training case	-
Martino (2016) [38]	RCT	SV on MI (227)	SAU (223)	R, I, T	10	Sign. greater increase in SVEe competency in intervention group (post, FU) N.s. differences in pat. Retention and substance abuse, MI-SV more cost-intensive	27 adverse events, unrelated
Milne (2011) [39]	N = 1 (ABA)	B: Evidence-based clinical SV (3)	A: CBT-SV	Q, R, I, O	14	Intervention perceived as better, experiential learning and high acceptance in both groups	Anxious, rushed, taxing
Milne (2013) [47]	S/A	S/A	S/A	R	15	Apparent SVor fidelity and perceived experiential learning in SVEes	-
Ng (2007) [40]	Pre- post	SV to CBT for psychosis (10)	-	R, CF	15	More acceptable case formulations and sign. Better therapeutic competences after SV	-
Rizvi (2016) [41]	N = 1 (ABA)	B: BITE-SV (1)	A: SAU	Q, R	11	Pat/SVEe perceived BITE as acceptable, SVEe perceived increase in DBT confidence, adequate adherence	-
Smith (2012) [42]	RCT	Live phone-SV on MI with standardized Pat	Audiotape-based phone-SV on MI with standard. Pat; No SV	R	10	Intervention with sign. Greater global MI integrity and skill than Audiotape-based SV than No SV	N/S



**Table 2** Study characteristics (main studies (bold type) and concomitant publications reported together) (Continued)

Publication	Design	Intervention group (n patient)	Control group (n patient)	Ass	Q	Main outcomes	Negative effects
Weick (2016) [43]	RCT	BITE-SV (19)	Delayed video-based SAU (23)	Q, R	11	Audiotope-based SV sign. Better in increasing complex reflections Sign. better therapeutic alliance and competence in intervention group Ns. differences when controlling for baseline scores and for pat. Outcomes	–
Jakob (2013) [48]	CS	BITE subgroup (10)	–	Q	13	High acceptance, perceived helpfulness and usefulness by Pat, SVEe, SVors	Split attention
Jakob (2015) [49]	CS	BITE subgroup (8)	–	I	10	Positive perception of an added value by BITE e.g., on therapeutic competence For SVEes, organizational efforts and anxiety at the beginning	Stress
Willutzki (2005) [44]	Cohort	Additionally requested CBT-SAU	Regular CBT-SAU (104 in total cohort)	Q	14	Perceived problematic therapeutic alliance before additionally requested SV (Pat, SVEe) Small effects on improved therapeutic alliance after SV, high satisfaction with SV	–
Zarbock (2009) [45]	CS	SAU: Multimodal BT (90)	–	Q	13	Supervisory relationship as best predictor of overall SV satisfaction Low correlation between SVor and SVEe ratings of SV	–

Ass assessment methods, SV supervision, SVsor supervisor, SVEe supervisee, SAU supervision as usual, Pat patient, – not applicable or no information, S/A see above, N/S indicated but not specified, PD psychodynamic, CBT cognitive / behavior therapy, DBT dialectical behavior therapy, MI motivational interviewing, RCT randomized-controlled trial, CS cross-sectional study, N = N of 1 trial, ABA withdrawal; Rat rating, Ques questionnaire, Int interview, Obs observation, T test, CF case formulation, FB feedback, MI motivational interviewing, BITE bug-in-the-eye, FU follow-up, Ns./sign. non/significant, Q methodological quality, 5 (lowest) to 15 (highest possible threat



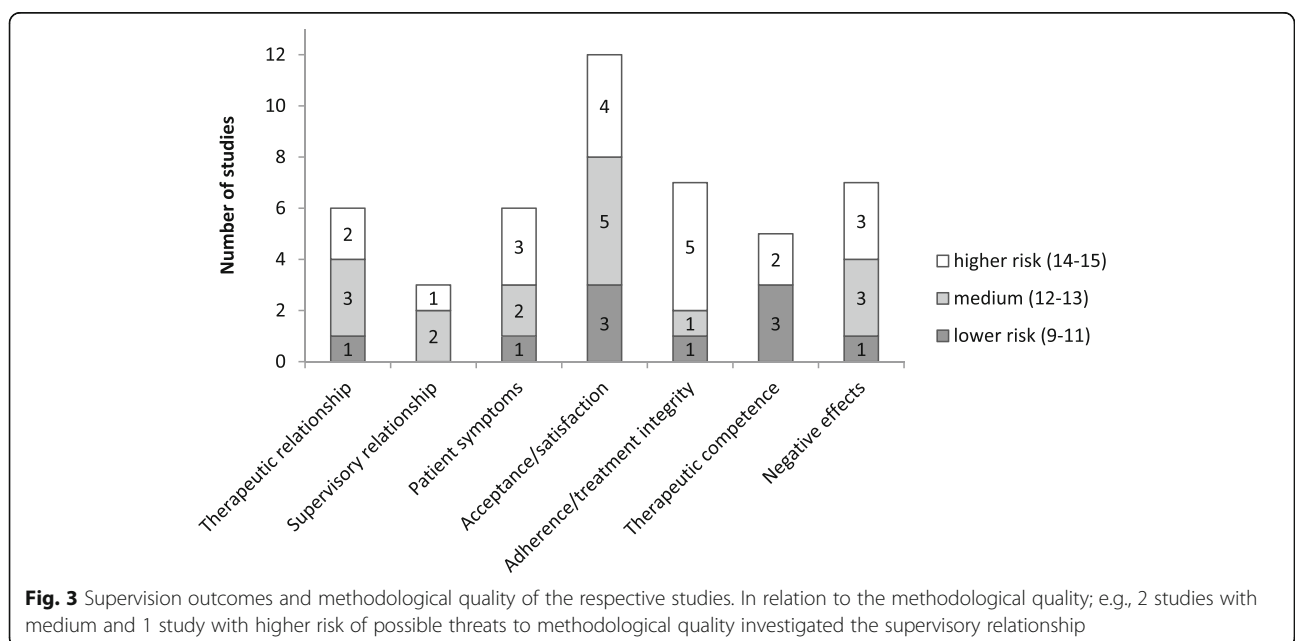
Whereas most publications did not describe negative or unexpected effects of supervision, two mentioned them without further specification [31, 42], two referred to unwanted effects as being unrelated to the outcome [33, 38], and three described limits to therapists’ cognitive capacity and perceived anxiety or stress during supervision [39, 48, 49] (recommendation to “Investigate diverse positive and negative supervision outcomes aside from acceptance”).

**Discussion**

The aim of the present study was to systematically review the status and quality of the current empirical

literature on clinical supervision and, based on the review findings, to draw conclusions for future studies. The current review identified 19 publications referring to 15 empirical studies on the status of clinical supervision. Despite using wide inclusion criteria, it is remarkable that only such a small number of studies could be included. In contrast to former reviews, our study was conducted systematically according to current guidelines, using a reproducible methodology and concise reporting. Compared to previous reviews, it was not limited to psychotherapeutic approaches or study designs.

Regarding the psychotherapeutic approaches of the supervisees, most interventions had a CBT background,



which still documents a research gap in studies on clinical supervision between CBT and other therapeutic approaches.

Aside from psychotherapy approaches, the meta-theoretical perspective of competency-based supervision, as proposed by the American Psychological Association [8], provides a more integrative and broader view. Their supervision guidelines involve seven key domains central to good-quality supervision, from supervisor competencies to diversity or ethical issues. Importantly, they describe supervision to be science-informed, which again underlines the importance of supervisors and supervisees to keep their evidence-based knowledge and skills up-to-date during profession-long learning.

Considering the conduction of supervision, face-to-face supervision was prevalent, but technological support was common as well, at least in published empirical studies. A variety of interventions was used, including less active ones such as case discussions and coaching, as well as more active ones such as feedback on patient outcomes or supervisee performance. It is clearly positive that active interventions (such as coaching and feedback) were implemented and evaluated because they have proven useful in active learning and therapist training [50]. Nevertheless, even more active methods, such as exercise or role play, were an exception [23]. Furthermore, it remains unclear which interventions are helpful in profession-long learning and maintenance of expertise [21, 23]. We found that central supervision characteristics, such as the training of supervisors or the manual used for supervision, were not described consistently. Although a detailed description of how studies were conducted seems intuitive, it is surprising that reporting guidelines are not referred to consistently.

Concerning design characteristics, most studies were uncontrolled or used small samples. Further constraints were associated with the lack of follow-up data and major inconsistencies in the evaluation of negative effects. Although external observers, which were only sometimes independent, were used, almost half of the studies relied exclusively on self-reported questionnaires. Another problem was that the heterogeneity in the designs and instruments hampered the quantitative summary of results. Methodological quality has been criticized in supervision research for years (e.g. [16, 17]), and inconclusive findings or relevant alternative explanations additionally impeded firm conclusions on supervision effects. Regarding the effects of clinical supervision, the review documents that supervision research clearly lags behind psychotherapy research in general; that is, we still have limited evidence on supervision effects, especially those regarding patient benefits [10], and we continue to search for active supervision ingredients [51].

Acceptance and satisfaction are crucial prerequisites for supervision effects, and they were the variables most frequently investigated. Although positive results in these domains may be considered stable [13], satisfaction may not be confused with effectiveness. Taken from health care-related conceptualizations [52], subjective satisfaction may depend on a number of variables, such as mutual expectations, communication, the supervisory relationship, the access to supervision or financial strains. In this sense, satisfaction is distinct from learning and competence development. Other important outcomes of supervision, such as the therapeutic relationship and competencies, treatment integrity, patient symptoms or unwanted effects, clearly need further investigation [10, 21]. Other ideas include considering not only the supervisory relationship but also supervisory expectations as important process variables across psychotherapeutic approaches [13].

### Limitations

We constructed a short tool for rating methodological quality, which enabled comparisons between the diverse designs of the studies included. Although inter-rater reliability was high, it lacks comparability with other reviews. Due to a stricter operationalization of the inclusion criteria, six studies were included in our previous scoping review [23], and three were included in another current review [14] that were not part of the current systematic review. More specifically, one study was not located via our search strategy, and the other publications did not describe explicitly if the patients were adults. As the excluded publications were mainly referring to CBT supervision, it generally reflects the stronger evidence-base of CBT that has its roots in basic research. Since the review aimed to illustrate the status and quality of supervision research, we did not restrict it to specific designs, but mapped the status quo. This necessarily increased heterogeneity, and especially regarding supervision effects, it limited the possibility to draw clear-cut conclusions or to combine the results statistically. Differences in the results of reviews may result not only from methodological aspects but also from diversity in the primary studies, which may be addressed only by better supervision research [14].

### Conclusions

The review provides a variety of starting points for future research. The recommendations derived mainly refer to the replicability of research (i.e., to conduct methodologically stringent empirical studies, and to include positive and negative supervision outcomes). Taking a competency-based view, the following are examples of significant foci of both future practice and supervision research [23, 53, 54]:

- Define, review and continuously develop supervisor competencies.
- Include active methods, live feedback and video-based supervision.
- Enhance the deliberate commitment to ethical standards to protect patients.
- Positively value and include scientific knowledge and progress.
- Foster profession-long learning of supervisees and supervisors.

Logistics may be an important issue in supervision research. Therefore, if large-scale quantitative studies are difficult to conduct or fund, methodologically sound pragmatic trials [3] and experimental studies may be feasible alternatives. Most of the results still speak to the lack of scientific rigor in supervision research. Thus, we consider competency-based supervision and research investigating the essential components of supervision as the major goals for future supervision research and practice.

#### Abbreviations

CBT: Cognitive-behavioral therapy; DBT: Dialectical behavioral therapy; MI: Motivational interviewing; PRISMA: Preferred reporting items for systematic reviews and meta-analyses; PROSPERO: International prospective register of systematic reviews; PST: Problem solving treatment; RCT: Randomized controlled trial; SV: Supervision

#### Acknowledgements

We would like to thank the two reviewers for their valuable and important contributions to a former version of the manuscript.

#### Authors' contributions

FK conceptualized the research goal, developed the design and the methodology, provided the resources needed for the study, supervised and managed the research, collected the data/evidence, analyzed, synthesized and visualized the study data and wrote the initial draft of the paper. JM and SW aided in collecting the data, in analyzing, synthesizing and visualizing the data and revised the work. FW took part in the conceptualization process, the coordination of the responsibilities, the validation and reviewing process and supervised the research activity. All authors read and approved the final manuscript.

#### Funding

We gratefully acknowledge the support of the Deutsche Forschungsgemeinschaft (DFG) and the Open Access Publishing Fund of the University of Potsdam.

#### Availability of data and materials

All data generated or analyzed during this study are included in the published article.

#### Ethics approval and consent to participate

Not applicable.

#### Consent for publication

Not applicable.

#### Competing interests

The authors declare that they have no competing interests.

Received: 4 March 2019 Accepted: 18 July 2019

Published online: 22 August 2019

#### References

1. Falender CA, Shafranske EP. Supervision essentials for the practice of competency-based supervision. Washington, DC: American Psychological Association; 2017.
2. Goodyear RK, Lichtenberg J, Hutman H, Overland E, Bedi R, Christiani K, et al. A global portrait of counselling psychologists' characteristics, perspectives, and professional behaviors. *Couns Psychol Q*. 2016;29(2):115–38.
3. Owen J. Supervisory processes in the training of psychotherapists: introduction to the special section. *Psychotherapy*. 2015;52(2):151–2.
4. Falender CA, Shafranske EP. Competence in competency-based supervision practice: construct and application. *Prof Psychol Res Pract*. 2007;38(3):232–40.
5. Falender CA, Cornish JAE, Goodyear R, Hatcher R, Kaslow NJ, Leventhal G, et al. Defining competencies in psychology supervision: a consensus statement. *J Clin Psychol*. 2004;60(7):771–85.
6. Milne DL, Watkins CE. Defining and understanding clinical supervision: a functional approach. In: Watkins CE, Milne DL, editors. *The Wiley international handbook of clinical supervision*. Chichester: Wiley; 2014. p. 3–19.
7. Bernard JM, Goodyear RK. *Fundamentals of clinical supervision*, 3rd ed. Needham Heights: Allyn & Bacon; 2004.
8. American Psychological Association. Guidelines for clinical supervision in health service psychology. *Am Psychol*. 2015;70(1):33–46.
9. International Society for Coaching Psychology. <https://www.isfcp.info/what-is-coaching-psychology/>. Accessed 19 June 2019.
10. Hill CE, Knox S. Training and supervision in psychotherapy. In: Lambert MJ, editor. *Bergin and Garfield's handbook of psychotherapy and behavior change*. 6. Hoboken: Wiley; 2013. p. 775–812.
11. Milne DL, James I. A systematic review of effective cognitive-behavioral supervision. *Br J Clin Psychol*. 2000;39(2):111–27.
12. Wheeler S, Richards K. The impact of clinical supervision on counsellors and therapists, their practice and their clients. A systematic review of the literature. *Couns Psychother Res*. 2007;7(1):54–65.
13. Watkins CE. How does psychotherapy supervision work? Contributions of connection, conception, allegiance, alignment, and action. *J Psychother Integr*. 2017;27(2):201.
14. Alfonsso S, Parling T, Spännargård Å, Andersson G, Lundgren T. The effects of clinical supervision on supervisees and patients in cognitive behavioral therapy: a systematic review. *Cogn Behav Ther*. 2017;47(3):1–23.
15. Moher D, Liberati A, Tetzlaff J, Altman DG. Prisma Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med*. 2009;6(7):e1000097.
16. Ellis MV, Krengel M, Ladany N, Schult D. Clinical supervision research from 1981 to 1993: a methodological critique. *J Couns Psychol*. 1996;43(1):35–50.
17. Freitas GJ. The impact of psychotherapy supervision on client outcome: a critical examination of 2 decades of research. *Psychotherapy*. 2002;39(4):354–67.
18. 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(5):496–516.
19. Reiser RP, Milne DL. A systematic review and reformulation of outcome evaluation in clinical supervision: applying the fidelity framework. *Train Educ Prof Psychol*. 2014;8(3):149–57.
20. Auckenthaler A. Supervision of psychotherapy: claims, facts, trends. [Supervision von Psychotherapie: Behauptungen-Fakten-trends]. *Psychotherapeut*. 1999;44(3):139–52.
21. Strauß B, Wheeler S, Nodop S. Clinical supervision. Review of the state of research. [Klinische supervision: Überblick über den stand der Forschung]. *Psychotherapeut*. 2010;55(6):455–64.
22. Alfonsso S, Spännargård Å, Parling T, Andersson G, Lundgren T. The effects of clinical supervision on supervisees and patients in cognitive-behavioral therapy: a study protocol for a systematic review. *Syst Rev*. 2017;6(1):1–6.
23. Kühne F, Maas J, Wiesenthal S, Weck F. Supervision in behavioral therapy. A scoping review for identification of research objectives. [Supervision in der Verhaltenstherapie. Ein Scoping Review zur Identifikation von Forschungszielen]. *Zeitschrift für Klinische Psychologie und Psychotherapie*. 2017;46(2):73–82.

24. Muse K, McManus F. A systematic review of methods for assessing competence in cognitive-behavioural therapy. *Clin Psychol Rev*. 2013;33(3):484–99.
25. Roth AD, Pilling S, Turner J. Therapist training and supervision in clinical trials: implications for clinical practice. *Behav Cogn Psychother*. 2010;38(3):291–302.
26. Perepletchikova F, Treat TA, Kazdin AE. Treatment integrity in psychotherapy research: analysis of the studies and examination of the associated factors. *J Consult Clin Psychol*. 2007;75(6):829–41.
27. Higgins JP, Green S. Cochrane handbook for systematic reviews of interventions version 5.1. 0 (updated March 2011). *Cochrane Collab*; 2011.
28. CASP Checklist. <http://www.casp-uk.net/checklists>. Accessed 08 Jan 2018.
29. Sanderson S, Tatt ID, Higgins J. Tools for assessing quality and susceptibility to bias in observational studies in epidemiology: a systematic review and annotated bibliography. *Int J Epidemiol*. 2007;36(3):666–76.
30. Wirtz MA, Caspar F. Beurteilerübereinstimmung und Beurteilerreliabilität: Methoden zur Bestimmung und Verbesserung der Zuverlässigkeit von Einschätzungen mittels Kategoriensystemen und Ratingskalen. Göttingen: Hogrefe; 2002.
31. Anderson T, Crowley MEJ, Patterson CL, Heckman BD. The influence of supervision on manual adherence and therapeutic processes. *J Clin Psychol*. 2012;68(9):972–88.
32. Bambling M, King R, Raue P, Schweitzer R, Lambert W. Clinical supervision: its influence on client-rated working alliance and client symptom reduction in the brief treatment of major depression. *Psychother Res*. 2006;16(3):317–31.
33. Davidson KM, Rankin ML, Begley A, Lloyd S, Barry SJE, McSkimming P, et al. Assessing patient progress in psychological therapy through feedback in supervision: the memos randomized controlled trial (measuring and monitoring clinical outcomes in supervision: memos). *Behav Cogn Psychother*. 2017;45(3): 209–24.
34. Grossl AB, Reese RJ, Norsworthy LA, Hopkins NB. Client feedback data in supervision: effects on supervision and outcome. *Train Educ Prof Psychol*. 2014;8(3):182–8.
35. Hiltunen AJ, Kocys E, Perrin-Wallqvist R. Effectiveness of cognitive behavioral therapy: an evaluation of therapies provided by trainees at a university psychotherapy training center. *Psych J*. 2013;2(2):101–12.
36. Locke LD, McCollum EE. Clients' views of live supervision and satisfaction with therapy. *J Marital Fam Ther*. 2001;27(1):129–33.
37. Lu W, Yanos PT, Gottlieb JD, Duva SM, Silverstein SM, Xie H, et al. Use of fidelity assessments to train clinicians in the CBT for PTSD program for clients with serious mental illness. *Psychiatr Serv*. 2012;63(8):785–92.
38. 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 Abus Treat*. 2016;68:11–23.
39. Milne DL, Reiser RP, Cliffe T, Breese L, Boon A, Raine R, et al. A qualitative comparison of cognitive-behavioural and evidence-based clinical supervision. *Cogn Behav Ther*. 2011;4(4):152–66.
40. Ng RMK, Cheung MSM. Supervision of cognitive behavioural therapy for psychosis: a Hong Kong experience. *Hong Kong J Psychiatry*. 2007;17(4):124–30.
41. Rizvi SL, Yu J, Geisser S, Finnegan D. The use of "bug-in-the-eye" live supervision for training in dialectical behavior therapy: a case study. *Clin Case Stud*. 2016;15(3):243–58.
42. Smith JL, Carpenter KM, Amrhein PC, Brooks AC, Levin D, Schreiber EA, et al. Training substance abuse clinicians in motivational interviewing using live supervision via teleconferencing. *J Consult Clin Psychol*. 2012;80(3):450–64.
43. Weck F, Jakob M, Neng JMB, Hofling V, Grikscheit F, Bohus M. The effects of bug-in-the-eye supervision on therapeutic alliance and therapist competence in cognitive-behavioural therapy: a randomized controlled trial. *Clin Psychol Psychother*. 2016;23(5):386–96.
44. Willutzki U, Tönnies B, Meyer F. Psychotherapy supervision and the therapeutic alliance-A process study. [Psychotherapiesupervision und die therapeutische Beziehung-Eine Prozessstudie.]. *Verhaltenstherapie & Psychosoziale Praxis*. 2005;37(3):507–16.
45. Zarbock G, Drews M, Bodansky A, Dahme B. The evaluation of supervision: construction of brief questionnaires for the supervisor and the supervisee. *Psychother Res*. 2009;19(2):194–204.
46. Anderson T, Crowley MEJ, Binder JL, Heckman BD, Patterson CL. Does the supervisor's teaching style influence the supervisee's learning prescribed techniques? *Psychother Res*. 2017;27(5):549–57.
47. Milne DL, Reiser RP, Cliffe T. An N=1 evaluation of enhanced CBT supervision. *Behav Cogn Psychother*. 2013;41(2):210–20.
48. Jakob M, Weck F, Bohus M. Live supervision: From the one-way mirror to video-based online-supervision. [Live-Supervision: Vom Einwegspiegel zur videobasierten Online-Supervision.]. *Verhaltenstherapie*. 2013;23(3):170–80.
49. Jakob M, Weck F, Schornick M, Krause T, Bohus M. When the supervisor is watching. Qualitative analysis of the acceptance of live supervision. [Wenn der supervisor zuschaut: qualitative analyse der Akzeptanz von live-supervision.]. *Psychotherapeut*. 2015;60(3):210–5.
50. 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):1–30.
51. Rakovshik SG, McManus F, Vazquez-Montes M, Muse K, Ougrin D. Is supervision necessary? Examining the effects of internet-based CBT training with and without supervision. *J Consult Clin Psychol*. 2016;84(3):191–9.
52. Brettschneider C, Lühmann D, Raspe H-H. Der Stellenwert von patient-reported outcomes (PRO) im Kontext von health technology assessment (HTA): DIMDI; 2010.
53. Gonsalvez CJ, Calvert FL. Competency-based models of supervision: principles and applications. *Promises Chall Aust Psychologist*. 2014;49(4):200–8.
54. Falender CA. Clinical supervision-the missing ingredient. *Am Psychol*. 2018; 73(9):1240.

### 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](https://biomedcentral.com/submissions)

