



Interactions with Standardized Patients to Evaluate Students' Psychotherapy-Competencies

Reliable Assessment and Valid Evaluation

Georg W. Alpers  and Kristina M. Hengen

School of Social Sciences, Department of Psychology, University of Mannheim, Germany

Abstract: The use of standardized patients (SPs) in the training of prospective practitioners is a well-established didactic tool in medical schools. Only recently have simulations of patients in psychotherapy been introduced into the training of psychologists. By integrating psychotherapy training into university-level master's programs, German law now requires licensing exams for psychotherapists (i.e., *Approbationsprüfung*) to include an assessment of therapeutic competencies in simulated interactions with SPs. Yet, it has not been examined whether these simulations are useful for a reliable assessment of competencies in psychotherapy trainees. Also, we need to develop standardized instruments to evaluate competencies in entry-level psychotherapists. As part of a university course, we trained master's-level students from three cohorts in clinical interviewing techniques (course title: *Klinisch-psychologische Gesprächsführung*). We analyzed videotaped 20-min sequences of $N = 104$ students while they interviewed one of $N = 38$ trained SPs. The students' task was to interview the SP, conduct a brief case history, and use the interviewing skills they had learned in class. Two independent raters evaluated their psychotherapeutic competencies with an adapted version of the German Cognitive Therapy Scale (CTS). Raters evaluated students' performance on two subscales and the total score with satisfactory interrater agreement (intraclass correlations). In general, students performed well in the interviews: They structured the sessions sufficiently, and their global psychotherapeutic competencies were satisfactory. However, the psychotherapeutic competencies of master's students fell short of the benchmark derived from experienced psychotherapists. This pilot study provides first evidence that simulated interviews with SPs may be a reliable tool in the assessment of practical competencies in psychotherapy trainees at an early stage of their training. Moreover, we found that the CTS, which has demonstrated validity to quantify competencies of psychotherapists, is applicable and reliable in this training context as well. In sum, this suggests that simulated interviews with SPs may be useful for evaluating psychotherapeutic competencies of psychotherapy trainees.

Keywords: standardized patients, psychotherapy training at universities, therapeutic competence, competency evaluation, German psychotherapy reform

Interviews mit standardisierten Patienten zur Bewertung der Kompetenzen von Studenten. Zuverlässige Beurteilung und valide Bewertung

Zusammenfassung: Der Einsatz von Schauspielpatienten (SP) ist in der praxisorientierten Ausbildung angehender Praktiker an den medizinischen Fakultäten gut etabliert. Erst kürzlich wurden simulierte Interaktionen mit Darstellern von Psychotherapiepatienten in der Ausbildung von Psychologen eingeführt. Das Thema ist besonders aktuell, weil das neue Approbationsstudium für Psychotherapie in Deutschland jetzt verlangt, dass die Approbationsprüfung eine Bewertung der therapeutischen Kompetenzen anhand von Interaktionen mit Schauspielpatienten beinhaltet. Bisher wurde aber nicht untersucht, ob solche Simulationen geeignet sind, um die Kompetenzen der Ausbildungskandidaten reliabel zu erfassen. Ebenfalls müssen zunächst Instrumente entwickelt werden, um Kompetenzen angehender Psychotherapeuten zu evaluieren. Als Teil eines universitären Kurses in klinisch-psychologischer Gesprächsführung unterrichteten wir Masterstudierende der Universität Mannheim. Wir analysierten Videomitschnitte von Gesprächen, die $N = 104$ Studierende aus drei Kohorten mit $N = 38$ dafür extra geschulten Schauspielpatienten führten. Die Aufgabe der Studierenden bestand darin, eine kurze Anamnese durchzuführen und dabei die im Kurs gelernten Gesprächsführungstechniken einzusetzen. Zwei voneinander unabhängige Beurteiler bewerteten die Leistung der Studierenden mithilfe der deutschen Version der Cognitive Therapy Scale (CTS). Die Beurteilung der Leistung der Studierenden ließ sich auf zwei Subskalen und dem Gesamtwert mit zufriedenstellender Übereinstimmung (intraclass correlations) bewerkstelligen. Insgesamt führten die Studierenden fachlich adäquate Gespräche mit den Schauspielpatienten; sie strukturierten die Sitzung genügend und die gezeigten psychotherapeutischen Kompetenzen waren zufriedenstellend. Allerdings zeigte sich auch, dass die psychotherapeutischen Kompetenzen der Studierenden noch klar unter dem Maßstab erfahrenerer Psychotherapeuten lag. Diese Pilotstudie legt nahe, dass sich simulierte Gespräche mit Schauspielpatienten als reliables Mittel zur Erfassung praktischer Kompetenzen in einem frühen Stadium der universitären Therapieausbildung anbieten. Darüber hinaus fanden wir, dass eine adaptierte Version der CTS-Skala, deren Validität zur Bewertung von therapeutischen Kompetenzen schon gut belegt war, ebenfalls im universitären Ausbildungskontext reliabel anwendbar ist. Zusammengefasst zeigt dies, dass simulierte Gespräche mit Schauspielpatienten nützlich sein dürften, um die Kompetenzen von Kandidaten gemäß der neuen Approbationsprüfung zu bewerten.

Schlüsselwörter: Standardisierte Patienten, Psychotherapieausbildung an Universitäten, therapeutische Kompetenz, Kompetenzbewertung, deutsche Psychotherapie reform, Parcoursprüfung

Psychotherapy training entails learning about models of psychopathology and the theory of psychological interventions. At the same time, students need to acquire practical skills to become therapists (Vogel & Alpers, 2009). German federal law recently restructured the training of psychotherapists (Der Bundestag, 2019), the most important consequence being that training is now integrated into university-based master's programs within established psychology departments. While university exams traditionally focus on the assessment of knowledge, this reform now requires the assessment of therapeutic performance as well (Vogel & Alpers, 2009). As a logical consequence, new licensing exams (*Approbationsprüfung*) require the formal assessment of students' practical skills. Highly specified licensing exams include a round of interviews of each student with trained standardized patients (SPs) who simulate a realistic scenario for a clinical interview (so-called *Parcourprüfung*). Future licensing is thus preceded by a reliable assessment of each candidate's competencies in these clinical interviews. We are unaware of similar practical exams, since previously licensing exams focused either on the acquisition of knowledge or on case formulations. Because the didactic method of SPs is relatively new to the field and formal assessments of competencies in such interviews have yet to be evaluated, we explored whether competencies can be evaluated based on interviews with SPs.

Psychotherapy training in Germany is regulated by federal law, and for the last 20 years, it has been carried out as highly structured on-the-job postgraduate training. Recently, new legislation further prescribed integrated academic training within psychology departments (Der Bundestag, 2019). Importantly, this recent development corresponds to international training models, in particular the so-called scientist-practitioner model, which is well established in the United States (see Alpers & Hofmann, 2007). By closing the gap between research and practice, this reform also addresses the criticism that current master's programs in clinical psychology are mostly theoretically oriented (Bergold, 2008; Rief et al., 2012; Schiefele & Jacob-Ebbinghaus, 2006; Wentura et al., 2013).

Federal law now defines the relevant content of the masters' programs and specifies which competencies are required to pass licensing exams after their completion. This is justified by the need to guarantee compliance with patients' safety, although graduates obtain their license much earlier than in the previous model of postgraduate training (Bundesministerium für Gesundheit, 2020). Therefore, the new master's program requires more practical training and consequentially the assessment of each graduate's therapeutic performance (Rief et al., 2014).

Several didactic approaches have been implemented to include practical skills training in the course of psychology

programs and psychotherapy training. Of course, sooner or later students should be exposed to real patients with real mental-health issues – the individuals they treat after graduation. Interacting with patients provides the most realistic impression of practical competencies as demanded to adequately perform psychotherapy. Internships and placements therefore form an integral part of any structured psychotherapy training. However, it is not always feasible to include real patients in classroom teaching because of practicability and didactic hurdles. First, enormous costs are involved for instructors and their patients; second, patients are often hesitant to participate in educational programs (see Alpers & Steiger-White, 2020).

Much easier to implement in teaching are simulations of interactions with patients. Peer-to-peer role-playing has traditionally been used to enact psychotherapy in the classroom. Indeed, role-playing has a strong track record to illustrate relevant aspects of professional interaction; it has proven to prepare the students for real-life interaction with patients with mental-health issues (Bennett-Levy et al., 2009; Fairburn & Cooper, 2011; Mitterhofer et al., 2011). In such role-playing, students alternate between the role of the patient and that of the psychotherapist. Although peer-to-peer role-playing is an important first step in practical training it suffers from specific limitations (Alpers & Steiger-White, 2020), and it is rather challenging to simulate authentic interactions. First, while students concentrate on their therapeutic role, they are typically not trained to authentically portray the patient role. Second, patient roles and their anamnestic history are usually developed ad-hoc. Third, and most importantly, peer-to-peer role-playing takes place in a familiar environment with peers of similar demographic characteristics who may know each other (in their role as a student).

More effective are external actors, who can be instructed to enact the patient role. As their roles and enactment can be well rehearsed and standardized, they are often referred to as SPs. Because they can be selected and trained to portray authentic patient roles, modules that include such SPs can offer an intermediate step between peer-to-peer role-playing and exposure to actual patients. Moreover, thanks to their high degree of standardization, they may be more suitable for exams than role-playing with other students as well as interactions with real patients.

The general approach to using SPs in training and also in testing is well established in medical schools (Barrows, 1993; Köllner et al., 2016; McNaughton et al., 2008), and this didactic method is now an integral part of medical training in Germany (Bundesministerium für Gesundheit, 2017). In medical schools, SPs are usually recruited from the general population; training procedures and role scripts have been established. This can be done so successfully

that well-trained actors with medical conditions cannot be differentiated from real patients (e.g., Ortwein et al., 2006).

SPs are much less established in psychology departments and the formal training of psychotherapists. One reason may be that it is much more difficult to enact a patient with a mental disorder than someone with, say, abdominal pain. Although we acknowledge that mental disorders are difficult to portray, we argue that the challenge posed by portraying mental disorders is one of the most important reasons for utilizing well-trained actors. While traditional curricula have heavily leaned on role-playing among peers, trained actors are likely the more realistic option. Indeed, positive experiences with this didactic method are growing, and many studies have started to include it in the current postgraduate training programs for psychotherapists (Eckel, Alpers et al., 2014; Eckel, Merod et al., 2014; Nikendei et al., 2019; Partscheffeld et al., 2013). We recently implemented a teaching module with SPs in our university's curriculum for the master's in psychology to foster teaching based on the scientist-practitioner model. Our motivation was the anticipation of the revised psychotherapy training (Alpers & Steiger-White, 2020).

From the perspective of a multilevel concept of psychotherapy training, we decided to first focus on practical competencies (Vogel & Alpers, 2009), respectively, on basic therapeutic techniques (Linden et al., 2007). As part of our class on therapeutic interviewing, students were first introduced to the theoretical models and then practiced basic interviewing skills in an anamnestic patient interview. At the end of the course, students were asked to use the knowledge they had acquired and what they had already practiced in peer-to-peer role-playing in a simulated interview with a trained SP. This module was experienced as particularly useful by the instructors and very well received by the students; course evaluations were excellent and revealed great satisfaction. Most importantly, the students' psychotherapeutic self-efficacy increased over time. Interestingly, this progress was most pronounced on those dimensions of therapeutic self-efficacy that were addressed in the class. These findings show first and promising evidence that the method of SPs is well applicable in a master's program of clinical psychology and may thus be useful for the new training format at German universities (Alpers & Steiger-White, 2020).

While there is a strong case that SPs are useful as a teaching method, it remains unclear whether interactions with SPs can be used to objectively evaluate therapeutic competencies in exams. A necessary first step in documenting its usefulness would thus be to identify measures

that allow a reliable assessment of students' competencies in simulated interactions.

The proper assessment of therapeutic competencies in psychotherapists has long been a research topic (Shaw & Dobson, 1988). From a research perspective, it is particularly important for treatment evaluation studies, where therapists' performance is typically examined concerning therapy outcome (Beidas & Kendall 2010; Dobson & Kazantzis, 2003; Kazantzis, 2003; Kuyken & Tsivrikos, 2009; Trepka et al., 2004). Previous research evaluated these skills with several heterogeneous methodological approaches, not all of which are well-validated (see Weck et al., 2010).

One of the more established tools for evaluating therapeutic competency is the Cognitive Therapy Scale (CTS; see Young & Beck, 1980, 1986). The rating scale has been used in many different settings, often in psychotherapy outcome studies, and, most importantly for our paper, it has been used to study the effects of psychotherapy training. Fortunately, there is an adapted and validated German version with very good psychometric properties (Weck et al., 2010). Although there is convincing evidence that the CTS is a useful tool to assess psychotherapeutic competencies in licensed psychotherapists (see Weck et al., 2010), it has not been well established whether the scale is also useful to evaluate performance at an earlier stage of trainees' university-level master's studies. Moreover, we are unaware of any application of the scale to evaluate interactions in simulations with SPs. However, we argue that such instruments may bear potential for the assessment in the mandatory licensing exam, which must be established for the new licensing procedure in Germany (Bundesministerium für Gesundheit, 2020).

We therefore evaluated the properties of the scale in this pilot study by coding videotaped material from our master's students' program in clinical psychology (Alpers & Steiger-White, 2020). To this end, we used the CTS to evaluate psychotherapeutic competencies in students.

Our first goal was to document the usefulness of an established scale regarding our students' performance. Only if ratings of competency can be achieved economically (practicability and time costs) would such an evaluation be reasonable for routine evaluations on a large scale. A standardized scale to evaluate therapeutic competency must be adaptable to the current setting and task. Second, we examined whether students' performance can be reliably measured on master's students. Third, as an indication of the validity of such ratings, we compared our students' scores with those of more experienced therapists (see Weck et al., 2016). Only if values within our sample had a certain range, and only if they plausibly differed from those obtained from more experienced psy-

chotherapists, would we conclude that the information we obtain is meaningful for future exams.

Taken together, this information is essential to determine whether simulations of interviews with SPs are applicable in licensing exams as now required nationwide in Germany. In the face of the growing international interest in better-structured training procedures, this may contribute to the dissemination of efficient and effective training and testing procedures for much-needed therapists.

Method

Participants

The samples of SPs and the master's students were based on the sample characterized in our previous report (Alpers & Steiger-White, 2020).

Standardized Patients and Role Scripts

A group of 39 volunteers was recruited as SPs, though one had to be excluded because of unsuitable acting performance and an inadequate display of the mental disorder to be portrayed. Thus, the analyzed videotaped sessions included 38 volunteered SPs. Most were women (71.1%), and they had a wide age range with a mean of 49.9 ($SD = 19.25$; range: 25–74).¹ None of them had been previously formally trained or professionally established as actors. Most of them were affiliated with our university but not part of the psychology program; rather, they had the status either of guest students or senior students ($N = 35$; 92.1%). Three SPs were psychology students who had a special interest in acting but had not been previously enrolled in this course. They received 8 € compensation for each session.

All SPs completed a 2-hour introductory workshop directed by an experienced clinical psychologist to learn about the specific mental disorders they were to simulate. In their training, we emphasized improvisational acting more than what would be necessary for actors who portray a patient with, for example, a broken leg. Therefore, a professional acting coach instructed the SPs with elements of improvisation skills, and they practiced their roles in a 4-hour workshop under his supervision. In addition, they were provided with detailed role scripts portraying nine highly prevalent mental disorders. The role scripts included characterizations of individuals with major depression, social anxiety, specific phobias, substance abuse

disorders, posttraumatic stress disorder, panic disorder and agoraphobia, obsessive-compulsive disorder as well as somatic stress disorder (Steiger-White & Alpers, 2020). On about 6–7 pages, the scripts provide information on the role's biography, case history, several characteristic statements that can be used in the interview, information about the disorder, typical body posture, and nonverbal signs. For a more extensive description of the role scripts, see Alpers and Steiger-White (2020).

Participants: Master Students and their Task

156 students took this class in three cohorts (2017, 2018, 2019). Thereof, $N = 122$ students consented to be videotaped while they conducted the simulated clinical assessments with the SPs. The videos of 13 students had to be excluded because all videos recorded with a particular SP turned out to be useless; 5 further videos were excluded because of poor sound quality (technical failure). Thus, a total of $N = 104$ videos was available for our evaluation.

All participating students were enrolled in our master's curriculum in clinical psychology at the University of Mannheim. In addition to our course, they had previously completed at least one internship in clinical psychology; most intended to engage in a professional carrier as a clinician.

Raters of Therapeutic Competencies

Two independent raters provided ratings on the CTS (one female and one male). One rater (KH) was an experienced psychologist with a master's degree in clinical psychology who was at an advanced stage in her psychotherapy training. The other rater was a clinically well-experienced (three clinical internships, each lasting at least 3 months) master's student (NS) who had previously participated in the course himself. The first rater (KH) instructed and trained the second rater on coding the CTS (Weck et al., 2010); after this initial calibration, the ratings were done independently. The two raters had a mean age of 27.50 ($SD = 4.95$).

Measuring Therapeutic Competency

We used the German version of the CTS (Weck et al., 2010) to assess psychotherapeutic competencies in the master's students' interviews. The CTS is a well-established measure to assess psychotherapeutic competencies necessary to administer cognitive behavioral therapy (CBT). It has been evaluated by psychologists in psychotherapy training as well as by practicing psychotherapists.

¹ Age was available for only 50% of the SPs because we did not ask for it in the beginning of data collection.

The scale consists of 14 items to evaluate the level of psychotherapeutic competencies on a 7-point rating scale (0 = *poor*, 1 = *barely adequate*, 2 = *mediocre*, 3 = *satisfactory*, 4 = *good*, 5 = *very good*, 6 = *excellent*). The 14 items cover conceptually derived specific tasks of successful CBT: (a) agenda setting, (b) dealing with problems/questions/objections, (c) clarity of communication, (d) pacing and efficient use of time, (e) interpersonal effectiveness, (f) resource activation, (g) reviewing previously set homework, (h) using feedback and summaries, (i) guided discovery, (j) focus on central cognitions and behavior, (k) rationale, (l) selecting appropriate strategies, (m) appropriate implementation of techniques, and (n) assigning homework. In addition, these subdomains can be aggregated into two subscales of psychotherapeutic competencies and a total score. Because of the better test properties, we report the averages for the subscales Session-structuring versus Global psychotherapeutic competencies. Although they do not necessarily contain the same number of items, averaging allows us to compare the subscales directly.

In a validation study of the German version, the CTS demonstrated satisfactory to good interrater reliability on the item level ($ICC(2,2) = .66-.95$), good interrater reliability on the subscale level ($ICC(2,2) = .85-.93$) as well as on the global scale ($ICC(2,2) = .90$) (Weck et al., 2010). This corresponds well with the good properties of the original CTS (Young & Beck, 1980).

Few adaptations to the scale were necessary for the present work. Because our simulated interview, in the format of an anamnestic session, simulates first contact between a therapist and a patient, we decided to exclude the items "reviewing previously set homework" and "assigning homework" from our evaluation. Moreover, we also excluded the item "selecting appropriate strategies" because interviewers in our setting had no choice as to which strategies to choose from; rather, they were all limited to using "communication and interviewing techniques." Thus, for the total CTS, we evaluated the remaining 11 items. Finally, the subscale Session-structuring competencies, included the items "agenda-setting," "pacing and efficient use of time," "guided discovery," "focus on central cognition and behavior," "rationale," and "appropriate implementation of techniques." The subscale Global psychotherapeutic competencies included the items "dealing with problems/questions/objections," "clarity of communication," "interpersonal effectiveness," "resource activation," and "using feedback and summaries." The interrater reliability of this adapted version is reported in the Results section.

Procedure

The Class

The interview to be evaluated for this study was part of a mandatory small-group seminar conducted in a block of two and a half days as part of the master's program in clinical psychology (*Gesprächsführungsseminar*). Students received course credit (two European Credit Transfer and Accumulation System points) for their participation. The general goal of the course was to teach communication and interviewing techniques and to help students to deal with difficult interviewing situations. During the seminar, there were several practical units where students worked in triads to practice their communication and interviewing skills. Figure 1 illustrates the general procedure.

The SP Interviews

On the last day of the course, each student conducted a 20-minute anamnestic interview with one of the SPs, who was trained to portray one of the roles described elsewhere (Steiger-White & Alpers, 2020). While one student interviewed the SP, another student observed the interview silently to learn from observation and later provide informal feedback. Only when both student and the SPs provided informed consent were the sessions videotaped for subsequent evaluation.

The students' task was to conduct a 20-minute interview with a patient to gain enough information on the problem, the probable diagnosis, and an overview of the case history. Importantly, the SPs were instructed to closely follow their role scripts and also simulate one of four particularly challenging behaviors, such as questioning the interviewer's therapeutic competence, excessive lamenting, suicidal tendencies, and striving for intimacy. These challenges were chosen based on conceptual considerations designed as representations of common interpersonal difficulties in psychotherapy sessions (Noyon & Heidenreich, 2013) and were categorized by one of the two independent raters.

At the end of the simulated interview, the master's students received feedback, first from the SPs, then from the observer.

Analyses

Once the general observations on the procedure and evaluation had been provided, we conducted the following statistical analyses. First, we calculated mean scores for the global scale of the CTS and the two subscales for each rater. Interrater reliability was calculated with Model 3 and a two-mixed model with an absolute agreement definition where two raters evaluated all videotaped

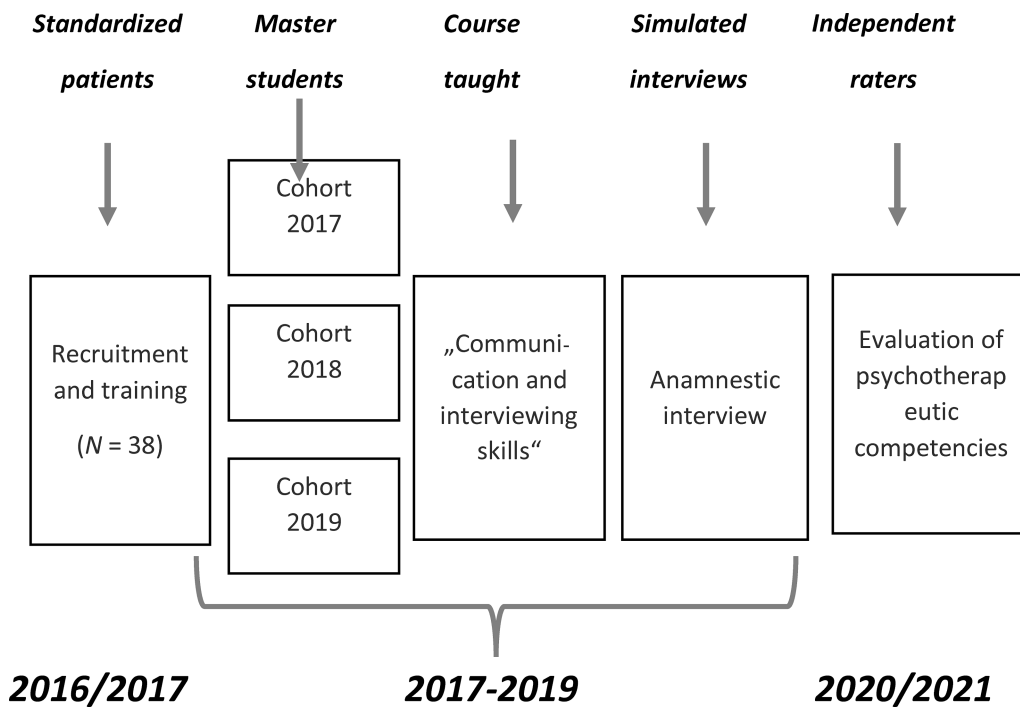


Figure 1. Timeline of the study procedure.

sessions ($ICC_{(3,2)}$) (see Shrout & Fleiss, 1979). The calculations were carried out using the 27th version of the statistical software SPSS Inc. (IBM, 2020). We used a 95% confidence interval to test for statistical significance.

Given satisfactory interrater reliability, we planned to calculate the mean scores of the ratings of the two independent raters for the following analyses. First, we calculated Cronbach's α for the total scale and the two subscales to test the internal consistency of our mean ratings. Furthermore, we analyzed the differences between the two subscales for all three cohorts to explore any potential inconsistencies after newly introducing a new didactic element. We calculated a mixed ANOVA with the between-factor Cohort and the within factor Scale. Significant effects were followed up with paired-sample t -tests.

To compare with a benchmark, we tested the global score of our students against a score obtained from published German studies of the evaluation of psychotherapeutic practitioners. For this purpose, we extracted scores of the CTS of an independent sample (Weck et al., 2014), dropping those items from their scores that were eliminated in our scales ($M = 3.88$, $SD = 1.03$). Using a t -test for two independent samples, we checked for significant differences between the psychotherapeutic competencies of the students and the more experienced practitioners.

Results

A Descriptive Evaluation of the Rating Procedure

The Efficiency of the Training Procedure

The two raters described the task as quite demanding but doable. Because both raters were previously trained, the calibration of the rating procedure required some time. The calibration included several meetings in which the raters practiced their rating procedure on psychotherapy sessions from standardized tutorial videos. Subsequently, the raters evaluate two videotaped sessions of the master's students explaining the cognitive rationale of exposure therapy and implementing an exposure session with SPs. This stepwise procedure sufficed to guarantee that the two raters were familiar enough with the videotaped material, and that there were no profound disagreements in their interpretation of the items.

Rating Procedure

It was very well possible to accomplish the evaluation of many videos in an economic fashion. Specifically, it took about 50 minutes to rate each 20-minute segment of video footage. Completing the task revealed that some segments appeared to be more diagnostically relevant than others. In particular, the raters had the impression that closely observing students during a difficult interview situation was most relevant to their overall impression.

Table 1. Means of global CTS score, the two subscales and single items – scored from 0 (*poor*) to 6 (*excellent*) and their interrater reliabilities based on 104 videotaped sessions, each rated twice

	Interrater reliabilities		
	<i>M (SD)</i>	<i>ICC</i> _(3,2)	<i>p</i> -value
Total score of the CTS	2.92 (0.64)	.89	< .001**
Subscale 1: Session-structuring competencies	2.66 (0.69)	.90	< .001**
Subscale 2: Global psychotherapeutic competencies	3.23 (0.77)	.85	< .001**
1. Agenda setting	0.98 (1.19)	.86	< .001**
2. Dealing with problems/questions/objections	3.15 (1.35)	.73	< .001**
3. Clarity of communication	4.79 (0.92)	.72	< .001**
4. Pacing and efficient use of time	2.71 (1.05)	.72	< .001**
5. Interpersonal effectiveness	3.87 (0.93)	.58	< .001**
6. Resource activation	1.25 (1.00)	.77	< .001**
7. Using feedback and summaries	3.09 (1.34)	.64	< .001**
8. Guided discovery	1.49 (1.04)	.72	< .001**
9. Focus on central cognitions and behavior	3.76 (1.09)	.75	< .001**
10. Rationale	2.85 (1.45)	.76	< .001**
11. Appropriate implementation of techniques	4.17 (1.00)	.68	< .001**

Note. Each item can vary between 0 and 6. Scores on scales are averages of their items.

Face Validity for Students' Interviews

The items on the scales of the CTS were subjectively applicable to what was theoretically being taught by the class and displayed by the students. Overall, the items had good face validity for the evaluation of the students' performance. In addition, the two raters agreed that the scale Global psychotherapeutic competencies in particular appeared to have the best face validity. More precisely, dealing with problems/questions/objections and Using feedback and summaries appeared to be the most relevant items to differentiate between students according to the raters.

Interrater Reliability of Competency Evaluation

The adapted scales internal consistency still had an acceptable Cronbach's α -value of .78 for the global score (Field, 2013, Nunnally, 1978). Table 1 shows the agreement of the evaluations of the raters who did their assessments independently from each other. All items had satisfactory to good interrater reliabilities for the ratings of the two raters ranging from .58 to .86. Thus, we were could reliably assess psychotherapeutic competencies in the master's students.

Students' Competency Scores

On a descriptive level, the students' average competency lay between a moderate to a satisfactory level of psychotherapeutic competencies at an item level (see Table 1); the two subscales and the total score were satisfactory. However, some items revealed poor performance on average, since some items may have addressed aspects to be expected in typical therapy sessions but not in the short interview task we gave our students. Importantly, the students' scores ranged broadly between 1.68–4.55 on the total score. The session-structuring competencies ranged between 1.25–4.75, and the global psychotherapeutic competencies ranged between 1.60–5.10. Apparently, there were no floor or ceiling effects. Figure 2 depicts the histogram of total scores.

In the ANOVA with the factors Scale and Cohort there was a significant main effect of Scale, $F(1, 101) = 69.45$, $p < .001$, $\eta^2 = 0.41$, but no significant effects of Cohort, $F(1, 101) = 1.65$, $p = .198$, $\eta^2 = 0.03$, and no interaction between both factors, $F(1, 101) = 2.35$, $p = .101$, $\eta^2 = 0.04$. The posthoc paired-sample t -test indicated that performance was better in global psychotherapeutic than in session-structuring competencies in all cohorts, all t -values ≥ 3.46 , all p -values $\leq .002$, all d -values ≥ 1.20 . The results of the ANOVA are illustrated in Figure 3.

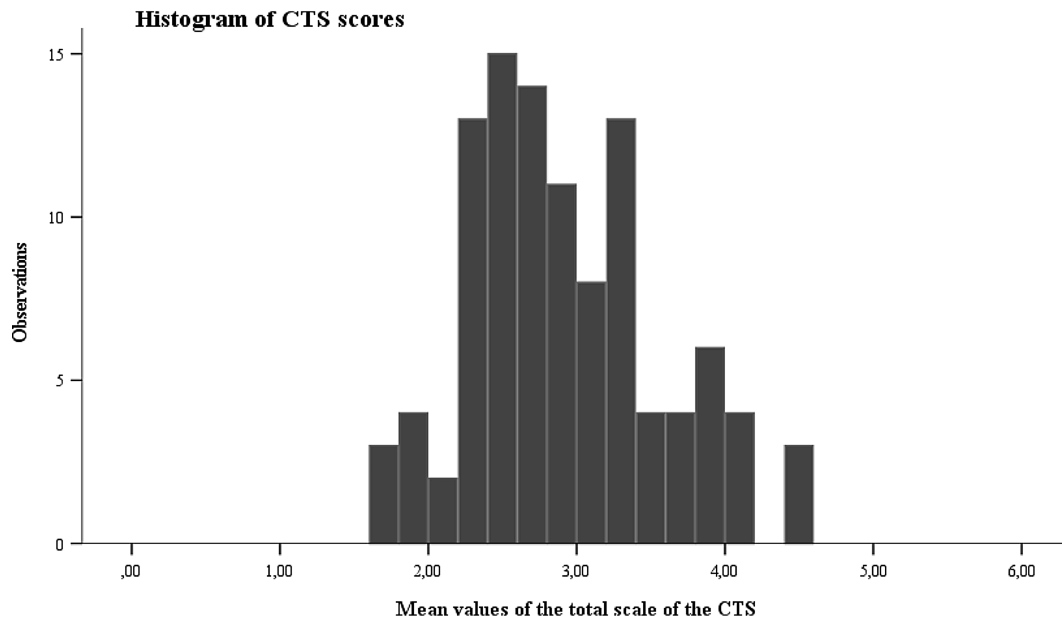


Figure 2. Histogram of CTS total scores (adapted version) in $N = 104$ averaged across two observers.

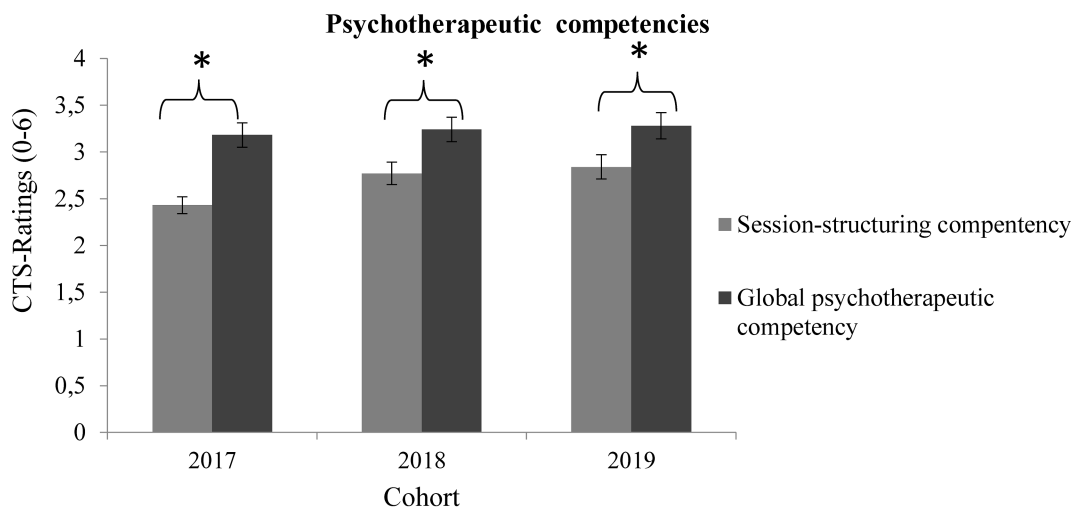


Figure 3. Means of the CTS-Ratings (adapted version) of the different cohorts of students. Error bars reflect the standard error of means. * mark significant differences.

Master Students Compared to the Benchmark

As a comparison with a benchmark of professional therapists, we calculated an t -test for independent samples to compare the students' performance with the performance of experienced practitioners on the total scale of the CTS (Weck et al., 2014). On the Global score, we found a significant difference between students' ($M = 2.92$, $SD = 0.64$) and practitioners' ($M = 3.88$, $SD = 1.03$) performance, $t(186) = -9.87$, $p < .001$, $d = -1.45$. Correspondingly, the test for the subscale Session-structuring competencies demonstrated that students' performed signi-

ficantly poorer ($M = 2.66$, $SD = 0.69$) than practitioners' ($M = 3.60$, $SD = 1.14$), $t(186) = -6.98$, $p < .001$, $d = 1.02$. Similarly, on the subscale Global psychotherapeutic competencies, students ($M = 3.23$, $SD = 0.77$) did not perform as proficiently as experienced practitioners ($M = 4.16$, $SD = 0.91$), $t(186) = -7.59$, $p < .001$, $d = -1.11$. These differences in psychotherapeutic performance between students and much more experienced practitioners support the validity of our ratings.

Discussion

Students in psychotherapy training must be instructed appropriately to acquire practical skills, and their competencies must be evaluated as part of any teaching program. This is particularly relevant regarding the implementation of the new psychotherapy training in Germany, which now requires evaluations of interviews with SPs as part of the official licensing procedure.

In agreement with the well-established scientist-practitioner model (Benjamin & Baker, 2000), the university-level curricula of clinical psychology and psychotherapy have started to focus more on combining theoretical and practical aspects (i.e., according to the scientist-practitioner model, see Alpers & Hofmann, 2007). To accomplish this, the field has started to experiment with new didactic methods to simulate realistic patient encounters. While teaching with SPs has been well established in medical schools, only recently was this method included in the curricula of psychology programs (Alpers & Steiger-White, 2020).

Importantly, German legislation now requires that SPs be part of the licensing exams (Bundesministerium für Gesundheit, 2020). However, to date, it remains unclear whether psychotherapeutic competencies can be evaluated in MSc-level university students. We recently implemented the method of SPs in our curriculum of clinical psychology and psychotherapy and evaluated 104 videotaped sessions of students and SPs simulating short anamnestic interviews with an SP.

This pilot study demonstrates that short interviews with SPs may be useful for evaluating students' therapeutic competencies. First, we identified a scale that appears to be useful for rating purposes (CTS; Weck et al., 2010) and adapted it for our purposes. Its time- and cost-effective scoring makes it possible to evaluate many observations as part of a standardized exam in routine licensing procedures.

Second, the interrater agreements we observed in numerous observations show that measurements can be obtained reliably. Only this can guarantee fair and replicable grading as part of a licensure procedure. The rating procedure is feasible, and evaluators can be trained efficiently.

Third, our data provide preliminary evidence for the validity of such evaluations. There is convincing face validity that the scales developed to evaluate professional psychotherapists can be transferred and adapted to evaluate students' performance. Moreover, a wide range of differences in students' performance can be mapped by the evaluators' ratings. Most importantly, our observation that students' proficiency appears to plausibly differ from

performance observed in much more experienced psychotherapists speaks to the validity of this assessment. At the same time, validity is limited because not all aspects to be expected in typical therapy sessions were part of the short interview task we gave our students. Obviously, the match between task and evaluation can be improved by more rigorous item selection.

Taken together, this information provides the essential first evidence that simulations of interviews with SPs may be applicable in student evaluation. This is particularly timely and relevant because such evaluations are now mandatory nationwide in Germany as part of the revised licensing procedures. Aside from these practical implications, this pilot study also provided us with several noteworthy observations.

Interestingly, our students appeared to fair better on global psychotherapeutic competencies than in arguably more basic session-structuring. This may be because the primary topic of the course focused on such global psychotherapeutic competencies (e.g., dealing with problems and interpersonal effectiveness) and less on session-structuring competencies (e.g., agenda-setting). Although we did not examine this as a formal a priori hypothesis, it suggests that our students performed better on those tasks that were explicitly part of the course, which may further indicate that the evaluation can capture relevant aspects of our instruction. Of course, future research should explicitly examine the sensitivity to improvements in performance, ideally in a pre-post design. Apart from this necessity, we found evidence for such specificity in our previous evaluation of self-reported therapeutic self-efficacy (Alpers & Steiger-White, 2020). In that study, students reported more psychotherapeutic self-efficacy for the specific skills they were actually taught. Importantly, the present data substantiate this finding with a more objective assessment of psychotherapeutic competencies by external raters.

Standardized ratings appear to be able to capture competencies, which were previously identified as important for psychotherapists (Vogel & Alpers, 2009). Future work building on these findings may identify subscales on which students' competencies are low, so that those areas may be targeted by specific instruction or practice sessions of the curriculum.

While such evaluations are not necessarily bound to the use of a specific scale we chose, we would like to discuss several observations regarding the scale's properties. It is apparently advantageous that the CTS provides a grade-related evaluation scheme. In our sample, we found a mediocre to satisfactory level of psychotherapeutic competencies in students who we obviously did not expect to perform as well as much more experienced practitioners. This finding may speak to the validity of the CTS to capture

and evaluate psychotherapeutic competencies at all levels of proficiency. As an extension to this research, future work should also examine the scale's sensitivity to change. Could they possibly capture students' progress from before to after a course?

Of course, there are several limitations to this pilot study that need to be considered. First, although both of our raters had a substantial level of clinical experience, they were not yet licensed themselves, a prerequisite for actual evaluations as they are legally defined as part of the licensing procedures. However, the agreement between the two raters was substantial and compares well with calibrated raters in other studies using the same scale (see Grikscheit et al., 2015; Weck et al., 2014, 2016). For licensing exams, which are legally binding, the scales first need to be evaluated with a larger number of much more extensively trained raters (see Strahl et al., 2019). Moreover, in a formal licensing procedure this would require continuous training and evaluation (see Strahl et al., 2018).

Second, for our purpose, we had to adapt the global scale of the CTS. However, we did take this into account in our comparisons between students and practitioners. Importantly, adapted versions of the CTS appear to be still functional. It is important to select those items (or to add to them according to the specific competencies to be examined according to one's model; see Linden et al. 2007; Vogel & Alpers, 2009). Also, the videos we rated were certainly not completely comparable to those used in the typical scenarios where the CTS is used. The students' interviews were not complete therapy sessions; they were obviously shorter and did not comprise all of the typical elements.

Third, in this pilot study, we did not completely control how the SPs acted out the difficult interpersonal challenge we instructed them to portray. Obviously, this resulted in more heterogeneous simulations compared to what would be required for formal licensing exams.

Fourth, the pairing of students and assigned SPs was not completely independent (38 SPs acted in simulated situations with 104 master's students). This might have resulted in a systematic bias for the evaluation of psychotherapeutic competencies. We did not control for subtle differences in difficulty, in pairings of sex, age, and other individual characteristics, most importantly the acting style of different SPs. Also, we were unable to examine order effects, which may be influenced by observing other models before a student had to perform themselves. While it might be possible to better standardize such boundary conditions at any one site, there will always be limits to the standardization across more than one assessment site.

Regarding the efficiency of such ratings, we acknowledge that there was a considerable cost and effort. Although the videotaped sessions lasted only 20 minutes,

the evaluation of the psychotherapeutic competencies of 104 master's students required considerable time resources. If longer exams were to be evaluated, these costs would increase linearly.

Regarding the generalization of our observations, it is important to note that conceptualizations of what constitutes "proficient" therapeutic behavior need to be considered in any evaluation. Only a limited subsection of therapeutic competencies was evaluated in this study, although there are, of course, many more aspects of professional expertise. Obviously, interviewing skills are important, albeit only one of many technical skills (Vogel & Alpers, 2009). Moreover, we did not study students' performance concerning equally important disorder-specific competencies (Linden et al., 2007), although adherence to specific manuals or a therapeutic rationale has repeatedly been proven to be relevant to outcome (e.g., Hauke et al., 2013; Weck et al., 2016). Another tradition emphasizes the importance of the therapeutic relationship (Luong et al., 2020). It remains to be evaluated whether students may vary in performance on profiles of different therapeutic skills and whether their performance results from an interaction of practical competencies with theoretical knowledge and interpersonal skills (Vogel & Alpers, 2009).

Although our teaching program was not specific to a particular therapeutic tradition, the instructors and raters were trained in CBT. Moreover, the scale we used was based on a CBT rationale and has not been extensively evaluated with therapists from other traditions. Future research is needed to explore whether similar scales can evaluate therapeutic performance rooted in other methods as well. While some basic techniques and concepts of therapeutic change clearly overlap (O'Donohue et al., 2000; Orlinsky & Howard, 1987), other aspects of therapeutic performance differ profoundly, and the specific skills to carry out basic techniques may require more specific attention (Linden et al. 2007).

Conclusion

Our observations provide a promising perspective for the evaluation of students' performance in professional interactions with SPs. Established instruments to evaluate psychotherapeutic performance can be adapted and applied efficiently and reliably for valid evaluations of student competencies.

References

- Alpers G. W. & Hofmann S. G. (2007). The American Way: Die Psychotherapieausbildung in den USA [The American Way: Psychotherapy training in the USA]. *Verhaltenstherapie und psychosoziale Praxis*, 39(1), 195–201.
- Alpers, G. W. & Steiger-White, F. (2020). Simulationspatienten im universitären Psychologiestudium: Einführung einer innovativen Methode zur Stärkung praxisorientierter Lehre in klinischer Psychologie und Psychotherapie [Standardized patients in university-level psychology programs: Introduction of an innovative method to strengthen practice-oriented teaching in clinical psychology and psychotherapy]. *Verhaltenstherapie*, 30(2), 104–116. <https://doi.org/10.1159/000507740>
- Barrows, H. S. (1993). An overview of the uses of standardized patients for teaching and evaluating clinical skills. *Academic Medicine-Philadelphia*, 68, 443–451. <https://doi.org/10.1097/00001888-199306000-00002>
- Beidas, R. S. & Kendall, P. C. (2010). Training therapists in evidence-based practice: A critical review of studies from a systems-contextual perspective. *Clinical Psychology: Science and Practice*, 17(1), 1–30. <https://doi.org/10.1111/j.1468-2850.2009.01187.x>
- Benjamin Jr., L. T. & Baker, D. B. (2000). Boulder at 50: Introduction to the section. *American Psychologist*, 55(2), 233–236. <https://doi.org/10.1037/h0087859>
- Bennett-Levy, J., McManus, F., Westling, B. E., & Fennell, M. (2009). Acquiring and refining CBT skills and competencies: Which training methods are perceived to be most effective? *Behavioural and Cognitive Psychotherapy*, 37(5), 571–583. <https://doi.org/10.1017/S1352465809990270>
- Bergold, J. B. (2008). Zurück in den Elfenbeinturm! Psychologiestudium ohne Praxis? [Back to the ivory tower! Psychology programs without application.] *Journal für Psychologie*, 16(1), 1–19. retrieved from <https://journal-fuer-psychologie.de/article/view/193>
- Bundesministerium für Gesundheit (2017). Approbationsordnung für Ärzte vom 27. Juni 2002 (BGBl. I S. 2405) (2017). Updated by article 5 of the law from July 17. *Bundesgesetzblatt I*, p. 2581, published in Bonn, Germany.
- Bundesministerium für Gesundheit (2020). Approbationsordnung für Psychotherapeutinnen und Psychotherapeuten (PschThAprO). *Bundesgesetzblatt I*, No. 11, published on March 12 in Bonn, Germany.
- Der Bundestag (2019). Gesetz zur Reform der Psychotherapeutenausbildung. *Bundesgesetzblatt I*, No. 40, published on November 22 in Bonn.
- Dobson, K. S. & Kazantzis, N. (2003). The therapist in cognitive-behavioral therapy: Introduction to a special section. *Psychotherapy Research*, 13(2), 131–134. <https://doi.org/10.1080/713869635>
- O'Donohue, W., Buchanan, J. A., & Fisher, J. E. (2000). Characteristics of empirically supported treatments. *Journal of Psychotherapy Practice and Research*, 9(2), 69–74.
- Eckel, J., Alpers, G. W., Merod, R., Vogel, H., & Neuderth, S. (2014). Einsatz Möglichkeiten von SchauspielpatientInnen in der Psychotherapieausbildung: Ergebnisse von ExpertInnenbefragungen. [Possible use of standardized patients in the training of psychotherapists: Results from a survey of experts' opinion]. *Verhaltenstherapie & Psychosoziale Praxis*, 46, 393–413.
- Eckel, J., Merod, R., Vogel, H., & Neuderth, S. (2014). Einsatz von Schauspielpatienten in den „Psych“-Fächern des Medizinstudiums – Verwendungsmöglichkeiten in der Psychotherapieausbildung? [Use of standardized patients in the psycho-social subjects of medical studies – Applicability of standardized patients in postgraduate psychotherapy Training Curricula?] *Psychother Psych Med*, 64, 5–11. <https://doi.org/10.1055/s-0033-1347210>
- Fairburn, C. G. & Cooper, Z. (2011). Therapist competence, therapy quality, and therapist training. *Behaviour Research and Therapy*, 49(6–7), 373–378. <https://doi.org/10.1016/j.brat.2011.03.005>
- Field, A. (2013). *Discovering statistics using SPSS* (4th ed.). Sage.
- Grikscheit, F., Lang, T., Kordt, A., Helbig-Lang, S., Geisler, K., & Weck, F. (2015). Entwicklung und Validierung von Instrumenten zur Überprüfung der therapeutischen Adhärenz und Kompetenz bei der Behandlung der Panikstörung [Development and validation of instruments to monitor therapeutic adherence and competence in the treatment of panic disorder]. *Zeitschrift für Klinische Psychologie und Psychotherapie*, 44, 88–96. <https://doi.org/10.1026/1616-3443/a000309>
- Hauke, C., Gloster, A. T., Gerlach, A. L., Hamm, A. O., Deckert, J., Fehm, L., Alpers, G. W., Kircher, T., Ströhle, A., Arolt, V., & Wittchen, H.-U. (2013). Therapist adherence to a treatment manual influences outcome and dropout rates: Results from a multicenter randomized clinical CBT trial for panic disorder with agoraphobia. *International Journal of Research Studies in Psychology*, 2(4), 3–16. <https://doi.org/10.5861/ijrsp.2013.335>
- IBM (2020). IBM SPSS Statistics for Windows, Version 27.0 (software). Author.
- Kazantzis, N. (2003). Therapist competence in cognitive-behavioural therapies: Review of the contemporary empirical evidence. *Behaviour Change*, 20(1), 1–12. <https://doi.org/10.1375/bech.20.1.1.24845>
- Köllner, V., Foltin, Y., Speidel, V., Müller, A. K., & Jäger, J. (2016). Anamnesegruppen als Einstieg in die Gesprächsführung und Arzt-Patient-Beziehung [Anamnestic group as an introduction to doctor-patient communication]. *Medizinische Welt*, 67, 9–13.
- Kuyken, W. & Tsvirikos, D. (2009). Therapist competence, comorbidity and cognitive-behavioral therapy for depression. *Psychotherapy and Psychosomatics*, 78(1), 42–48. <https://doi.org/10.1159/000172619>
- Linden, M., Langhoff, C., & Milew, D. (2007). Das Mehrebenen-Modell psychotherapeutischer Kompetenz [The multilevel concept of psychotherapy]. *Verhaltenstherapie*, 17(1), 52–59. <https://doi.org/10.1159/000098158>
- Luong, H. K., Drummond, S. P. A., & Norton, P. J. (2020). Elements of the therapeutic relationship in CBT for anxiety disorders: A systematic review. *Journal of Anxiety Disorders*, 76, 102322. <https://doi.org/10.1016/j.janxdis.2020.102322>
- McNaughton, N., Ravitz, P., Wadell, A., & Hodges, B. D. (2008). Psychiatric education and simulation: A review of the literature. *The Canadian Journal of Psychiatry*, 53(2), 85–93. <https://doi.org/10.1177/2F070674370805300203>
- Mitterhofer, H., Andreatta, P., & Seewald, M. (2011). Das Rollenspiel in der Psychodramausbildung im Fokus der Sozialforschung [Role-play in the training of psychodrama in social science research]. *Zeitschrift für Psychodrama und Soziometrie*, 10(1), 225–237. <https://doi.org/10.1007/s11620-011-0123-8>
- Nikendei, C., Huber, J., Ehrenthal, J. C., Herzog, W., Schauenburg, H., Schultz, J.-H., & Dinger, U. (2019). Intervention training using peer role-playing and standardised patients in psychodynamic psychotherapy trainees. *Counselling & Psychotherapy Research*, 19(4), 508–522. <https://doi.org/10.1002/capr.12232>
- Noyon, A. & Heidenreich, T. (2013). *Schwierige Situationen in Therapie und Beratung: 30 Probleme und Lösungsvorschläge* [Difficult situations in therapy and counseling: 30 problems and solutions]. Beltz.
- Nunnally, J. C. (1978). *Psychometric Theory*. McGraw-Hill.
- Orlinsky, D. E. & Howard, K. I. (1987). A generic model of psychotherapy. *Journal of Integrative & Eclectic Psychotherapy*, 6(1), 6–27.

- Ortwein, H., Fröhmel, A., & Burger, W. (2006). Einsatz von Simulationspatienten als Lehr-, Lern- und Prüfungsform [Using standardized patients in teaching and evaluation]. *Psychotherapie-Psychosomatik-Medizinische Psychologie*, 56(1), 23–29. <https://doi.org/10.1055/s-2005-867058>
- Partschfeld, E., Strauß, B., Geyer, M., & Philipp, S. (2013). Simulationspatienten in der Psychotherapieausbildung. [Standardized patients in the training of psychotherapists] *Psychotherapeut*, 58, 438–445.
- Rief, W., Abele-Brehm, A., Fydrich, T., Schneider, S., & Schulte, D. (2014). Praxisanteile im Direktstudium Psychotherapie: Auf welche Lehr-Erfahrungen kann aufgebaut werden? [Practical elements in the university training of psychotherapy: Which teaching experience can be built upon?]. *Psychotherapeutenjournal*, 1, 31–36.
- Rief, W., Schulte, D., Vogel, H., & Kuhr, A. (2012). Pro und contra Direktausbildung Psychotherapie [Pros and cons of integrated training in psychotherapy]. *Verhaltenstherapie*, 22(1), 56–63. <https://doi.org/10.1159/000336473>
- Schiefele, U. & Jacob-Ebbinghaus, L. (2006). Lernermerkmale und Lehrqualität als Bedingungen der Studienzufriedenheit [Characteristics of learning and teaching quality as determinants of satisfaction with studies]. *Zeitschrift für Pädagogische Psychologie*, 20(3), 199–212. <https://doi.org/10.1024/1010-0652.20.3.199>
- Shaw, B. F. & Dobson, K. S. (1988). Competency judgments in the training and evaluation of psychotherapists. *Journal of Consulting and Clinical Psychology*, 56(5), 666–672. <https://doi.org/10.1037/0022-006X.56.5.666>
- Shrout, P. E. & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin*, 86(2), 420–428. <https://doi.org/10.1037/0033-2909.86.2.420>
- Steiger-White F. & Alpers G. W. (2020). *Schauspielpatienten-manual für das Masterstudium klinische Psychologie und Psychotherapie* [Standardized patients manual for the masters program in clinical psychology and psychotherapy]. Universität Mannheim.
- Strahl, A., Gerlich, C., Alpers, G. W., Ehrmann, K., Gehrke, J., Müller-Garnn, A., & Vogel, H. (2018). Development and evaluation of a standardized peer-training in the context of peer review for quality assurance in work capacity evaluation. *BMC Medical Education*, 18(1), 135. <https://doi.org/10.1186/s12909-018-1233-z>
- Strahl, A., Gerlich, C., Alpers, G. W., Gehrke, J., Müller-Garnn, A., & Vogel, H. (2019). An instrument for quality assurance in work capacity evaluation: development, evaluation, and inter-rater reliability. *BMC Health Services Research*, 19, 556. <https://doi.org/10.1186/s12913-019-4387-4>
- Trepka, C., Rees, A., Shapiro, D. A., Hardy, G. E., & Barkham, M. (2004). Therapist competence and outcome of cognitive therapy for depression. *Cognitive Therapy and Research*, 28(2), 143–157. <https://doi.org/10.1023/b:cotr.0000021536.39173.66>
- Vogel, H. & Alpers, G. W. (2009). Welche Kompetenzen braucht ein Psychotherapeut? Überlegungen zu den Zielen der Psychotherapieausbildung. [Which competencies do psychotherapists need? Considerations with respect to the goals of training in psychotherapy] *Psychodynamische Psychotherapie*, 8(3), 124–134. <https://doi.org/99.120110/pdp-8-3-124>
- Weck, F., Grikscheit, F., Höfling, H., Kordt, A. Hamm, A. O., Gerlach, A. L., Alpers, G. W., Arolt, V., Kircher, T. Pauli, P., Rief, W., & Lang, T. (2016). The role of treatment delivery factors in exposure-based cognitive behavioral therapy for panic disorder with agoraphobia. *Journal of Anxiety Disorders*, 42, 10–18. <https://doi.org/10.1016/j.janxdis.2016.05.007>
- Weck, F., Grikscheit, F., Höfling, V., & Stangier, U. (2014). Assessing treatment integrity in cognitive-behavioral therapy: Comparing session segments with entire sessions. *Behavior Therapy*, 45, 541–552. <https://doi.org/10.1016/j.beth.2014.03.003>
- Weck, F., Hautzinger, M., Heidenreich, T. & Stangier, U. (2010). Erfassung psychotherapeutischer Kompetenzen: Validierung einer deutschsprachigen Version der Cognitive Therapy Scale [Assessment of therapists' competencies: Validation of a German version of the Cognitive Therapy Scale]. *Zeitschrift für Klinische Psychologie und Psychotherapie*, 29(4), 244–250. <https://doi.org/10.1026/1616-3443/a000055>
- Wentura, D., Ziegler, M., Scheuer, A., Bölte, J., Rammsayer, T., & Salewski, C. (2013). Bundesweite Befragung der Absolventinnen und Absolventen des Jahres 2011 im Studiengang BSc Psychologie [Nationwide assessment of graduates of the year 2011 of BSc psychology programs]. *Psychologische Rundschau*, 64(2), 103–112. <https://doi.org/10.1026/0033-3042/a000158>
- Young, J. & Beck, A. T. (1980). *Cognitive Therapy Scale rating manual*. Unpublished manuscript, Center of Cognitive Therapy, Philadelphia, PA.
- Young, J. & Beck, A. T. (1986). Cognitive therapy scale: Psychometric properties. *Journal of Consulting and Clinical Psychology*, 54(3), 381–385. <https://doi.org/10.1037/0022-006X.54.3.381>

Acknowledgments


We would like to thank those who volunteered to take on roles as standardized patients, the course instructors from our team, and of course our students at the University of Mannheim. Thank you to Dr. Frauke Steiger-White for developing the curriculum with us. Also, we thank Julia Derkau for helping to secure funding for the project.

Funding

Open access publication enabled by University of Mannheim. Prof. Alpers received funding by two grants from the HAREBE consortium (Handlungswissen, Reflexion und Berufsorientierung in der universitären Lehre) and RE2, both sponsored by the Ministry for Science and Culture of the State of Baden-Wuerttemberg (Ministerium für Wissenschaft und Kultur des Landes Baden-Württemberg).

ORCID

Univ.-Prof. Dr. Georg W. Alpers

 <https://orcid.org/0000-0001-9896-5158>

Univ.-Prof. Dr. Georg W. Alpers

Lehrstuhl Klinische und Biologische Psychologie und Psychotherapie
 Fakultät für Sozialwissenschaften
 Department of Psychology
 School of Social Sciences
 University of Mannheim
 L13, 17
 68131 Mannheim
 alpers@uni-mannheim.de