The Cognitive Behaviour Therapist (2025), vol. 18, e25, page 1 of 16
doi:10.1017/S1754470X25000145

BABCP
www.babcp.com

ORIGINAL RESEARCH

Supervisees' evaluations of a structured video-supervision technique

Dominik Henrich, Julia A. Glombiewski and Saskia Scholten

Department of Clinical Psychology and Psychotherapy, University of Kaiserslautern-Landau (RPTU), Landau, Germany
Corresponding author: Dominik Henrich; Email: d.henrich@rptu.de

(Received 6 August 2024; revised 7 March 2025; accepted 10 March 2025)

Abstract

Video-recordings are considered crucial for effective supervision, but empirical evidence is limited. We thus conducted a naturalistic study to assess supervisees' evaluations of a structured video-supervision technique ('Give Me 3' [GM3]) compared with video-supervision as usual (VSAU) and supervision based on self-report (SAU). Twenty-four participants in postgraduate training in cognitive behavioural therapy conducted each supervision method and provided quantitative ratings of supervision satisfaction. A subsample of nine trainees participated in semi-structured qualitative interviews. Repeated-measure ANOVAs revealed no difference in post-session ratings of supervisees' satisfaction with supervision. Retrospective comparisons of the three supervision methods resulted in significant differences in perceived effort and comfort but not in usefulness. The qualitative interviews suggested that the supervisees generally appreciated the benefits of video-supported supervision, while evaluations of GM3 and VSAU were mixed. VSAU and GM3 both appeared to have distinct advantages which might enrich supervision if aligned with the therapeutic context and supervisees' goals.

Key learning aims

(1) How does trainees' supervision satisfaction differ between supervision with or without the review of video-recorded therapy sessions?
(2) Is structured video-supervision perceived by trainee supervisees as being more satisfactory than unstructured video-supervision?
(3) How do supervisees perceive the differences between structured versus unstructured (video) supervision? When is each method deemed most useful?

Keywords: psychotherapy training; supervision satisfaction; video-supervision

Introduction

Supervision is a core component of psychotherapy training, with positive effects on patient outcome and therapeutic competence (Foa *et al.*, 2020; Henrich *et al.*, 2023). The use of video-recordings in supervision is considered by many experts an important component of effective supervision (e.g. Aveline, 1992; Haggerty and Hilsenroth, 2011; Huhra *et al.*, 2008). Video-review allows the analysis of a therapeutic interaction as it happened, which can facilitate other supervisory interventions (e.g. competency-based feedback, role-playing, guided self-exploration). However, while videos have been intensively used in supervision for decades, few studies have evaluated the effect of specific video-supervision techniques on the process (e.g. supervisee satisfaction) or outcome of supervision (e.g. therapist competence, patient outcome) (Aafjes-van Doorn *et al.*, 2022).

© The Author(s), 2025. Published by Cambridge University Press on behalf of British Association for Behavioural and Cognitive Psychotherapies

https://doi.org/10.1017/S1754470X25000145 Published online by Cambridge University Press

The effectiveness of video- or audio-supported supervision

Qualitative studies suggest that therapists generally appreciate the benefits of recording-based supervision; supervisees have emphasized that it allowed them to notice otherwise hidden aspects of therapy (Brown *et al.*, 2013), aided self-reflection and self-awareness (Gossman and Miller, 2012), or facilitated the acquisition of clinical skills and reassured supervisees of their strengths (Routon, 2016). However, the results of quantitative analyses are mixed. With a randomized analogue experimental design, Bearman *et al.* (2017) found that supervision which incorporated role-playing, modelling, and video-based feedback led to greater increases in psychology students' observer-rated fidelity in cognitive behavioural therapy (CBT) compared with supervision based on case discussion. However, this effect could not be attributed entirely to the use of tape-based feedback, as the supervision conditions differed in other aspects, as well. In contrast, Weck *et al.* (2024) selectively manipulated video-use in supervision in a randomized controlled trial with a similar sample and comparable design. Seventy-three bachelor's- and master's-level psychology students were trained in CBT skills (behavioural activation, cognitive restructuring), applied them in role-plays with standardized patients, and were randomly assigned to one session of either video-supervision or supervision based on self-report. The authors found no significant difference in supervision effectiveness based on self- and observer-rated measures of supervisees' competence, empathy, state-trait anxiety, and the quality of the therapeutic alliance during the standardized patient role-plays. This result was corroborated by analyses based on two samples of licensed mental health professionals. In a randomized, controlled implementation trial, Monson *et al.* (2018) found no difference in patient outcome between therapists whose supervision included audio-tape review and supervision based on self-report. Accordingly, Miller *et al.* (2004) found three enhanced training groups (individual audio-based feedback and/or individual coaching) equally effective in strengthening observer-rated competence compared with workshop-only and a control condition. These results cast doubt on whether video-supervision is generally more effective than supervision based on self-report.

There is some indication that a smaller dosage of video or audio review and feedback might be more efficient than maximizing its use in supervision. Weck *et al.* (2021) conducted a randomized controlled trial to analyze the effect of structured video-based competence feedback on therapists' observer-rated competence and patient outcome. The authors found a significant effect on competence but not on patient outcome. The increase in competence occurred after the first feedback and did not change significantly thereafter. This is in line with results from Stirman *et al.* (2017), who evaluated two cohorts of therapists trained to implement CBT in a community mental health setting. The first cohort received weekly 1-hour individual tape-based feedback plus 2 hours of group supervision; the second cohort participated in 2 hours of weekly group supervision that included the mandatory review of short audio segments. Both cohorts achieved the same gain in observer-rated competence. A study by Carpenter *et al.* (2012) indicated that the effects of audio-based versus live supervision on competence in motivational interviewing might be moderated by therapist-level attributes (e.g. educational degree, verbal repertoire). Verbal repertoire was measured with a multiple-choice test where participants had to choose a word with a similar meaning compared with a target word. The authors found that audio-based supervision was most effective for clinicians with a graduate degree, while live supervision was most effective for clinicians without a graduate degree and with stronger vocabulary performance.

Taken together, current evidence for the effectiveness of video-supported supervision is mixed and few studies have rigorously compared the use of recordings with supervision based on self-report. There is some indication that the effects of recording-based supervision might be moderated by therapist-level variables and by how and when recordings are used.

Recommendations on the use of video in supervision

There is a general consensus that the use of video in supervision demands some methodical considerations among supervisors (e.g. Haggerty and Hilsenroth, 2011; Huhra *et al.*, 2008). It has

https://doi.org/10.1017/S1754470X25000145 Published online by Cambridge University Press

been highlighted that video-review can make supervisees feel uncomfortable and anxious (e.g. Brown *et al*., 2013; Topor *et al*., 2017) and that the depth of information that a video provides should be managed by focusing on sequences, certain interventions, or objectives (Bernard and Goodyear, 2019). Video-recordings are a supervision tool, not a technique, and video-supported supervision can take on many forms, which is why a precise operationalization of a video-supported supervision technique can facilitate our ability to evaluate its effects. Examples of two specific, structured video-supervision techniques, the 'I-Spy'technique and the 'Give Me 5' technique, were described by Gonsalvez *et al.* (2016). The 'I-Spy' technique includes focusing the video-review on the occurrence of a specific, pre-defined, in-session behaviour (e.g. a therapeutic microskill). The 'Give Me 5' technique employs a broader focus. Supervisees are instructed to select a 15- to 20-minute video-sequence that is representative of the respective therapy session. The supervisee reviews the sequence prior to supervision and notes the five most relevant observations on an observation sheet. In supervision, the sequence is replayed, and the supervisor also notes their five most relevant observations. The supervisor and supervisee then compare their observations and discuss conclusions. Depending on the observations, other active learning strategies (e.g. role-playing) can be utilized to teach skills or solve difficult therapeutic situations (Gonsalvez *et al*., 2016). Both techniques have not yet been evaluated empirically.

Objectives

Little is known about the effects of video-supported clinical supervision *per se*, and even less is known about the effects of specific video-supported supervision techniques. In the current study, we aimed to strengthen the evidence base by assessing supervisees' evaluations of a modified version of the 'Give Me 5' technique (Gonsalvez *et al*., 2016). We referred to the modified protocol as the 'Give Me 3' (GM3) technique (see the Method section of the main study for a detailed description). We chose the technique because it is a structured video-supervision procedure based on experiential evidence provided by Gonsalvez *et al.* (2016) and combines several advantages of video-review (e.g. objective sequence review, facilitation of supervisee's self-reflection, competency-orientated comparison of supervisor's and supervisee's perceptions).

In the *main study*, we used a naturalistic design to compare supervisees' evaluations of GM3 with supervision based on self-report (supervision as usual [SAU]) and unstructured video-supported supervision (video-supervision as usual [VSAU]). The study was pre-registered at AsPredicted (Henrich *et al*., 2020). We hypothesized that supervisees' satisfaction with GM3 supervision would be greater than with VSAU and SAU.

To better understand the results of the main study, we conducted semi-structured qualitative interviews with a subsample of participants (supervisees) from the main study (*interview study*). Our primary aim was to explore how SAU, VSAU, and GM3 differed in terms of feasibility, comfort, and perceived usefulness and under which conditions supervisees would prefer each method.

Main study

Method

Supervision conditions

The GM3 supervision condition was derived from the 'Give Me 5' procedure described by Gonsalvez *et al.* (2016). In the original procedure, a 15- to 20-minute video-sequence is reviewed by supervisee and supervisor, followed by a discussion of the five most important observations. We modified the procedure to fit within the time frame of regular case supervision in German psychotherapy training (approximately 22 minutes). Initially, we only reduced the length of the

https://doi.org/10.1017/S1754470X25000145 Published online by Cambridge University Press

video-review to 10–15 minutes. However, a pilot study with six supervisees and three supervisors showed this did not leave enough time for comparing the observations and working on supervision goals. Therefore, in the current study we further reduced the length of the presented video-sequence to 6 minutes and reduced the number of observations to three (see Table 1). Both the VSAU and the SAU conditions were markedly less structured, leaving it to the supervisee to decide which topics were discussed in supervision. In the VSAU condition, supervisees were instructed to show at least one video-sequence from the most recent therapy session. The supervisees' instructions for each supervision method appear in Table 1.

Procedure

The supervisees selected a current patient to take part in this study. During the study, the patient's treatment was discussed in three consecutive supervision sessions. Supervision sessions were usually 4 weeks apart, with about four therapeutic sessions in between. Each supervisee was supervised by one supervisor, who, during the course of the study, conducted individual supervision. The order of the supervision conditions was randomized, and both supervisees and supervisors received written instructions before each supervision session (Table 1). All supervision sessions were audio-taped and rated for adherence by the first author.

Directly after each supervision session, supervisees evaluated the session with an online questionnaire. Two weeks after the third supervision, a follow-up assessment was conducted to collect retrospective evaluations of all three supervision conditions. Data were collected between August 2020 and September 2023. Participating supervisees were compensated with €30. The protocol was approved by the department of psychology's ethics committee (approval number: LEK-316). All supervisees, patients, and supervisors provided written informed consent.

Participants

We invited 80 supervisees to take part in this study. The supervisees were postgraduate CBT students currently treating patients under supervision as part of their practical training in a German university out-patient clinic. Overall, 27 supervisees participated, and 24 (88.9%) completed all three supervision sessions. A flowchart of attrition is included in the Supplementary material (fig. 1). Most of the supervisees were females (83.0%) aged between 26 and 49 years ($M = 32.3$, $SD = 6.1$). They had an average of 4.9 years ($SD = 5.2$) of clinical experience and had already performed 336 hours ($SD = 160$), on average, of CBT under supervision as part of their current practical training. Supervisees had previously provided an average of 20.8 ($SD = 13.6$) therapy sessions to their study patients. The sessions were supervised by the same supervisor who then provided supervision during the study. During this previous supervision process pertaining to the patient, 10 supervisees (41.7%) had used video-supervision (1–3 times). The patients were 63% female, with ages varying from 21 to 64 years ($M = 41.4$, $SD = 12.9$). Co-morbidity was high (75.0%), and depressive (62.5%) and anxiety disorders (25.0%) were the most common focus of treatment. The supervisors were 12 licensed psychotherapists (age: $M = 46.2$, $SD = 7.7$; gender: 75% female) with an average of 15.6 years ($SD = 7.2$) of clinical experience and 9.4 years ($SD = 6.8$) of supervision experience.

Measures

Supervisees' post-session *supervision satisfaction* was measured with the Session Evaluation Questionnaire (SEQ; Stiles, 1980), the Supervision Outcomes Survey (SOS; Worthen and Isakson, 2003), and the Session Questionnaire for General and Differential Individual Psychotherapy [Stundenbogen für die Allgemeine und Differentielle Einzelpsychotherapie (für die Supervision)] (STEP-SV; Zarbock *et al.*, 2009).

https://doi.org/10.1017/S1754470X25000145 Published online by Cambridge University Press

Table 1. Supervisees' instructions for the supervision conditions (abbreviated)

Instruction	Give Me 3 (GM3)	Video supervision as usual (VSAU)	Supervision as usual (SAU)
Before supervision	• Select a representative 6-min video from the most recent therapeutic session • Watch the video, note your therapeutic strengths and areas of development • Select the 3 most important observations (at least 1 strength and at least 1 area of development)	• Select one or more video sequences from the most recent therapeutic session, for which there is a supervision question • Note your supervision questions	• Reflect on the therapy process since the last supervision • Note important observations/ events/issues and your supervision questions • Do not prepare a video
During supervision	• Watch the 6-min video with your supervisor, uninterrupted. S/he will similarly note 3 observations for discussion • Compare your 3 observations with those of your supervisor	• Watch the sequence(s) with your supervisor. Use it to discuss your supervision questions	• Report your supervision questions and key aspects of the treatment process
	• Discuss what personal goals emerge from viewing and discussing the video(s) (GM3/VSAU)/ the supervision questions (SAU) • Use the rest of the supervision time to work on your goals		

The SEQ (Stiles, 1980) was originally designed to measure the impact of a therapeutic session on client and therapist but has also been used repeatedly in supervision process research (e.g. Daniel *et al.*, 2015; Strozier *et al.*, 1993; Sumerel and Borders, 1996). It originally consisted of four 5-item scales: two for the evaluation of the session (*depth*, *smoothness*) and two for the assessment of post-session mood (*positivity*, *arousal*). All items are rated on a 7-point semantic differential. The *depth* scale measures a session's perceived effect on the dimensions 'worthless–valuable', 'shallow–deep', 'empty–full', 'weak–powerful', and 'ordinary–special'. The smoothness scale assesses the perceived ease or comfort during the session (e.g. 'pleasant–unpleasant', 'rough–smooth'). The positivity scale captures the subject's feelings after the session (e.g. 'afraid–confident', 'sad–happy') and the arousal scale includes items like 'peaceful–energetic' and 'excited–calm'. In this study, we used the German version of the SEQ, the SEQ-D (Hartmann *et al.*, 2013). We omitted the *arousal* subscale because it lacked reliability in the SEQ-D validation study (Hartmann *et al.*, 2013). Our primary outcome was the *depth* scale as it assesses a supervisee's evaluations of a session's impact and can be interpreted as an indicator of perceived effectiveness (Stiles *et al.*, 1994; Strozier *et al.*, 1993). The responses to the *smoothness* and *positivity* scales indicated good reliabilities (Cronbach's alpha: *smoothness* = .84, *positivity* = .90), but responses on the *depth* scale indicated only questionable reliability (*depth* = .63).

The SOS (Worthen and Isakson, 2003) is a 20-item supervisee self-report scale designed to measure the perceived effects of supervision on a supervisee's professional development and functioning (e.g. 'My supervisor helps me feel strengthened and affirmed in my efforts to become a professional'). It additionally includes items on general supervision satisfaction as well as satisfaction with the supervisor as a role model and the supervisory relationship. Items are assessed with a 7-point Likert-type scale, assessing approval from 1 (*not at all*) to 7 (*greatest degree possible*). We translated the SOS into German, following guidelines by Guillemin *et al.* (1993). This included back-translation into English by a native-speaking bachelor's-level psychologist and the discussion of differences in a committee review by which minor adjustments were made. In this study, the responses to the translated SOS indicated excellent reliability (Cronbach's alpha = .89), which is in line with a previous application of the original scale by Reese *et al.* (2009).

The STEP-SV (Zarbock *et al.*, 2009) is a 12-item German self-report measure designed to evaluate the effect of a single supervision session. Most items focus on the supervision's perceived

https://doi.org/10.1017/S1754470X25000145 Published online by Cambridge University Press

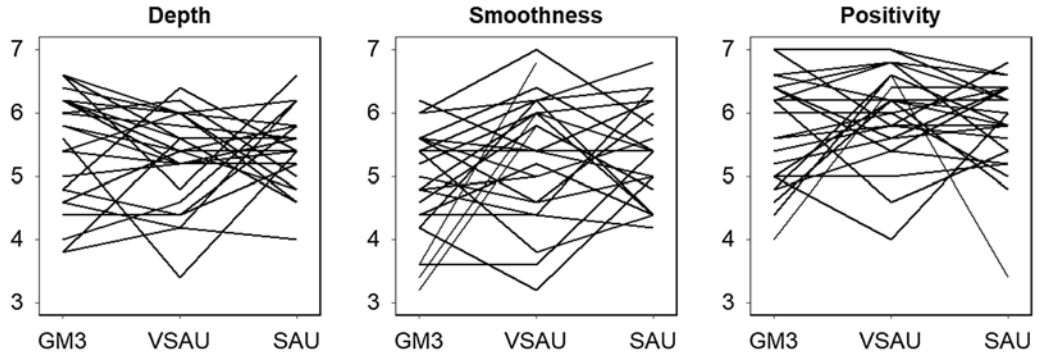


Figure 1. Spaghetti plots of supervisees' post-session SEQ ratings per supervision condition.
$n = 24$ supervisees. Each scale ranged from 1 to 7. Abbreviations of the supervision methods: GM3, Give Me 3; VSAU, video supervision as usual; SAU, supervision as usual. Order of supervision conditions varied randomly.

effect on clarification and problem-solving (e.g. improving the supervisee's case conceptualization and intervention planning), while some also cover aspects of the supervisory relationship. Items are rated on a 7-point Likert-type scale, assessing approval from 1 (*not at all*) to 7 (*greatest degree possible*). In our study, supervisees' responses on the STEP-SV indicated excellent reliability (Cronbach's alpha = .86), similar to the value reported in the original publication (Zarbock *et al.*, 2009).

The assessment of supervision satisfaction directly post-session has the advantage that the experience which underlies the supervisee's evaluation is still fresh and vivid. Conversely, this post-session evaluation is influenced not only by the supervision technique (which was systematically varied) but is also confounded by the context of supervision (e.g. whether the supervisee's current supervision goal matched the supervision provided). As we were unable to control for these external variables, we implemented a second evaluation at follow-up, in which we asked each supervisee to retrospectively evaluate the three supervision techniques (GM3, VSAU, SAU) in direct comparison. We expected this evaluation to be less influenced by the specificities of a single supervision experience and to reflect a more general view on the different supervision strategies. Supervisees retrospectively evaluated each supervision technique with four slider items (*usefulness*, *effort*, *comfort*, each ranging from 0 = 'not at all' to 100 = 'very much', and *general evaluation* ranging from 0 = 'bad' to 100 = 'good').

Data analysis

We pre-registered a multi-level analysis with a targeted sample size of 90 participants (Henrich *et al.*, 2020). However, due to unexpected sample size limitations, we instead calculated exploratory repeated-measures ANOVAs to analyse differences between the supervision conditions for each outcome. Sensitivity analyses, conducted with G*Power (Faul *et al.*, 2007; power = .80, $\alpha = .05$, $n = 24$), resulted in a minimal effect size of $f = .65$ for each analysis.

Results

Supervision satisfaction

We excluded one extreme outlier from the analyses. Repeated-measures ANOVAs revealed no significant effect of supervision condition on any measure of supervision satisfaction, including SEQ *depth*: $F_{2,44} = 0.66$, $p = .53$, $\eta^2 = .02$; SEQ *smoothness*: $F_{2,44} = 1.98$, $p = .15$, $\eta^2 = .05$; SEQ *positivity*: $F_{2,44} = 0.76$, $p = .47$, $\eta^2 = .02$; SOS: $F_{2,44} = 0.17$, $p = .84$, $\eta^2 = .00$; STEP-SV: $F_{2,44} = 0.32$, $p = .72$, $\eta^2 = .01$). Figure 1 shows large inter-individual variance in SEQ ratings for each supervision condition.

https://doi.org/10.1017/S1754470X25000145 Published online by Cambridge University Press

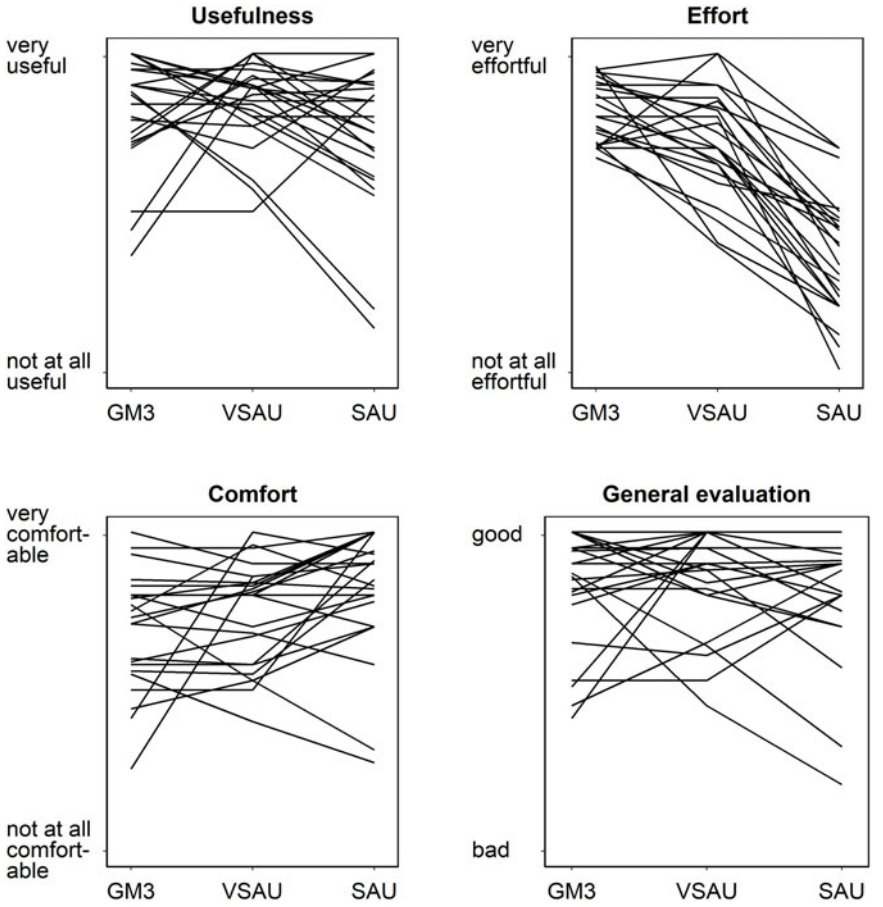


Figure 2. Spaghetti plots of supervisees' retrospective evaluation of the supervision methods. $n = 24$ supervisees. All items were rated on a slider for each supervision method, which ranged from 0 to 100. Abbreviations of the supervision methods: GM3, Give Me 3; VSAU, video supervision as usual; SAU, supervision as usual.

Retrospective evaluations of the supervision methods

Mauchly's test indicated that the assumption of sphericity was met for all dimensions, except *general evaluation*, $W = 0.56$, $p < .01$. For the latter, we applied the Greenhouse-Geiser correction ($\varepsilon = .69$). The results showed that *perceived usefulness* did not differ between SAU, VSAU, and GM3, $F_{2,44} = 1.44$, $p = .25$, $\eta^2 = .04$. Analyses revealed large differences in *perceived effort*, $F_{2,44} = 81.28$, $p < .001$, $\eta^2 = .62$, with SAU requiring significantly less effort than VSAU, $t_{22} = 9.85$, $p < .001$, and GM3, $t_{22} = 11.30$, $p < .001$. There was a non-significant trend of GM3 requiring more effort than VSAU, $t_{22} = 2.53$, $p = .06$. There was a significant but small effect of supervision method on *perceived comfort*, $F_{2,44} = 5.52$, $p < .01$, $\eta^2 = .11$. Post-hoc *t*-tests showed significantly greater comfort in SAU compared with GM3, $t_{22} = 2.74$, $p < .05$, and a similar non-significant trend compared with VSAU, $t_{22} = 2.50$, $p = .06$. GM3 and VSAU did not differ in perceived comfort, $t_{22} = 1.34$, $p = .58$. The *general evaluation* of the three supervision conditions did not differ, $F_{1.38,30.44} = 0.69$, $p = .54$, $\eta^2 = .01$. Correlational analyses indicated relationships between ratings of *comfort* and *effort* (Spearman's $\rho = -.30$, $p < .05$) and between *comfort* and *usefulness* (Spearman's $\rho = .34$, $p < .01$). There was considerable inter-individual variance in the ratings of *usefulness*, *comfort*, and *general evaluation* (see Fig. 2).

https://doi.org/10.1017/S1754470X25000145 Published online by Cambridge University Press

Conclusion

Interpretation of the results requires particular caution due to our small sample size, the limited reliability of the main outcome measure (SEQ *depth*), and the situatedness of each supervision session within a sequence of interdependent sessions. The results seem to indicate that post-session supervision satisfaction did not differ between GM3, VSAU, and SAU. Supervision was generally rated as having high depth, smoothness, and positivity. However, retrospective evaluations of the three supervision methods showed that while the supervisees judged the SAU condition as being significantly less effortful and more comfortable than GM3 or VSAU, ratings of usefulness and overall evaluation did not differ. Descriptive analyses showed large variance in the ratings of each supervision method, pointing towards individual and contextual factors that might determine supervisees' satisfaction with each method.

Interview study

Method

Procedure

We designed a semi-structured interview guide which we pre-tested with a postgraduate CBT trainee. The interviews lasted 20 to 30 minutes and were conducted in a university out-patient clinic or online via videoconference. All interviews were conducted by an advanced graduate student under supervision of the first author. Data were collected between August and October 2022. Participants received €10 in compensation. Each interview was recorded and transcribed, using the automatic speech recognition *f4x* (https://www.audiotranskription.de/en/f4x/), followed by manual correction. The protocol was approved by the department of psychology's ethics committee (approval number: LEK-428). All subjects provided written informed consent.

Participants

At the time of the study, 15 supervisees had already completed participation in the main study. Of these, nine supervisees (60.0%) participated. The participants were mostly female (66.7%) and had already performed an average of 555 hours ($SD = 142$) of CBT under 134 hours ($SD = 19$) of supervision as part of their current practical training.

Measures

The semi-structured interview guide consisted of five main questions, each with several optional follow-up questions. It started with an open recall of the supervision experiences from the main study, followed by questions regarding *expectations toward 'good' supervision* and the *perceived usefulness, degree of comfort, and the quality of personal feedback* experienced during SAU, VSAU, and GM3. Following each main question, differences between the three supervision methods were actively explored via follow-up questions.

Data analysis

The transcripts were analysd with hybrid deductive–inductive thematic analysis (Kuckartz and Rädiker, 2022). First, during initiating text work, each transcript was summarized, and important features of the text were recorded in memos. Second, the main categories were formed deductively (based on the interview guide) and inductively (derived from the initiating text work). Third, the main categories were tested on part of the material and then used to code the entire material in the first coding process. Fourth, based on the codes within each main category, subcategories were

https://doi.org/10.1017/S1754470X25000145 Published online by Cambridge University Press

developed. During the development of the subcategories, adjustments to the main category system were made. Definitions for each category were created, and the material was fully coded in a second coding process. Coding was done first by an advanced graduate student. The first author reviewed the coding and made minor adjustments to the category system. Then, the entire material was recoded and compared descriptively with the coding from the first coder. Discrepancies were found by the first author and coding was adjusted where this seemed appropriate to the first author.

We aimed for a sample size of 12 participants, as this number has been shown to be sufficient to reach saturation in rather homogeneous samples (Guest *et al*., 2006). Our final sample size was slightly smaller than hoped for but was still well above the minimum number of *six* needed to identify the most prevalent themes (Guest *et al*., 2006).

Results

The supervisees described differences in perceived comfort, effort, and usefulness between the three conditions (see Table 2 for an overview). Most supervisees reported some degree of increased nervousness during VSAU compared with SAU. However, the supervisees described discomfort as being generally low and tolerable.

Four supervisees described VSAU as requiring markedly more effort compared with SAU, which was perceived as a major drawback, like for Supervisee 6:

> Time is an important factor; preparation time, too. I have to watch it [the session recording] all over again. It just takes an insanely long time. Yes, that really is the main obstacle [. . .] really the contextual conditions, because otherwise I would actually say it [VSAU] would always be helpful. (para 69)

Interestingly, while most supervisees valued the unique advantages of VSAU, evaluations of GM3 were mixed. Some supervisees emphasized pleasant feelings like curiosity and encouragement during GM3, while others felt more discomfort, like Supervisee 9: 'I found the additional task [GM3] even more difficult. The nervousness was even higher [. . .] because I wasn't sure whether the strengths I saw were the same the supervisor saw or whether she would assess them differently' (para 71). Per usefulness, some supervisees valued that GM3 supervision had a stronger focus on the supervisee's competencies, making it more instructive for some, like for Supervisee 8:

> Completely different aspects somehow came up that I don't think would otherwise have come up in the normal reporting [SAU] or normal watching of the video [VSAU]; also, from the aspect that I thought about it beforehand, what went well, what didn't go so well, or what could go better, that you move away much more from the content of the conversation, what I do with the patients in the therapy session and talk much more about interaction or my own behavior or my reaction to some behavior of the patient. That's exactly what I found very helpful. (para 52)

However, three described the supervisor's feedback as being less authentic and credible than in VSAU. For instance, Supervisee 7 shared:

> This structured approach [GM3] always smacks of sterility to me. And if the supervisor then gives me 3 points and one of them is definitely positive and one is like this, then for me it's less serious or a bit artificial compared to an open one [VSAU]; at least I don't know whether she might set a quota for herself internally anyway, I don't know. But if it's openly disclosed that you get this kind of feedback, then it somehow has less value for me. (para 91)

https://doi.org/10.1017/S1754470X25000145 Published online by Cambridge University Press

Table 2. Comparison of subjective comfort, effort and usefulness between the conditions

Category	VSAU compared with SAU	GM3 compared with VSAU
Comfort	*Increased nervousness* during preparation for VSAU (8) Minor degree of *shame* while showing videos in VSAU (2)	*Greater discomfort* during GM3 (3) *More pleasant feelings* (curiosity, encouragement) during GM3 (2) *No difference* in nervousness/shame (1)
Effort	VSAU causing markedly *more effort* (4)	GM3 *more time consuming* during preparation and supervision (3)
Usefulness	VSAU *more objective*, less reliant on self-report (7)	GM3 with *stronger focus on therapist's competencies* (4)
	VSAU allows the *processing of additional information* (9)	GM3 helped to *identify own therapeutic strengths* (5)
	VSAU enables supervisor to gain a *personal impression* of the patient (5)	Supervisor's *feedback less authentic* and credible in GM3 (3)
	VSAU allows *better focus on the therapeutic interaction*, body language and micro sequences (5)	*Less room* for specific supervision questions in GM3 (4)
	Supervisor's *feedback more specific* in VSAU (5)	
	VSAU *facilitates therapist's self-reflection* (4)	

Number indicates how many of the $n = 9$ participants provided responses that were included in the respective category.

Also, four supervisees pointed out that there was less room for specific supervision questions to the supervisee in the GM3 condition which made it less feasible than VSAU.

Eight supervisees described conditions under which they would each use the supervision technique. Supervisees differed strongly in their opinion, with some emphasizing the general usefulness of video-supported supervision, while others saw it rather as a supplement for very specific situations.

Three supervisees characterized SAU as best suited when supervisees felt confident in therapy and had no or only very specific questions to discuss. One supervisee emphasized that an effective use of SAU demanded a strong willingness of the supervisee to disclose one's own faults and insecurities. Another supervisee saw an advanced level of therapeutic competence as a precondition for the usefulness of SAU.

In contrast, VSAU was described by six supervisees as best suited if there were specific supervision questions regarding the therapeutic relationship, problematic interactions, or diagnostics, like Supervisee 2:

> Yes, if I realized [. . .] that something was wrong; that the patient was suddenly different. And I missed it, but I think that would be something for an open one [VSAU], where I can really just pull out the sequence and say, 'Something happened, and somehow I don't know what'. (para 79)

One supervisee emphasized that VSAU would be useful at the beginning of each treatment to get the supervisor to know the patient. They also pointed out that VSAU demanded more experience than GM3, as supervisees had to be able to choose the most relevant sequences for supervision.

All supervisees indicated that GM3 should be used only selectively in supervision. Five supervisees argued it would be best suited if a supervisee had strong insecurities or self-doubt or if supervision was to focus on the development of competence. Three supervisees pointed out that GM3 should be used at times when there was nothing else to discuss in supervision (e.g. at the end of a therapeutic process), like Therapist 2:

> And the structured one [GM3], I think, is well suited for, let's say, quiet phases of therapy, where perhaps a lot of things are going according to plan in terms of method and process and there aren't such acute concerns regarding the patient. I think you can use it really well to

https://doi.org/10.1017/S1754470X25000145 Published online by Cambridge University Press

> reflect on your own therapeutic skills and simply work on them in general, which in turn benefits the other patients. (para 79)

Finally, one therapist deemed GM3 helpful to structure supervision at the beginning of psychotherapy training.

Conclusion

The results of the interview study added valuable context to the quantitative outcome of the main study. One limitation was the strong dependence on the coding of the first author due to the lack of a second concurrent rater. Results showed that the supervision strategies yielded very different outcomes, depending on the person of the therapist, the supervisor, and the therapeutic context. Supervisees emphasized the advantages of video-supported supervision but also valued the feasibility and subjective usefulness of supervision based on self-report. For some supervisees, the competency-focus of GM3 was unique and advantageous, while others experienced the technique as less revealing and overly rigid.

Discussion

The present studies added to a limited body of research on the effects of specific supervision strategies on the supervision process. This was the first project in which supervisees' satisfaction with different methods of delayed video-supervision was compared. The results of the main study suggested that supervisees' satisfaction with the effect of video-supported supervision (VSAU, GM3) did not differ significantly compared with supervision based on self-report (SAU). Ratings differed considerably between the supervisees, which suggested that contextual factors moderate supervisees' evaluations of supervision outcome. Retrospective evaluations at follow-up revealed that SAU was perceived as significantly less effortful and slightly more comfortable than VSAU and GM3. The results of the interview study suggested that supervisees generally appreciated the benefits of video-supported supervision and would mostly prefer it over SAU, if aspects of feasibility (organizational and time-related burden) were ignored.

Supervisees' evaluations of the structured video-supervision technique (GM3) differed greatly. The qualitative interviews shed some light on possible moderators of the effectiveness of GM3, such as the perceived authenticity of the supervisor's feedback, the novelty of observations gained through the sequence review, and the degree to which supervisees had other pressing supervision concerns which they were unable to discuss in GM3.

Our results were entirely based on supervisees' subjective evaluations, and we could not evaluate the objective outcomes (e.g. observer-rated therapist competence, patient outcome) of the different supervision strategies. We are therefore unable to assess whether the three supervision conditions differed in their effectiveness. However, our results complement previous research (Monson *et al.*, 2018; Weck *et al.*, 2024) which found that recording-based supervision was not significantly more effective than supervision based on self-report. This seems to be reflected in the finding that SAU was rated as useful and feasible by supervisees, while there was no significant difference in perceived usefulness between SAU and video-supported conditions. Despite the limitations of our studies, particularly the small size of our sample, we can conclude that supervisees do not always find video-supervision more useful or satisfying than SAU. In our interview study, most participants appreciated the benefits of video-supervision, but also criticized the effort involved. The results could help to understand the discrepancy between an expert consensus favouring video-supervision (e.g. Bernard and Goodyear, 2019; Haggerty and Hilsenroth, 2011), and surveys which show the infrequent use of video in regular supervision (e.g. Amerikaner and Rose, 2012; Weck *et al.*, 2017). The use of video might often be perceived as useful by supervisees, but the benefits may not always outweigh the costs. Other barriers such as

https://doi.org/10.1017/S1754470X25000145 Published online by Cambridge University Press

anxiety and shame as well as technical and organizational variables (e.g. data protection policies) and insufficient socialization might contribute to the infrequent use of video-supervision (Henrich *et al.*, 2024). Taken together, our results do, therefore, contribute to a more balanced view on video-supported supervision as a potentially highly useful but also resource-intensive supervisory tool that should be used selectively and purposively.

The GM3 technique is a supervision method aimed at enhancing therapist competence. Contrary to expectations, the supervisees did not rate GM3 higher than VSAU; in fact, ratings of GM3 were mixed. Some supervisees valued the competency focus, while others preferred the less structured, more flexible supervision in the VSAU condition. In our study, the competency focus was intertwined with the structured procedure of the GM3 condition. Some supervisees might have benefited more from a competency-based VSAU condition, which we did not implement. Developmental models of psychotherapeutic competence highlight the tension between autonomy and dependence in training and professional development (e.g. Stoltenberg and McNeill, 1997). Huhra *et al.* (2008) noted that novice therapists benefit from structured supervision, while advanced learners need more flexibility and focus on therapeutic challenges. Therapist-level factors, such as professional confidence or fear of negative evaluation, may also influence individual evaluations of supervision conditions. Based on a belief updating perspective, Herzog *et al.* (2023) suggested that supervisees who hold strong beliefs (e.g. about the therapeutic process or their own competence) may be less inclined to adjust their assumptions to new, divergent information (e.g. to a supervisor's competency-based feedback). Instead, these therapists would tend to disregard or avoid feedback. Such strong prior assumptions might lessen the effect of video-based competency feedback, as provided in the GM3 condition, and it could explain why some therapists dismissed the supervisor's feedback as superficial or irrelevant. Professional humility, on the other hand, has been found to be an inconsistent but possibly important predictor of therapist effectiveness, an effect that could be mediated by an increased effort to improve (Wampold and Owen, 2021). However, excessive self-doubt, which manifests itself in a strong fear of negative evaluation, can also lead therapists to avoid competency-based video-supported supervision. Prior assumptions about their own incompetence can increase anxiety about supervision and reduce supervisees' ability to integrate constructive feedback. These supervisees may need to be gently introduced to video-supervision, for example, by directly addressing the effects of performance anxiety, engaging in a video-based learning contract, or modelling therapeutic fallibility by reviewing a supervisor's video-recordings in supervision (Topor *et al.*, 2017).

Limitations and future directions

In this article we described the first evaluation of a structured video-supervision technique that used quantitative and qualitative data analysis. We based our analysis on postgraduate training supervision of real therapeutic processes, which lends external validity to our results. Despite these strengths, both studies had important limitations. First, our results were entirely based on supervisees' self-reports. Self-report measures of supervision satisfaction are susceptible to bias. Distributions of supervision satisfaction ratings are often left-skewed towards the positive pole (e.g. O'Donovan and Kavanagh, 2014), and demand characteristics of the measurement context might influence self-report ratings, like if participants guess the hypotheses of the respective study. Other outcomes of supervision (e.g. observer-rated competence, patient outcome) and other perspectives on the supervision process (e.g. supervisors, independent raters) would be needed for a valid and holistic assessment of different supervision methods' quality and effectiveness.

Second, interpretation of study results is limited by small sample sizes. The results of the main study in particular should be interpreted with caution due to the lack of statistical power. Sampling was hampered by a low participation rate. Only 33.8% of the supervisees who were invited participated. Selection effects (e.g. lower evaluation anxiety, greater openness to feedback, less life stress) might have impaired the generalizability of our results.

https://doi.org/10.1017/S1754470X25000145 Published online by Cambridge University Press

Finally, we were unable to control for important confounding variables due to the limited size of our sample and due to restrictions in our studies' designs. For example, descriptive analyses of rating distributions across supervisors and conditions indicated relevant variance on the supervisor level (see fig. 2 in the Supplementary material). Perhaps there might be undetected differences in supervisee satisfaction between the supervision conditions, depending on the supervisor and sources of therapist-level variance (e.g. session context, current supervision goal). Our results might have been confounded by the fact that most supervisees in our sample had used video-supervision with only a fraction of their current caseload prior to participation ($M = 28.8\%$, $Mdn = 25.4\%$). Limited prior experience with video-supervision might increase discomfort and the perceived effort of video-supervision. More experienced therapists might also be able to select and use videos more effectively, which would increase the perceived usefulness and depth of supervision. The generalizability of our results might therefore be limited to therapists with little previous use of video-supervision (e.g. postgraduate trainees). Other sources of unexplained variance could be due to uncontrolled differences in supervisors' supervision competence and supervisees' selection of video sequences.

The results of our studies suggest several avenues for future research. First, researchers should explore the effectiveness of GM3 and VSAU compared with SAU. Also, they should replicate the results with a larger sample and add objective measures of supervision effectiveness, like change in observer-rated therapist competence or patient outcome. Multiple baseline designs seem especially suited to test the effects of different forms of supervision on the therapist and the therapeutic process (e.g. Alfonsson *et al.*, 2020; Wittenborn *et al.*, 2022). Also, future studies should include qualitative interviews as an obligatory part of a two-phase fixed mixed-methods design, which would ensure greater generalizability of the results obtained.

Second, future research should identify which supervision strategies fit best for individual therapists. There is no lack of theoretical models which provide hypotheses on how different therapist variables (e.g. evaluation anxiety, reflective ability) are best addressed in supervision (e.g. Bennett-Levy, 2006; Stoltenberg and McNeill, 1997). Future studies should test these hypotheses, such as by analysing large datasets of supervision evaluations or by conducting *n*-of-1 trials in which supervisees with particular characteristics (e.g. high evaluation anxiety) receive a tailored supervision intervention.

Finally, future studies should also include the person of the supervisor. Joint research efforts are needed to collect large samples that allow multi-level analyses of the influence of supervisors on supervisee satisfaction and the effectiveness of supervision. Moreover, future investigations should explore supervisors' perspectives on different video-supported supervision strategies and when these strategies are actually chosen in supervision.

Conclusion

The present studies contribute to the limited evidence for the effects of video supervision on supervisee satisfaction. Our analyses showed that a structured form of video supervision, which was designed to maximize the positive effect of video-review, did not lead to significantly greater satisfaction compared with video-supported SAU (supervision as usual). We also found that supervisee-rated usefulness of video-supported supervision did not differ significantly from supervision based on self-report. However, our qualitative inquiries into therapists' supervision evaluations revealed that therapists saw unique advantages and disadvantages in each of the supervision strategies. Future research should further explore contexts in which different forms of video-supported supervision are most useful for therapists and supervisors.

https://doi.org/10.1017/S1754470X25000145 Published online by Cambridge University Press

Key practice points

(1) The use of video in supervision does not necessarily lead to higher supervision satisfaction.
(2) Video-supported supervision is significantly more time-consuming and potentially anxiety-inducing for supervisees, which is why it should be introduced and used in a targeted manner.
(3) In postgraduate training, supervisors and supervisees should be flexible in deciding which supervision method best suits the supervision context and the supervisee's goals.

Further reading

Alfonsson, S., Lundgren, T., & Andersson, G. (2020). Clinical supervision in cognitive behavior therapy improves therapists' competence: a single-case experimental pilot study. *Cognitive Behaviour Therapy*, *49*, 1–14. https://doi.org/10.1080/16506073.2020.1737571

Gonsalvez, C. J., Brockman, R., & Hill, H. R. M. (2016). Video feedback in CBT supervision: review and illustration of two specific techniques. *the Cognitive Behaviour Therapist*, *9*. https://doi.org/10.1017/S1754470X1500029X

Haggerty, G., & Hilsenroth, M. J. (2011). The use of video in psychotherapy supervision. *British Journal of Psychotherapy*, *27*, 193–210. https://doi.org/10.1111/j.1752-0118.2011.01232.x

Supplementary material. The supplementary material for this article can be found at https://doi.org/10.1017/S1754470X25000145

Data availability statement. The data that support the findings of this study are available on request from the corresponding author, D.H. The data are not publicly accessible, as the transcripts of the interviews could compromise the privacy of the research participants despite the removal of identifying information.

Acknowledgements. With many thanks to all the supervisees, supervisors and patients who participated in this study, and to Alexander Baumgart and Franziska Diehl for their help with data collection and management.

Author contributions. **Dominik Henrich:** Conceptualization (lead), Data curation (lead), Formal analysis (lead), Investigation (lead), Methodology (lead), Project administration (lead), Resources (lead), Visualization (lead), Writing - original draft (lead), Writing - review & editing (lead); **Julia Glombiewski:** Supervision (supporting), Writing - review & editing (supporting); **Saskia Scholten:** Conceptualization (supporting), Methodology (supporting), Supervision (lead), Writing - review & editing (supporting).

Financial support. This manuscript is based upon work supported by a grant from the Verein für Forschung und Lehre in Psychologischer Psychotherapie (WiPP e.V.). The funding body was not involved in the design of the study, collection, analysis, or interpretation of data, or in writing the manuscript.

Competing interests. Julia A. Glombiewski is head of the CBT out-patient clinic and the university training program in which the study was conducted. Dominik Henrich is employed at the psychotherapy outpatient clinic. He was involved in the planning and procurement of permanently installed video-technology in the therapy rooms. Saskia Scholten declares that she has no competing interests.

Ethical standards. All contributors abided by code of conduct as set out by the Declaration of Helsinki. The studies were approved by the Department of Psychology's ethics committee (reference numbers: 316, 428). All participants gave informed consent to participate in the study and for the results to be published.

References

Aafjes-van Doorn, K., Liu, A., & Kamsteeg, C. (2022). Video recorded treatment sessions for professional development. *Counselling Psychology Review*, *37*, 4–20. https://doi.org/10.53841/bpscpr.2022.37.1.4

Alfonsson, S., Lundgren, T., & Andersson, G. (2020). Clinical supervision in cognitive behavior therapy improves therapists' competence: a single-case experimental pilot study. *Cognitive Behaviour Therapy*, *49*, 1–14. https://doi.org/10.1080/16506073.2020.1737571

Amerikaner, M., & Rose, T. (2012). Direct observation of psychology supervisees' clinical work: a snapshot of current practice. *The Clinical Supervisor*, *31*, 61–80. https://doi.org/10.1080/07325223.2012.671721

Aveline, M. (1992). The use of audio and videotape recordings of therapy sessions in the supervision and practice of dynamic psychotherapy. *British Journal of Psychotherapy*, *8*, 347–358. https://doi.org/10.1111/j.1752-0118.1992.tb01198.x

https://doi.org/10.1017/S1754470X25000145 Published online by Cambridge University Press

Bearman, S. K., Schneiderman, R. L., & Zoloth, E. (2017). Building an evidence base for effective supervision practices: an analogue experiment of supervision to increase EBT fidelity. *Administration and Policy in Mental Health and Mental Health Services Research*, *44*, 293–307. https://doi.org/10.1007/s10488-016-0723-8

Bennett-Levy, J. (2006). Therapist skills: a cognitive model of their acquisition and refinement. *Behavioural and Cognitive Psychotherapy*, *34*, 57–78. https://doi.org/10.1017/S1352465805002420

Bernard, J. M., & Goodyear, R. K. (2019). *Fundamentals of Clinical Supervision* (6th edn). Pearson.

Brown, E., Moller, N., & Ramsey-Wade, C. (2013). Recording therapy sessions: what do clients and therapists really think? *Counselling & Psychotherapy Research*, *13*, 254–262. https://doi.org/10.1080/14733145.2013.768286

Carpenter, K. M., Cheng, W. Y., Smith, J. L., Brooks, A. C., Amrhein, P. C., Wain, R. M., & Nunes, E. V. (2012). 'Old dogs' and new skills: how clinician characteristics relate to motivational interviewing skills before, during, and after training. *Journal of Consulting and Clinical Psychology*, *80*, 560–573. https://doi.org/10.1037/a0028362

Daniel, L., Borders, L. D., & Willse, J. (2015). The role of supervisors' and supervisees' mindfulness in clinical supervision. *Counselor Education and Supervision*, *54*, 221–232. https://doi.org/10.1002/ceas.12015

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*, 175–191. https://doi.org/10.3758/bf03193146

Foa, E. B., McLean, C. P., Brown, L. A., Zang, Y., Rosenfield, D., Zandberg, L. J., Ealey, W., Hanson, B. S., Hunter, L. R., Lillard, I. J., Patterson, T. J., Rosado, J., Scott, V., Weber, C., Wise, J. E., Zamora, C. D., Mintz, J., Young-McCaughan, S., & Peterson, A. L. (2020). The effects of a prolonged exposure workshop with and without consultation on provider and patient outcomes: a randomized implementation trial. *Implementation Science*, *15*, article 1. https://doi.org/10.1186/s13012-020-01014-x

Gonsalvez, C. J., Brockman, R., & Hill, H. R. M. (2016). Video feedback in CBT supervision: review and illustration of two specific techniques. *the Cognitive Behaviour Therapist*, *9*. https://doi.org/10.1017/S1754470X1500029X

Gossman, M., & Miller, J. H. (2012). 'The third person in the room': recording the counselling interview for the purpose of counsellor training – barrier to relationship building or effective tool for professional development? *Counselling & Psychotherapy Research*, *12*, 25–34. https://doi.org/10.1080/14733145.2011.582649

Guest, G., Bunce, A., & Johnson, L. (2006). How many interviews are enough?: an experiment with data saturation and variability. *Field Methods*, *18*, 59–82. https://doi.org/10.1177/1525822X05279903

Guillemin, F., Bombardier, C., & Beaton, D. (1993). Cross-cultural adaptation of health-related quality of life measures: literature review and proposed guidelines. *Journal of Clinical Epidemiology*, *46*, 1417–1432. https://doi.org/10.1016/0895-4356(93)90142-n

Haggerty, G., & Hilsenroth, M. J. (2011). The use of video in psychotherapy supervision. *British Journal of Psychotherapy*, *27*, 193–210. https://doi.org/10.1111/j.1752-0118.2011.01232.x

Hartmann, A., Leonhart, R., Hermann, S., Joos, A., Stiles, W. B., & Zeeck, A. (2013). Die Evaluation von Therapiesitzungen durch Patienten und Therapeuten: Faktorstruktur und Interpretation des SEQ-D [The evaluation of therapy sessions by patients and therapists: factor structure and interpretation of the SEQ-D]. *Diagnostica*, *59*, 45–59. https://doi.org/10.1026/0012-1924/a000078

Henrich, D., Glombiewski, J. A., & Scholten, S. (2023). Systematic review of training in cognitive-behavioral therapy: summarizing effects, costs and techniques. *Clinical Psychology Review*, 102266. https://doi.org/10.1016/j.cpr.2023.102266

Henrich, D., Glombiewski, J. A., & Scholten, S. (2024). The use of video in supervision: is it all about social norm? *Training and Education in Professional Psychology*. https://doi.org/10.1037/tep0000491

Henrich, D., Scholten, S., & Glombiewski, J. A. (2020). *The evaluation of structured video-feedback in psychotherapy supervision.* https://aspredicted.org/blind.php?x = INF_LXE

Herzog, P., Kube, T., & Rubel, J. (2023). Why some psychotherapists benefit from feedback on treatment progress more than others: a belief updating perspective. *Clinical Psychology: Science and Practice*, *30*. https://doi.org/10.1037/cps0000174

Huhra, R. L., Yamokoski-Maynhart, C. A., & Prieto, L. R. (2008). Reviewing videotape in supervision: a developmental approach. *Journal of Counseling & Development*, *86*, 412–418. https://doi.org/10.1002/j.1556-6678.2008.tb00529.x

Kuckartz, U., & Rädiker, S. (2022). Die inhaltlich strukturierende qualitative Inhaltsanalyse [The structuring qualitative content analysis]. In *Qualitative Inhaltsanalyse. Methoden, Praxis, Computerunterstützung* (5th edn, pp. 129–155). Beltz.

Miller, W. R., Yahne, C. E., Moyers, T. B., Martinez, J., & Pirritano, M. (2004). A randomized trial of methods to help clinicians learn motivational interviewing. *Journal of Consulting and Clinical Psychology*, *72*, 1050–1062. https://doi.org/10.1037/0022-006X.72.6.1050

Monson, C. M., Shields, N., Suvak, M. K., Lane, J. E. M., Shnaider, P., Landy, M. S. H., Wagner, A. C., Sijercic, I., Masina, T., Wanklyn, S. G., & Stirman, S. W. (2018). A randomized controlled effectiveness trial of training strategies in cognitive processing therapy for posttraumatic stress disorder: impact on patient outcomes. *Behaviour Research and Therapy*, *110*, 31–40. https://doi.org/10.1016/j.brat.2018.08.007

O'Donovan, A., & Kavanagh, D. J. (2014). Measuring competence in supervisees and supervisors. In *The Wiley International Handbook of Clinical Supervision* (pp. 458–467). John Wiley & Sons. https://doi.org/10.1002/9781118846360.ch22

Reese, R. J., Usher, E. L., Bowman, D. C., Norsworthy, L. A., Halstead, J. L., Rowlands, S. R., & Chisholm, R. R. (2009). Using client feedback in psychotherapy training: an analysis of its influence on supervision and counselor self-efficacy. *Training and Education in Professional Psychology*, *3*, 157–168. https://doi.org/10.1037/a0015673

https://doi.org/10.1017/S1754470X25000145 Published online by Cambridge University Press

Routon, M. (2016). *A phenomenological inquiry of counseling trainees' experiences of reviewing video-recorded psychotherapy sessions in supervision* (Publication No. 2017-19722-091) [Doctoral Dissertation, Fielding Graduate University]. ProQuest Information & Learning.

Stiles, W. B. (1980). Measurement of the impact of psychotherapy sessions. *Journal of Consulting and Clinical Psychology*, *48*, 176–185. https://doi.org/10.1037/0022-006X.48.2.176

Stiles, W., Reynolds, S., Hardy, G., Rees, A., Barkham, M., & Shapiro, D. (1994). Evaluation and description of psychotherapy sessions by clients using the session evaluation questionnaire and the session impacts scale. *Journal of Counseling Psychology*, *41*, 175–185. https://doi.org/10.1037/0022-0167.41.2.175

Stirman, S. W., Pontoski, K., Creed, T., Xhezo, R., Evans, A. C., Beck, A. T., & Crits-Christoph, P. (2017). A non-randomized comparison of strategies for consultation in a community-academic training program to implement an evidence-based psychotherapy. *Administration and Policy in Mental Health and Mental Health Services Research*, *44*, 55–66. https://doi.org/10.1007/s10488-015-0700-7

Stoltenberg, C. D., & McNeill, B. W. (1997). Clinical supervision from a developmental perspective: research and practice. In *Handbook of Psychotherapy Supervision* (pp. 184–202). Wiley.

Strozier, A. L., Kivlighan Jr, D. M., & Thoreson, R. W. (1993). Supervisor intentions, supervisee reactions, and helpfulness: a case study of the process of supervision. *Professional Psychology: Research and Practice*, *24*, 13–19. https://doi.org/10.1037/0735-7028.24.1.13

Sumerel, M. B., & Borders, L. D. (1996). Addressing personal issues in supervision: impact of counselors' experience level on various aspects of the supervisory relationship. *Counselor Education and Supervision*, *35*, 268–286.

Topor, D. R., AhnAllen, C. G., Mulligan, E. A., & Dickey, C. C. (2017). Using video recordings of psychotherapy sessions in supervision: strategies to reduce learner anxiety. *Academic Psychiatry*, *41*, 40–43. https://doi.org/10.1007/s40596-016-0605-0

Wampold, B. E., & Owen, J. (2021). Therapist effects: history, methods, magnitude, and characteristics of effective therapists. In *Bergin and Garfield's Handbook of Psychotherapy and Behavior Change* (7th edn, pp. 297–326). John Wiley & Sons, Inc.

Weck, F., Junga, Y. M., Kliegl, R., Hahn, D., Brucker, K., & Witthöft, M. (2021). Effects of competence feedback on therapist competence and patient outcome: a randomized controlled trial. *Journal of Consulting and Clinical Psychology*, *89*, 885–897. https://doi.org/10.1037/ccp0000686

Weck, F., Kaufmann, Y. M., & Witthöft, M. (2017). Topics and techniques in clinical supervision in psychotherapy training. *the Cognitive Behaviour Therapist*, *10*, 1–17. https://doi.org/10.1017/S1754470X17000046

Weck, F., Maaß, U., Paunov, T., Heinze, P. E., & Kühne, F. (2024). Clinical supervision based on video vs. verbal report: a randomized controlled trial. *Cognitive Behaviour Therapy*. https://doi.org/10.1080/16506073.2024.2434016

Wittenborn, A. K., Subramaniam, S., Morgan, P. C., & Tseng, C. (2022). Effects of adding video feedback to emotionally focused therapy supervision: a concurrent multiple-baseline across subjects design. *Journal of Marital and Family Therapy*. https://doi.org/10.1111/jmft.12586

Worthen, V. E., & Isakson, R. L. (2003). *Enhancing supervisory relationships* [Paper presentation]. American Psychological Association, 111th Annual Conference, Toronto, ON, Canada. https://eric.ed.gov/?id = ED479919

Zarbock, G., Drews, M., Bodansky, A., & Dahme, B. (2009). The evaluation of supervision: construction of brief questionnaires for the supervisor and the supervisee. *Psychotherapy Research*, *19*, 194–204. https://doi.org/10.1080/10503300802688478

Cite this article: Henrich D, Glombiewski JA, and Scholten S (2025). Supervisees' evaluations of a structured video-supervision technique. *The Cognitive Behaviour Therapist* 1–16. https://doi.org/10.1017/S1754470X25000145

https://doi.org/10.1017/S1754470X25000145 Published online by Cambridge University Press