Original Paper

A Blended Learning Course on the Diagnostics of Mental Disorders: Multicenter Cluster Randomized Noninferiority Trial

Gabriel Bonnin[1,2*], MSc; Svea Kröber[1,2*], PhD; Silvia Schneider[1,2], PhD; Jürgen Margraf[1,2], PhD; Verena Pflug[1,2], PhD; Alexander L Gerlach[3], PhD; Timo Slotta[3], PhD; Hanna Christiansen[2,4], PhD; Björn Albrecht[2,4], PhD; Mira-Lynn Chavanon[2,4], PhD; Gerrit Hirschfeld[5], PhD; Tina In-Albon[6], PhD; Meinald T Thielsch[7], PhD; Ruth von Brachel[1,2], PhD

[1]Mental Health Research and Treatment Center (FBZ), Faculty of Psychology, Ruhr University Bochum, Bochum, Germany

[2]German Center for Mental Health (DZPG), Bochum-Marburg, Germany

[3]Clinical Psychology and Psychotherapy, Department of Psychology, University of Cologne, Cologne, Germany

[4]Clinical Child and Adolescent Psychology, Department of Psychology, Philipps University Marburg, Marburg, Germany

[5]Faculty of Business, University of Applied Sciences Bielefeld, Bielefeld, Germany

[6]Clinical Child and Adolescent Psychology and Psychotherapy, Department of Psychology, University of Mannheim, Mannheim, Germany

[7]Work and Environmental Psychology, Department of Psychology, University of Wuppertal, Wuppertal, Germany

[*]these authors contributed equally

Corresponding Author:
Gabriel Bonnin, MSc
Mental Health Research and Treatment Center (FBZ)
Faculty of Psychology
Ruhr University Bochum
Massenbergstraße 9-13
Bochum, 44787
Germany
Phone: 49 (0)234 32 217
Email: gabriel.bonnin@rub.de

Abstract

Background: Clinical diagnoses determine if and how therapists treat their patients. As misdiagnoses can have severe adverse effects, disseminating evidence-based diagnostic skills into clinical practice is highly important.

Objective: This study aimed to develop and evaluate a blended learning course in a multicenter cluster randomized controlled trial.

Methods: Undergraduate psychology students (N=350) enrolled in 18 university courses at 3 universities. The courses were randomly assigned to blended learning or traditional synchronous teaching. The primary outcome was the participants' performances in a clinical diagnostic interview after the courses. The secondary outcomes were diagnostic knowledge and participants' reactions to the courses. All outcomes were analyzed on the individual participant level using noninferiority testing.

Results: Compared with the synchronous course (74.6% pass rate), participation in the blended learning course (89% pass rate) increased the likelihood of successfully passing the behavioral test (odds ratio 2.77, 95% CI 1.55-5.13), indicating not only noninferiority but superiority of the blended learning course. Furthermore, superiority of the blended learning over the synchronous course could be found regarding diagnostic knowledge (β=.13, 95% CI 0.01-0.26), course clarity (β=.40, 95% CI 0.27-0.53), course structure (β=.18, 95% CI 0.04-0.32), and informativeness (β=.19, 95% CI 0.06-0.32).

Conclusions: Blended learning can help to improve the diagnostic skills and knowledge of (future) clinicians and thus make an important contribution to improving mental health care.

Trial Registration: ClinicalTrials.gov NCT05294094; https://clinicaltrials.gov/study/NCT05294094

Introduction

The reliable and valid diagnosis of mental disorders is associated with a more favorable therapeutic course and outcome [1,2]. However, although structured clinical interviews are acknowledged as the gold standard for diagnosing mental disorders [3-5], clinicians often rely on unstructured, experience-based explorations of symptoms [6,7]. As a result, both the under- and overdiagnosis of mental disorders are common [8-10], leading to undertreatment [11] or inappropriate or unnecessary psychotherapy or medication [12-14]. In view of the high rates of misdiagnoses, there is an urgent need to improve diagnostics of mental disorders by disseminating evidence-based assessment procedures into clinical practice. While there is increasing awareness of the importance of disseminating evidence-based treatment [15,16], the foundation of successful treatment, namely evidence-based diagnostics, has not been sufficiently addressed in dissemination research [17,18]. Therefore, the aim of this study was to develop and evaluate a blended learning course to disseminate evidence-based diagnostics of mental disorders.

Learning environments can be classified into 4 groups based on their modality (offline vs online) and synchronicity (synchronous vs asynchronous). Traditional classroom teaching combines synchronous with offline teaching, while webinars are an example of synchronous online teaching. Asynchronous teaching can be realized online through learning management systems or offline with the use of printed learning materials. Recent meta-analytical evidence indicates that there are no significant differences in learning outcomes between synchronous offline and online learning [19,20], as well as asynchronous online instruction [20]. In addition, there are various combinations of these learning environments, including blended learning, which combines (synchronous) offline and (asynchronous) online learning [21]. This approach addresses the limitations of both traditional offline and online learning, such as reduced student engagement and inconvenient time and space requirements [22]. By combining the best of both approaches, blended learning provides a more effective learning experience. Recent meta-analytical evidence suggests that blended learning is consistently more effective than either synchronous online or offline learning alone and is better accepted than traditional offline teaching [22-24].

In addition, a blended learning approach may be especially appropriate for disseminating evidence-based diagnostics, as it addresses challenges commonly associated with its training and dissemination in a face-to-face setting [25]. First, time and costs that are required for training in evidence-based methods act as important barriers for attendance [26]. Web-based or blended training methods can overcome this barrier by allowing clinical knowledge and skills to be trained in a time- and cost-efficient, easily accessible, flexible, and highly standardized way [27-29]. Second, viewing case studies of and practicing diagnostic situations are essential for acquiring diagnostic skills. As videos of diagnostic situations with simulated patients can be included in the asynchronous online part, whereas practicing in role plays can occur during the synchronous face-to-face part of the course, a blended learning approach allows learning practical diagnostic skills in a flexible and time-saving way. Third, clinicians underestimate patient acceptance of structured interviews and seem to have various preconceptions against their use [6]. These preconceptions can be reduced by addressing them explicitly and by intensifying training in the implementation of structured interviews [18,26,30]. By conveying content in a highly accessible and standardized way, blended learning courses can contribute to the intensification and standardization of training in structured diagnostic interviews and hereby reduce prejudices against their use.

Although the use and acceptance of online teaching methods increased globally during the COVID-19 pandemic [31,32], until today, only a few studies evaluated blended learning in randomized controlled trials [33-36]. Despite the high relevance of disseminating evidence-based diagnostics into clinical practice, to our knowledge, blended learning for teaching diagnostic skills was not yet evaluated at all. We aimed to fill this gap by conducting a cluster randomized controlled trial at 3 German universities and comparing a blended learning course with regular face-to-face teaching in a noninferiority analysis at the individual participant level. As there is evidence that the impact of training on more experienced practitioners does not last over time [37,38], and it is considered that such training may be more effective for those at the beginning of their clinical careers [39], we targeted a relatively inexperienced sample of preprofessionals, specifically undergraduate psychology students.

We hypothesize that students' interviewing skills, knowledge acquisition, and reactions in a novel blended learning course will be noninferior to those in traditional synchronous courses.

Methods

Study Design

The study was a multicenter cluster randomized controlled trial, comparing 2 university teaching formats: a blended learning course and a traditional synchronous course. A cluster randomization of courses was chosen because individual randomization of participants was not feasible given the constraints of the existing university setting. Clusters were 18 courses in clinical diagnostics at the 3 cooperating universities. Courses were randomly assigned to 1 of the 2 teaching conditions, stratified by study site. Participants could choose between courses in the online registration systems of the respective universities. To minimize any selection bias, course information available to participants (eg, content and instructor) was held constant in both conditions. Importantly, participants had no information about whether the teaching condition was synchronous or blended. Since the study was conducted at 3 different universities with different numbers of students and

teachers, the number and size of courses at each center varied. There were 3 assessments—before the start of the courses (t1), before the last course session (t2), and after the last course session (t3).

While both teachers and participants were aware of the teaching condition to which they were assigned, key aspects of the assessment process were conducted under blinded conditions. Specifically, the actors portraying patients in the diagnostic skills test were unaware of the participants' assigned teaching conditions to ensure impartiality in their interactions. In addition, the outcome assessors responsible for scoring the videotaped simulated diagnostic scenarios were also blinded to the participants' group assignments, ensuring that the assessment of diagnostic skills was not influenced by knowledge of the training method.

The trial is registered on ClinicalTrials.gov (NCT05294094). Due to governmental protective restrictions in Germany related to the COVID-19 pandemic, synchronous sessions that were originally planned as face-to-face classes had to be conducted as online webinars. This deviation from the registered study protocol is addressed in the *Discussion* section of this paper, where the potential consequences and limitations are analyzed in detail.

In accordance with the journal guidelines for reporting randomized controlled trials of eHealth interventions, the CONSORT-EHEALTH checklist is included in Multimedia Appendix 1.

Participants

A total of 3 universities took part in the study (Ruhr University Bochum, University of Cologne, and Philipps University Marburg). The eligibility criterion for clusters was an undergraduate course on clinical diagnostics at cooperating universities. Eligibility criteria for individual participants were (1) age >18 years, (2) undergraduate psychology students at a cooperating university, and (3) willingness to give informed consent online.

Participation in a course on the diagnostics of mental disorders was mandatory in the curriculum of the undergraduate psychology program at all cooperating universities. In total, 18 courses were offered, 10 of which focused on the diagnostics of mental disorders in adulthood and 8 of which focused on the diagnostics of mental disorders in childhood and adolescence. Participants were recruited over the course of 2 semesters between April 2021 and February 2022. During this period, the courses were attended by 400 students. Participation was possible for all students at each of the 3 measurement time points separately.

Procedure

Overview

Before the start of the course, written informed consent was obtained. Study participation was voluntary and compensated with a test participant certificate (mandatory part of the study program) and a shopping voucher (€10-20 [US $10.84-21.68], value depending on the scope of study participation). While the synchronous online classes were held weekly from the beginning, the blended learning course started 6 weeks into the current semester for organizational reasons. The asynchronous online part of the blended learning course was accessible to all students through the Moodle learning platform.

Experimental Condition: Blended Learning Course

Overview

The blended learning course followed a flipped classroom model, in which asynchronous online lessons focused on content delivery and synchronous online sessions were used to apply and deepen clinical skills under the guidance of an instructor [40]. The course consisted of 8 asynchronous online lessons and 3 synchronous online sessions and was designed considering the current knowledge regarding the conditions under which blended learning is effective (eg, including case studies, interactive elements with personalized feedback, or collaborative activities during synchronous sessions [32,41]) and well accepted by students (eg, user-friendly and functional design [42]).

A detailed overview of the course content is illustrated in Table 1. Access to the asynchronous online course can be provided by the corresponding author (GB) on request.

Table 1. Structure and content of the blended learning course.

Part and lesson	Adulthood	Childhood and adolescence
Part I: Asynchronous lessons		
Lessons 1-3: Diagnostic fundamentals and evidence-based assessment	1. Introduction to classificatory diagnostics, diagnostic approaches, and classification systems	1-3. Same as adulthood
	2. The diagnostic process, standardized clinical assessment, and biasing influences on the diagnostic process	
	3. Structure, conduction, and development of the (Kinder-)DIPS-OA[a]	
Lessons 4-7: Diagnostic criteria and conduction of the (Kinder-)DIPS-OA for specific disorders	4. Panic disorder, agoraphobia, social anxiety disorder, and generalized anxiety disorder	4. ADHD[b], oppositional defiant disorder, and conduct disorder
	5. Bipolar disorders, major depression, persistent depressive disorder, and OCD[c]	5. Separation anxiety disorder, specific phobia, and social anxiety disorder
	6. PTSD[d], somatic symptoms disorder, and illness anxiety disorder	6. Generalized anxiety disorder, selective mutism, and major depression
	7. Anorexia nervosa, bulimia nervosa, and alcohol use disorder	7. PTSD, OCD, and anorexia nervosa
Lesson 8: Evaluation of the (Kinder-)DIPS-OA	8. Evaluation of the (Kinder-)DIPS-OA, feedback of a diagnosis, difficult situations conducting the (Kinder-)DIPS-OA, and acceptance and psychometric properties of the (Kinder-)DIPS-OA	8. Same as adulthood
Part II: Synchronous Sessions		
Lessons 9-10	9. Apply skills and conduct the (Kinder-)DIPS-OA as the interviewer and as a patient with fellow students. Get direct feedback from peers and teacher.	9-10. Same as adulthood
	10. Other nonspecified content was based on the students' questions and interests (eg, questions regarding the diagnostic criteria, the diagnostic process, and how to conduct the [Kinder-]DIPS).	

[a](Kinder-)DIPS-OA: Diagnostic Interview for Mental Disorders (in Children and Adolescents) – Open Access.

[b]ADHD: attention-deficit/hyperactivity disorder.

[c]OCD: obsessive compulsive disorder.

[d]PTSD: posttraumatic stress disorder.

Blended Learning Course—Asynchronous Lessons

In total, 2 separate versions of the asynchronous online course were developed—a version with a focus on the diagnostics of mental disorders in childhood and adolescence and a version with a focus on the diagnostics of mental disorders in adulthood. Both versions were parallel in content, except for age-specific diagnostic procedures and some of the disorders presented, as they typically occur at different developmental stages (Table 1). Furthermore, both versions focused on teaching the conduction of a semistructured diagnostic interview, the Diagnostic Interview for Mental Disorders – Open Access 1.2 (DIPS-OA1.2) [43], and the Diagnostic Interview for Mental Disorders in Children and Adolescents – Open Access (Kinder-DIPS-OA) [44].

Content, usability, and design of the online courses were formatively evaluated during development by students and research associates from the Ruhr University Bochum and the University of Koblenz-Landau.

Each lesson included an introduction and conclusion sequence, a downloadable handout, and final evaluation questions. In the disorder-specific lessons (4-7), video-based case studies (played by actors) were presented to illustrate the conduction of a structured interview and to allow participants to test and apply their acquired knowledge through interactive elements (eg, multiple choice questions, automatic feedback, and matching tasks). Participants were able to navigate through the lessons and subchapters independently; however, working through the course content in sequential order was recommended. A tutorial video was provided, explaining how to navigate through the course, as well as how to use the various interactive course elements. In addition, the course included a forum where participants could ask questions about the course content, which were answered by the first 2 study authors (GB and SK).

Blended Learning Course—Synchronous Sessions

Following the asynchronous online course, participants of the experimental condition took part in 3 weekly synchronous online sessions (90 min each). In these sessions, they could discuss questions about the asynchronous course content with a lecturer and apply their skills in role plays with the other participants.

Control Condition: Synchronous University Course

The synchronous university course took place in attendance and consisted of 11 weekly online sessions (90 min each), representing the usual teaching of clinical diagnostic knowledge and skills at the 3 cooperating universities. The teachers were instructed to work through mandatory content, which was based on the asynchronous online course to ensure comparability between the 2 conditions. Before the start of the course, a training session was held for the teachers. In addition, course material was provided in the form of Microsoft PowerPoint slides. In addition to the mandatory content, teachers were

allowed to provide additional information relevant to the field of clinical diagnostics.

Measures and Assessments

Primary Outcome: Practical Diagnostic Skills

The primary outcome was the students' performance in a simulated structured diagnostic interview. At t2, course participants individually conducted a 15-minute section of a structured clinical interview (Diagnostic Interview for Mental Disorders [in Children and Adolescents] - Open Access; [Kinder-]DIPS-OA) with patients played by previously trained actors through video chat. All actors were blinded to the assigned teaching condition. Patient roles were based on 1 out of 3 case vignettes distributed evenly across courses, each for a different disorder (generalized anxiety disorder, obsessive-compulsive disorder, or major depression; "Section A" in Multimedia Appendix 2). Each case vignette included instructions to the actors to simulate difficult interview situations (eg, "Miss the point with your answer to this question."; "Section A" in Multimedia Appendix 2). The interviews were videotaped and then rated by 4 blinded and independent evaluators using a coding scheme ("Section B" in Multimedia Appendix 2), which assessed 2 facets of interview performance—formal interviewing skills (10 items; eg, "The interviewer asks relevant additional questions beyond the interview guide to assess the presence of the diagnostic criteria.") and interpersonal interviewing skills (9 items; eg, "The interviewer uses non-verbal and paraverbal interviewing techniques."). Both dimensions were assessed on scales ranging from 0 to 100. To succeed in adequately conducting the structured interview, participants had to score at least 50% correct on both scales. The cutoff of 50% is commonly used in the German education system.

All the outcome assessors had a master's degree in psychology, were certified and experienced conducting the (Kinder-)DIPS-OA, and received at least 2 years of postgraduate cognitive behavioral therapy training. Interrater reliability for each item was calculated based on 40 jointly coded interviews, with Fleiss κ ranging between fair (0.34) and almost perfect (0.96) agreement between outcome assessors [45].

Secondary Outcomes

Diagnostic Knowledge

Two parallel 15-item versions of a test of basic clinical diagnostic knowledge were created, which participants answered at t1 and t3 (refer to "Section C" in Multimedia Appendix 2 for example items). The format of the items varied (single choice, multiple select, and multiple-true-false) and the items were previously piloted with laypersons (30 undergraduate students in their first semester) and experts (44 therapists in postgraduate training). Items were selected based on item-scale correlation and discrimination between these 2 groups. In addition, at t1, the self-reported diagnostic knowledge was assessed on an 11-point Likert-type scale ("How knowledgeable are you in the area of 'clinical diagnostics'?"; 0= "I don't know anything about it", 10= "I am very knowledgeable in this area").

Participants' Reactions

Participants' reactions to the courses and the estimated patient acceptance of structured interviews were evaluated at t3 by means of an online questionnaire, which consisted of 32 selected items (Table 2) from several instruments [6,46-50]. There were 8 additional items only administered in the blended learning condition. Unless otherwise described, a 7-point Likert-type scale was used for the items, ranging from 1="strongly disagree" to 7="strongly agree", with higher scores indicating a better outcome.

Table 2. Overview of the items assessing participants' reactions.

Questionnaire and subscale	Items, n	Example	Cronbach α
MFE-Sr[a] [46]			
Intent to recommend	1	"I would recommend this course to other students."	__[b]
Experience of overload	3	"The content of this course was too difficult for me."[c]	0.77
Subjective learning success	1	"I learned a lot in this course."	—
Web-CLIC[d] [46]			
Clarity	3	"The contents of the course are clearly presented"	0.83
Likeability	3	"The course arouses my interest"	0.93
Informativeness	3	"The information is of high quality"	0.91
Credibility	3	"I can trust the information in the course"	0.95
Short scale for academic course evaluation [47]			
Course structure	3	"The course was clearly structured."	0.73
UMUX-Lite[e] [48]			
Usability	2	"This system is easy to use"[f]	0.82-0.83
VisAWI-S[g] [49]			
Visual Aesthetics	4	"The layout is professional"[f]	0.76
Items designed by the study authors			
Visual Aesthetics	1	"DiSkO is designed to be visually appealing"[f]	—
Credibility	1	"I completely trusted the content in DiSkO"[f]	—
Overall impression	1	"Overall: I give the course an overall grade of …"[c,h]	—
Acceptance of structured interviews questionnaire [6]			
Global satisfaction rating	1	"Please indicate on the accompanying scale how satisfied you think patients are or would be with structured diagnostic interviews in general."[i]	—
Mental effort and emotional reaction to structured interviews	10	"After a structured interview, patients feel more confused than before."[j]	—

[a]MFE-Sr: Münster Questionnaire for the Evaluation of Seminars – revised.

[b]Not applicable

[c]Lower scores indicate a better outcome.

[d]Web-CLIC: Website-Clarity, Likeability, Informativeness, and Credibility.

[e]UMUX-Lite: Usability Metric for User Experience – Lite.

[f]These items were only administered in the blended learning condition.

[g]VisAWI-S: Visual Aesthetics of Websites Inventory – Short.

[h]This item used the German grading system ranging from 1 (excellent) to 6 (insufficient).

[i]Visual analog scale ranging from 0 (not at all satisfied) to 100 (completely satisfied).

[j]4-point Likert scale ranging from 0 (disagree) to 3 (completely agree).

Statistical Analyses

All outcomes were evaluated at the individual participant level using noninferiority analyses. To assess noninferiority between the teaching conditions, 95% CIs were calculated. Although our hypothesis regarding the noninferiority of the blended learning course is inherently 1-sided, the use of a 2-sided 95% CI is standard practice [51] and recommended by the guidelines of the European Medicines Agency and the US Food and Drug Administration [52,53]. This approach allows for the interpretation of noninferiority if the entire 95% CI lies above the prespecified noninferiority margin.

A logistic regression was conducted, predicting the primary outcome (passing the performance test) based on the predictor's teaching condition (blended learning vs synchronous), study site (center 1 vs center 2 vs center 3), course focus (adulthood vs childhood and adolescence), study year, self-reported diagnostic knowledge, and the score in the knowledge test at t1. To account for the effects of cluster randomization, the

course variable was included as a random effect (random-intercept). Odds ratios (ORs) were calculated by unconditional maximum likelihood estimation and 95% CI using normal approximation. Based on experience with the traditional synchronous course format in diagnostic teaching, a passing rate of 85% was assumed for the synchronous course. As the passing rate after blended learning should be at least as good as that in traditional face-to-face instruction due to the positive effects of blended learning on learning outcomes [24], a 90% passing rate was assumed in the blended learning course. To test for noninferiority, the assumed passing rates and noninferiority margin of 5% were transferred to ORs [54,55]:

$$\frac{\frac{80}{20}}{\frac{85}{15}} = .71.$$

Accordingly, a power analysis for noninferiority trials with dichotomous data revealed for expected success rates of 85% and 90%, respectively; noninferiority margin of 5%; $\alpha=.05$; and power of 80% a required sample size of 135 per treatment group [56].

All secondary outcome measures were tested using multiple linear regression models with the following predictors: teaching condition, study site, course focus, study year, self-reported diagnostic knowledge, and the score in the knowledge test at t1. Noninferiority of the blended learning course was assumed, when the lower bound of the CI of the predictor teaching condition was larger than $\beta=-.10$, corresponding to a small negative effect.

To test for systematic differences between teaching conditions at baseline, *t* tests and Fisher exact tests were conducted. Furthermore, we tested for differences in assigned teaching condition and practical diagnostic test performance between completers (participation in t1, t2, and t3) and noncompleters using chi-square tests.

All available data were analyzed for each statistical test performed. All analyses were run in R (R Core Team) [57]. The anonymized dataset [58] and R code [59] are available online.

Ethical Considerations

The authors declare that all procedures contributing to this work comply with the ethical standards of the relevant national and institutional committees on human experimentation and with the Helsinki Declaration of 1975, as revised in 2008. This study protocol was reviewed and approved by the local ethics committee of the faculty of psychology at the Ruhr University Bochum (2021/686). Written informed consent was obtained from participants to participate in the study.

Results

Participant Characteristics at Baseline

Figure 1 shows the distribution of participants among the courses and the sample sizes at the measurement time points. A total of 350 participants took part in at least 1 of the 3 measurement time points, 203 of whom participated in all of them. Demographic data were missing from 17 participants because they did not complete the online survey that was part of the behavioral test. Participants (n=333) had a mean age of 23.6 (SD 4.52) years and the majority identified as women (279/333, 83.8%). The average study year was 2.82 (SD 0.93).

Figure 1. CONSORT (Consolidated Standards of Reporting Trials) flowchart of students enrolled in the courses and participating in the study.
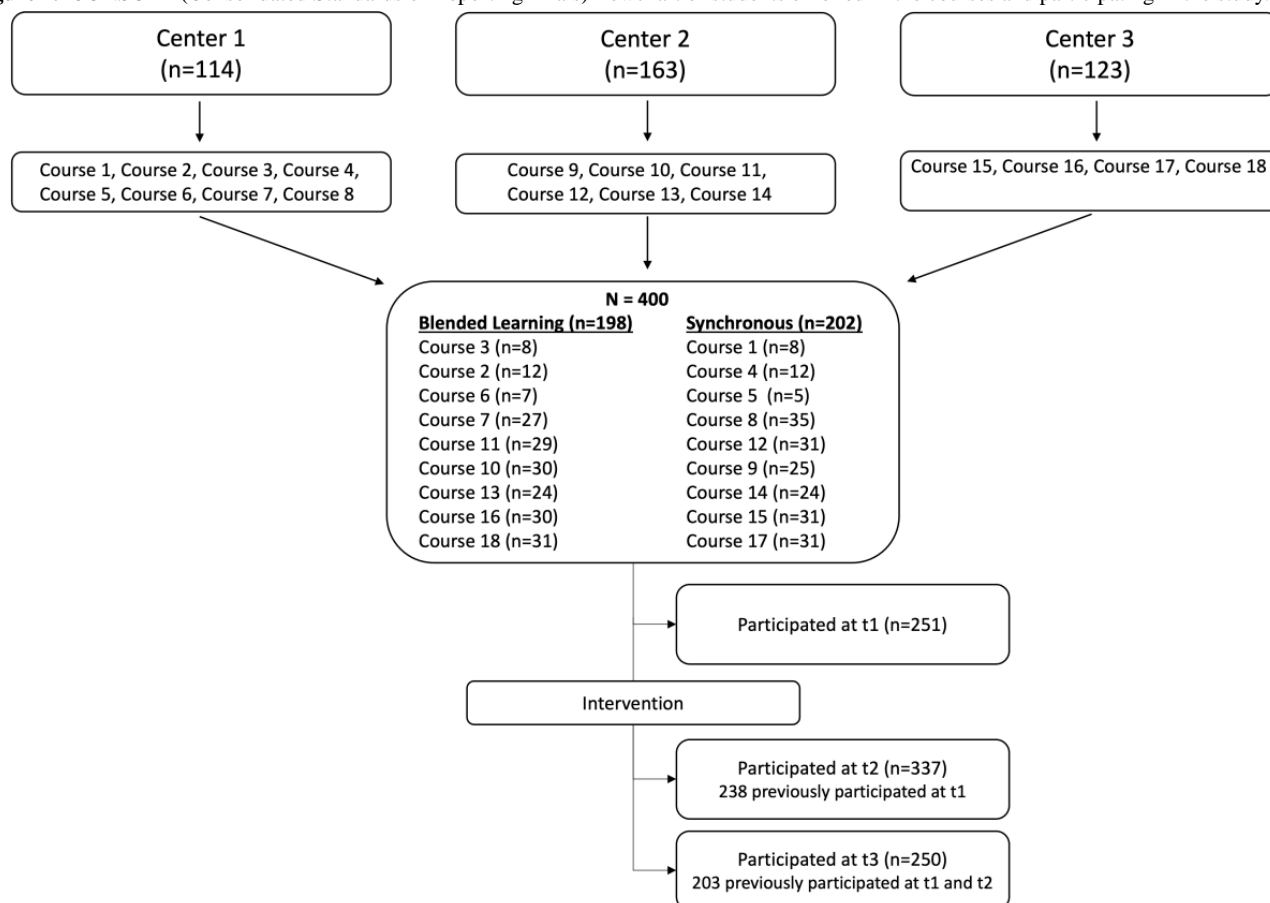
The t test and Fisher exact test revealed no significant differences between teaching conditions at baseline (all $P>.05$; Table 3). Chi-square tests showed no significant differences between completers and noncompleters regarding teaching condition ($\chi^2_1=0.79$, $P=.37$) and behavioral performance ($\chi^2_1=0.82$, $P=.37$).

Table 3. Participants' baseline demographic characteristics and diagnostic knowledge.

Characteristic	Participants, n	Blended learning	Synchronous	t test (df)	P value
Age (years), mean (SD)	333	24.1 (4.20)	23.2 (4.76)	1.81 (331)	.07
Study year, mean (SD)	333	2.88 (0.93)	2.76 (0.94)	1.09 (331)	.28
Diagnostic knowledge, mean (SD)					
Test score (t1)	251	9.44 (1.71)	9.36 (1.79)	0.37 (249)	.71
Self-rating	333	3.46 (2.07)	3.51 (2.18)	–0.22 (331)	.82
Gender, n (%)	350				.05
Missing		11 (6.47)	7 (3.89)	—[a]	
Female		127 (74.71)	152 (84.44)	—	
Male		31 (18.23)	18 (10.0)	—	
Diverse		1 (0.59)	3 (1.67)	—	

[a]Not applicable.

Primary Outcome—Practical Diagnostic Skills

Overall, participants showed high levels of interpersonal (blended learning: mean 74.8, SD 16.2; synchronous: mean 70.7, SD 18.1) and formal skills (blended learning: mean 86.1, SD 14.4; synchronous: mean 82.8, SD 16.6). The passing rate was 89% in the blended learning condition and 74.6% in the synchronous condition, corresponding to an OR of 2.77 (95% CI 1.52-5.03).

We furthermore tested whether this finding still held up when several covariates were considered. For this, we fitted a logistic mixed model (adjusted model; n=320) including the course variable as random effect to take cluster randomization into

account (Table 4; for all model coefficients, refer to "Section D" in Multimedia Appendix 2). As the model resulted in convergence errors with all predictors, the knowledge test score at t1 and the study center had to be excluded. More complex models accounting for the nesting of courses within universities were attempted to be fitted but resulted in convergence errors. Intraclass correlation is not reported as the results of the generalized linear mixed model did not contain the residual variance required to calculate the intraclass correlation.

Table 4. Odds ratio (OR) and 95% CI for the unadjusted and adjusted model.

Predictors	Unadjusted model, OR (95% CI)	Adjusted model, OR (95% CI)
Teaching condition	2.77 (1.55-5.13)	3.20 (1.56-6.71)
Center 2	—^{a}	—
Center 3	—	—
Course focus	—	0.42 (0.17-0.96)
Study year	—	1.36 (0.82-2.42)
Self-reported knowledge	—	1.06 (0.89-1.25)
Knowledge test (t1)	—	—
Random effects		
σ^2	—	3.29
N_{course}	—	18
Observations	337	320
Tjur D	0.035	0.214

^{a}Not applicable.

Secondary Outcomes

Overview

To describe the magnitude of the differences in secondary outcomes between the groups, multiple linear regression models were calculated (Table 5). For the complete covariate-adjusted models, refer to "Section E" in Multimedia Appendix 2.

Table 5. Means, SDs, and β-coefficients with CIs for all secondary outcomes.

Secondary outcome	Range	Blended learning (n=117), mean (SD)	Synchronous (n=132), mean (SD)	Teaching condition, β (95% CI)
Diagnostic knowledge				
Knowledge test (t3)	0-15	12.0 (1.65)	11.4 (1.68)	.13 (0.01 to 0.26)
Participants' reactions				
Intent to recommend	0-7	6.21 (1.06)	6.09 (1.13)	.09 (–0.05 to 0.22)
Experience of overload[a]	0-7	2.70 (1.04)	2.40 (1.04)	.20 (0.07 to 0.34)
Subjective learning success	0-7	5.89 (1.02)	5.80 (1.06)	.04 (–0.096 to 0.18)
Clarity	0-7	6.18 (0.74)	5.48 (0.91)	.40 (0.27 to 0.53)
Likeability	0-7	6.09 (0.96)	5.86 (1.21)	.09 (–0.04 to 0.23)
Informativeness	0-7	6.38 (0.68)	6.19 (0.72)	.19 (0.06 to 0.33)
Credibility	0-7	6.46 (0.64)	6.45 (0.62)	.08 (–0.05 to 0.22)
Course structure	0-7	6.30 (0.72)	5.98 (0.92)	.18 (0.04 to 0.32)
Overall impression	1-6	1.56 (0.81)	1.72 (0.75)	–.12 (–0.26 to 0.01)
Visual Aesthetics	0-7	5.87 (0.91)	__b	—
Usability	0-100	86.3 (14.1)	—	—
Acceptance				
Global rating	0-100	77.7 (14.2)	76.8 (16.2)	.04 (–0.095 to 0.18)
"More confused"[a]	0-3	0.26 (0.48)	0.42 (0.58)	–.11 (–0.25 to 0.03)
"questioned out"[a]	0-3	1.18 (0.74)	1.20 (0.85)	.05 (–0.08 to 0.19)
"too many questions"[a]	0-3	1.11 (0.81)	1.06 (0.76)	.03 (–0.11 to 0.16)
"exhausting"[a]	0-3	1.15 (0.71)	1.07 (0.77)	.06 (–0.08 to 0.20)
"taken seriously"	0-3	2.33 (0.88)	2.33 (0.84)	–.01 (–0.15 to 0.13)
"positive relationship"	0-3	1.95 (0.80)	2.04 (0.75)	–.10 (–0.24 to 0.04)
"not report everything"[a]	0-3	1.33 (0.81)	1.32 (0.86)	.04 (–0.10 to 0.18)
"better understanding"	0-3	1.56 (0.76)	1.50 (0.80)	.07 (–0.07 to 0.21)
"enough detail"	0-3	2.30 (0.75)	2.36 (0.72)	–.05 (–0.19 to 0.09)
"helpful"	0-3	2.12 (0.74)	2.14 (0.72)	–.03 (–0.17 to 0.12)

[a]Lower scores indicate better outcome. For negatively-keyed items, noninferiority of the blended learning course can be assumed if the upper bound of the CI is greater than 0.10.

[b]Not applicable.

Diagnostic Knowledge

Participation in the blended learning course increased the knowledge score at t3 (β=.13, 95% CI 0.01-0.26). Thus, noninferiority and superiority of the blended learning course regarding diagnostic knowledge at t3 can be assumed.

Participants' Reactions

Noninferiority of the blended learning course regarding participants' reactions to the courses could be observed in most measures collected. Only with regard to the experience of overload the blended learning course was inferior to the synchronous course, with lower scores indicating a more favorable outcome (β=.20, 95% CI 0.07-0.34). Furthermore, the superiority of the blended learning over the synchronous course could be found in the following subscales: clarity (β=.40, 95% CI 0.27-0.53), course structure (β=.18, 95% CI 0.04-0.32), and informativeness (β=.19, 95% CI 0.06-0.32).

Regarding the estimated patient acceptance of structured interviews, noninferiority of the blended learning course was observed for the global acceptance rating (β=.04, 95% CI –0.095 to 0.18) and the items "After a structured interview, patients feel more confused than before" (β=–0.11, 95% CI –0.25 to 0.03) and "Patients have the feeling that they understand themselves and their problems better, after a structured interview" (β=.07, 95% CI –0.07 to 0.21). For the other items, student's estimation did not differ between the blended learning and the synchronous courses.

Discussion

Principal Findings

The aim of the present study was to establish whether a blended learning course with a reduced personal contact time results in comparable clinical diagnostic skills as a traditional synchronous online course.

The results of this study are in line with and extend the existing literature on blended learning [24,35,36]. First, noninferiority and superiority of the blended learning course over the synchronous course could be found for the primary outcome measure—the performance in a simulated structured diagnostic interview. Second, noninferiority and superiority were also observed for the diagnostic knowledge test score at t3 and several reaction measures, such as clarity, informativeness, and structure of the course. Furthermore, noninferiority could be observed regarding the intention to recommend the course to other students, subjective learning success, likeability, credibility, overall impression of the course, and 3 items of the estimated patient acceptance of structured clinical interviews. Third, inferiority of the blended learning compared with the synchronous course was found for the participants' experience of overload.

Despite the described differences, participants in both courses showed high levels of interpersonal and formal skills, good diagnostic knowledge, positive reactions to the courses, and high-estimated patient acceptance. While therapists were found to underestimate patient acceptance of structured interviews [6], estimated patient acceptance ratings in this study correspond more closely to patients' actual acceptance ratings [60], indicating that participants of the present study estimated patient acceptance more accurately than did therapists in the aforementioned study.

Limitations and Strengths

The study has some limitations that should be mentioned. First, the blended learning course was presented as a block, meaning that participants had 3 weeks of time to work through the asynchronous online content followed by 3 weekly synchronous online sessions. In contrast, the synchronous online course consisted of 11 weekly sessions. The fact that participants only had 3 weeks for the online content, which was equivalent to 8 sessions of 90 minutes each, might be considered a disadvantage for the blended learning course. This might explain why participants in the blended learning course reported higher levels of overload than those in the synchronous online course. To reduce the experience of overload and possibly even further enhance students' performance, the blended learning course should be provided on a continuous basis in the future, ensuring continuous student activity [61]. Second, adherence in the control condition was not assessed. Although teachers were informed about the mandatory content in a training session and received course material before the start of the course, it remains unclear whether all mandatory content was in fact taught in the control condition. In contrast, the asynchronous online component of the blended learning course was meticulously developed. It was ensured that the blended learning course contained all the necessary content for the diagnostic skills and knowledge tests. In addition, formative evaluations were conducted to assess the content, usability, and design of the course. Based on the feedback received, the course was subsequently revised. Therefore, the blended course was designed and implemented with more effort than the control condition, which followed a "teaching as usual" rationale. Third, we did not evaluate how the participants used the courses and materials for preparation, repetition, and reflection of the lectures. The convenient repetition offered by the asynchronous materials in the blended-learning course may have been particularly beneficial for the practical skills and knowledge tests. Fourth, all synchronous sessions in both the experimental and control conditions had to be conducted online due to governmental restrictions associated with the COVID-19 pandemic. As a result, the experimental condition had to be adapted to a hybrid combination of synchronous and asynchronous online teaching, although it was initially intended to be a combination of synchronous face-to-face teaching and asynchronous online teaching. Thus, it could be argued that our experimental condition does not strictly qualify as blended learning since it did not include face-to-face teaching. To evaluate the potential implications of this adaption, it is important to consider several differences between synchronous online and offline teaching. While both provide the benefit of immediate educational support and feedback [62], online teaching has logistical, instructional, and financial benefits over offline teaching [63]. However, online learning's logistical flexibility also has the potential downside of causing social isolation for the learner [64]. In addition, the integration of technology in educational settings inevitably increases the likelihood of technical issues, which can decrease satisfaction and participation [64,65]. Recent meta-analyses suggest that there are no significant differences in learning success between online and offline teaching [19,20]. Furthermore, online teaching received significantly higher satisfaction ratings compared with offline teaching [19]. It is important to note that our blended learning course primarily consisted of asynchronous online sessions. It can be assumed that the impact of 3 synchronous sessions, regardless of whether they are taught online or offline, is relatively small. However, it remains unclear how the blended learning course would compare with a traditional face-to-face course.

Besides these limitations, the study also has some notable strengths. First, as a multicenter cluster randomized controlled trial, it makes an important contribution to the scarce evidence on the efficacy of blended learning in general [33-35] and, more specifically, for teaching evidence-based diagnostics. Second, the design of the evaluation study was developed very carefully. For instance, the outcome measures were assessed with reliability and validity in mind, the case vignettes for students' performance tests included very precise instructions, and actors were trained beforehand to ensure a high standardization. In addition, the items of the knowledge test were piloted on laypersons and therapists. Third, as undergraduate psychology students from 3 German universities attending a mandatory seminar of the diagnostics of mental disorders were invited to participate, a large sample of 337 participants could be included in the analysis of the primary outcome and 203 participants took part at all 3 measurement time points. Fourth, as the study was

conducted in an ongoing university setting, a high external validity and generalizability of the study results can be assumed.

Clinical Implications and Future Research

As the results indicate that the blended learning course can be used to teach evidence-based diagnostics, we aim to disseminate the blended learning course open access throughout Germany—at universities (undergraduate and graduate courses), at institutions of tertiary education, and among practicing psychotherapists. In order to facilitate the adoption of the blended learning course [66], a technical infrastructure was chosen which is available free of charge and provides ongoing technical support. In addition, an interesting question for future research is whether structured interviews are in fact used more frequently after attending the blended learning course. Increasing the use of structured interviews in clinical practice is an important goal as therapists appeared to use structured interviews only with 14.8% (55/370) of their patients [6]. Until today, research on therapist training is limited, especially when

it comes to web-based training [67]. Therefore, to extend the promising findings of this study, future research should also focus on the development and evaluation of further blended learning courses to improve evidence-based practice in clinical psychology in general.

Conclusion

In conclusion, this study suggests that a blended learning course, compared with a synchronous online course in a cluster randomized controlled trial, can be used to efficiently teach evidence-based diagnostics. The results indicate that the blended learning approach was more effective than synchronous online teaching in acquiring practical diagnostic skills and diagnostic knowledge, and that it was well received by the students. The blended learning course can therefore help to improve the skills and knowledge of (future) clinicians in a time- and cost-efficient way and thus make an important contribution to improving the diagnostics of mental disorders and the mental health care situation in the long term.

Acknowledgments

Conflicts of Interest

None declared.

Multimedia Appendix 1

CONSORT-eHEALTH checklist (V 1.6.x).
[PDF File (Adobe PDF File), 1043 KB-Multimedia Appendix 1]

Multimedia Appendix 2

Case vignette and instructions for the actors, coding scheme, knowledge test, logistic regression coefficients for the primary outcome, and multiple linear regression coefficients for secondary outcomes.
[DOCX File , 100 KB-Multimedia Appendix 2]

References

1. Jensen-Doss A, Weisz JR. Diagnostic agreement predicts treatment process and outcomes in youth mental health clinics. J Consult Clin Psychol. 2008;76(5):711-722. [doi: 10.1037/0022-006X.76.5.711] [Medline: 18837589]
2. Pogge DL, Wayland-Smith D, Zaccario M, Borgaro S, Stokes J, Harvey PD. Diagnosis of manic episodes in adolescent inpatients: structured diagnostic procedures compared to clinical chart diagnoses. Psychiatry Res. 2001;101(1):47-54. [doi: 10.1016/s0165-1781(00)00248-1] [Medline: 11223119]
3. Ehlert U. Eine Psychotherapie ist immer nur so gut wie ihre Diagnostik. Psychotherapy is only as good as its diagnostics. Verhaltenstherapie. 2007;17(2):81-82. [doi: 10.1159/000103156]
4. Merten EC, Schneider S. Clinical interviews with children and adolescents. In: Hofmann SG, editor. Clinical Psychology: A Global Perspective. Hoboken, NJ. Wiley; 2017.
5. Rettew DC, Lynch AD, Achenbach TM, Dumenci L, Ivanova MY. Meta-analyses of agreement between diagnoses made from clinical evaluations and standardized diagnostic interviews. Int J Methods Psychiatr Res. 2009;18(3):169-184. [FREE Full text] [doi: 10.1002/mpr.289] [Medline: 19701924]
6. Bruchmüller K, Margraf J, Suppiger A, Schneider S. Popular or unpopular? Therapists' use of structured interviews and their estimation of patient acceptance. Behav Ther. 2011;42(4):634-643. [doi: 10.1016/j.beth.2011.02.003] [Medline: 22035992]
7. Jensen-Doss A, Youngstrom EA, Youngstrom JK, Feeny NC, Findling RL. Predictors and moderators of agreement between clinical and research diagnoses for children and adolescents. J Consult Clin Psychol. 2014;82(6):1151-1162. [FREE Full text] [doi: 10.1037/a0036657] [Medline: 24773574]

8. Merten EC, Cwik JC, Margraf J, Schneider S. Overdiagnosis of mental disorders in children and adolescents (in developed countries). Child Adolesc Psychiatry Ment Health. 2017;11:5. [FREE Full text] [doi: 10.1186/s13034-016-0140-5] [Medline: 28105068]

9. Mojtabai R. Clinician-identified depression in community settings: concordance with structured-interview diagnoses. Psychother Psychosom. 2013;82(3):161-169. [doi: 10.1159/000345968] [Medline: 23548817]

10. Ruggero CJ, Zimmerman M, Chelminski I, Young D. Borderline personality disorder and the misdiagnosis of bipolar disorder. J Psychiatr Res. 2010;44(6):405-408. [FREE Full text] [doi: 10.1016/j.jpsychires.2009.09.011] [Medline: 19889426]

11. Vermani M, Marcus M, Katzman MA. Rates of detection of mood and anxiety disorders in primary care: a descriptive, cross-sectional study. Prim Care Companion CNS Disord. 2011;13(2):PCC.10m01013. [FREE Full text] [doi: 10.4088/PCC.10m01013] [Medline: 21977354]

12. Berardi D, Menchetti M, Cevenini N, Scaini S, Versari M, de Ronchi D. Increased recognition of depression in primary care. Comparison between primary-care physician and ICD-10 diagnosis of depression. Psychother Psychosom. 2005;74(4):225-230. [doi: 10.1159/000085146] [Medline: 15947512]

13. Bruchmüller K, Margraf J, Schneider S. Is ADHD diagnosed in accord with diagnostic criteria? Overdiagnosis and influence of client gender on diagnosis. J Consult Clin Psychol. 2012;80(1):128-138. [doi: 10.1037/a0026582] [Medline: 22201328]

14. Margraf J, Schneider S. From neuroleptics to neuroscience and from Pavlov to psychotherapy: more than just the "emperor's new treatments" for mental illnesses? EMBO Mol Med. 2016;8(10):1115-1117. [FREE Full text] [doi: 10.15252/emmm.201606650] [Medline: 27621275]

15. McMain S, Newman MG, Segal ZV, DeRubeis RJ. Cognitive behavioral therapy: current status and future research directions. Psychother Res. 2015;25(3):321-329. [doi: 10.1080/10503307.2014.1002440] [Medline: 25689506]

16. Weisz JR, Ng MY, Bearman SK. Odd couple? Reenvisioning the relation between science and practice in the dissemination-implementation Era. Clinical Psychological Science. 2013;2(1):58-74. [doi: 10.1177/2167702613501307]

17. Hunsley J, Mash EJ. Introduction to the special section on developing guidelines for the evidence-based assessment (EBA) of adult disorders. Psychol Assess. 2005;17(3):251-255. [doi: 10.1037/1040-3590.17.3.251] [Medline: 16262451]

18. Jensen-Doss A, Hawley KM. Understanding barriers to evidence-based assessment: clinician attitudes toward standardized assessment tools. J Clin Child Adolesc Psychol. 2010;39(6):885-896. [FREE Full text] [doi: 10.1080/15374416.2010.517169] [Medline: 21058134]

19. He L, Yang N, Xu L, Ping F, Li W, Sun Q, et al. Synchronous distance education vs traditional education for health science students: a systematic review and meta-analysis. Med Educ. 2021;55(3):293-308. [doi: 10.1111/medu.14364] [Medline: 32881047]

20. Ebner C, Gegenfurtner A. Learning and satisfaction in webinar, online, and face-to-face instruction: a meta-analysis. Front. Educ. 2019;4:92. [doi: 10.3389/feduc.2019.00092]

21. Laster S, Otte G, Picciano A, Sorg S. Redefining blended learning. 2005. Presented at: 2005 Sloan-C Workshop on Blended Learning; April 18, 2005; Chicago, IL.

22. Yu Z, Xu W, Sukjairungwattana P. Meta-analyses of differences in blended and traditional learning outcomes and students' attitudes. Front Psychol. 2022;13:926947. [FREE Full text] [doi: 10.3389/fpsyg.2022.926947] [Medline: 36186290]

23. Vallée A, Blacher J, Cariou A, Sorbets E. Blended learning compared to traditional learning in medical education: systematic review and meta-analysis. J Med Internet Res. 2020;22(8):e16504. [FREE Full text] [doi: 10.2196/16504] [Medline: 32773378]

24. Schneider M, Preckel F. Variables associated with achievement in higher education: a systematic review of meta-analyses. Psychol Bull. 2017;143(6):565-600. [doi: 10.1037/bul0000098] [Medline: 28333495]

25. Shafran R, Clark DM, Fairburn CG, Arntz A, Barlow DH, Ehlers A, et al. Mind the gap: improving the dissemination of CBT. Behav Res Ther. 2009;47(11):902-909. [doi: 10.1016/j.brat.2009.07.003] [Medline: 19664756]

26. Stewart RE, Chambless DL, Baron J. Theoretical and practical barriers to practitioners' willingness to seek training in empirically supported treatments. J Clin Psychol. 2012;68(1):8-23. [FREE Full text] [doi: 10.1002/jclp.20832] [Medline: 21901749]

27. Fairburn CG, Cooper Z. Therapist competence, therapy quality, and therapist training. Behav Res Ther. 2011;49(6-7):373-378. [FREE Full text] [doi: 10.1016/j.brat.2011.03.005] [Medline: 21492829]

28. Jackson CB, Quetsch LB, Brabson LA, Herschell AD. Web-based training methods for behavioral health providers: a systematic review. Adm Policy Ment Health. 2018;45(4):587-610. [FREE Full text] [doi: 10.1007/s10488-018-0847-0] [Medline: 29352459]

29. Khanna MS, Kendall PC. Bringing technology to training: web-based therapist training to promote the development of competent cognitive-behavioral therapists. Cognitive and Behavioral Practice. 2015;22(3):291-301. [doi: 10.1016/j.cbpra.2015.02.002]

30. Seehagen S, Pflug V, Schneider S. Psychotherapy and science - harmony or dissonance? [Article in German]. Z Kinder Jugendpsychiatr Psychother. 2012;40(5):301-306. [doi: 10.1024/1422-4917/a000186] [Medline: 22869223]

31. Mali D, Lim H. How do students perceive face-to-face/blended learning as a result of the COVID-19 pandemic? Int J Manag Educ. 2021;19(3):100552. [doi: 10.1016/j.ijme.2021.100552]

32. Singh J, Steele K, Singh L. Combining the best of online and face-to-face learning: hybrid and blended learning approach for COVID-19, post vaccine, and post-pandemic world. J Educ Technol Syst. 2021;50(2):1-32. [doi: 10.1177/00472395211047865]

33. Ilic D, Nordin RB, Glasziou P, Tilson JK, Villanueva E. A randomised controlled trial of a blended learning education intervention for teaching evidence-based medicine. BMC Med Educ. 2015;15:39. [FREE Full text] [doi: 10.1186/s12909-015-0321-6] [Medline: 25884717]

34. Liu Q, Peng W, Zhang F, Hu R, Li Y, Yan W. The effectiveness of blended learning in health professions: systematic review and meta-analysis. J Med Internet Res. 2016;18(1):e2. [FREE Full text] [doi: 10.2196/jmir.4807] [Medline: 26729058]

35. Lozano-Lozano M, Fernández-Lao C, Cantarero-Villanueva I, Noguerol I, Álvarez-Salvago F, Cruz-Fernández M, et al. A blended learning system to improve motivation, mood state, and satisfaction in undergraduate students: randomized controlled trial. J Med Internet Res. 2020;22(5):e17101. [FREE Full text] [doi: 10.2196/17101] [Medline: 32441655]

36. Ma L, Lee CS. Evaluating the effectiveness of blended learning using the ARCS model. Computer Assisted Learning. 2021;37(5):1397-1408. [doi: 10.1111/jcal.12579]

37. Beidas RS, Kendall PC. Training therapists in evidence-based practice: a critical review of studies from a systems-contextual perspective. Clin Psychol (New York). 2010;17(1):1-30. [doi: 10.1111/j.1468-2850.2009.01187.x] [Medline: 20877441]

38. Chu BC, Crocco ST, Arnold CC, Brown R, Southam-Gerow MA, Weisz JR. Sustained implementation of cognitive-behavioral therapy for youth anxiety and depression: long-term effects of structured training and consultation on therapist practice in the field. Prof Psychol Res Pr. 2015;46(1):70-79. [FREE Full text] [doi: 10.1037/a0038000] [Medline: 26366037]

39. McCarty RJ, Cooke DL, Lazaroe LM, Guzick AG, Guastello AD, Budd SM, et al. The effects of an exposure therapy training program for pre-professionals in an intensive exposure-based summer camp. Cogn Behav Ther. 2022;15:e5. [doi: 10.1017/s1754470x22000010]

40. Karabulut‐Ilgu A, Jaramillo Cherrez N, Jahren CT. A systematic review of research on the flipped learning method in engineering education. Brit J Educational Tech. 2017;49(3):398-411. [doi: 10.1111/bjet.12548]

41. van der Kleij FM, Feskens RCW, Eggen TJHM. Effects of feedback in a computer-based learning environment on students' learning outcomes. Rev Educ Res. 2015;85(4):475-511. [doi: 10.3102/0034654314564881]

42. Diep AN, Zhu C, Struyven K, Blieck Y. Who or what contributes to student satisfaction in different blended learning modalities? Brit J Educational Tech. 2016;48(2):473-489. [doi: 10.1111/bjet.12431]

43. Margraf J, Cwik JC, von Brachel R, Suppiger A, Schneider S. DIPS Open Access 1.2: Diagnostic Interview for Mental Disorders. Bochum, Germany. Mental Health Research and Treament Center, Ruhr-Universität Bochum; 2021.

44. Schneider S, Pflug V, In-Albon T, Margraf J. Kinder-DIPS Open Access: Diagnostisches Interview bei psychischen Störungen im Kindes- und Jugendalter. Bochum, Germany. Forschungs- und Behandlungszentrum für psychische Gesundheit, Ruhr-Universität Bochum; 2017.

45. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics. 1977;33(1):159-174. [Medline: 843571]

46. Hirschfeld G, Thielsch M. Glöckner-Rist A, editor. Münsteraner Fragebogen zur Evaluation von Seminaren - revidiert (MFE-Sr). Bonn, Germany. Zusammenstellung sozialwissenschaftlicher Items und Skalen ZIS Version 140; 2010.

47. Thielsch MT, Hirschfeld G. Facets of website content. Human–Computer Interaction. 2018;34(4):279-327. [doi: 10.1080/07370024.2017.1421954]

48. Zumbach J, Spinath B, Schahn J, Friedrich M, Kögel M. Entwicklung einer Kurzskala zur Lehrevaluation. Development of a short scale for teaching evaluation. In: Psychologiedidaktik und Evaluation VI. Psychological Didactics and Evaluation VI. Göttingen, Germany. V & R Unipress; 2007:317-325.

49. Lewis JR, Utesch BS, Maher DE. UMUX-LITE - When there's no time for the SUS. 2013. Presented at: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems; April 27, 2013; Paris, France. [doi: 10.1145/2470654.2481287]

50. Moshagen M, Thielsch M. A short version of the visual aesthetics of websites inventory. Behav Inf Technol. 2013;32(12):1305-1311. [doi: 10.1080/0144929x.2012.694910]

51. Cuzick J, Sasieni P. Interpreting the results of noninferiority trials-a review. Br J Cancer. 2022;127(10):1755-1759. [FREE Full text] [doi: 10.1038/s41416-022-01937-w] [Medline: 36104512]

52. Non-inferiority clinical trials to establish effectiveness: guidance for industry. U.S. Food and Drug Administration. 2016. URL: https://www.fda.gov/regulatory-information/search-fda-guidance-documents/non-inferiority-clinical-trials [accessed 2024-08-13]

53. Choice of a non-inferiority margin - scientific guideline. European Medicines Agency. 2005. URL: https://www.ema.europa.eu/en/choice-non-inferiority-margin-scientific-guideline [accessed 2024-08-13]

54. D'Agostino RB, Massaro JM, Sullivan LM. Non-inferiority trials: design concepts and issues - the encounters of academic consultants in statistics. Stat Med. 2003;22(2):169-186. [doi: 10.1002/sim.1425] [Medline: 12520555]

55. Rief W, Hofmann SG. Some problems with non-inferiority tests in psychotherapy research: psychodynamic therapies as an example. Psychol Med. 2018;48(8):1392-1394. [doi: 10.1017/S0033291718000247] [Medline: 29439745]

56. Laster LL, Johnson MF, Kotler ML. Non-inferiority trials: the 'at least as good as' criterion with dichotomous data. Stat Med. 2006;25(7):1115-1130. [doi: 10.1002/sim.2476] [Medline: 16381070]

57. R Core Team. R: A language and environment for statistical computing. Vienna, Austria. R Foundation for Statistical Computing; 2022.

58. Bonnin G, Kröber S, Schneider S, Margraf J, Pflug V, Gerlach A, et al. Dataset for: Blended learning for diagnostic skills: a multicenter cluster randomized non-inferiority trial. PsychArchives. 2023. URL: https://psycharchives.org/en/item/5cecec0f-b732-4012-a1e3-661a3e3c5182 [accessed 2024-11-19]

59. Bonnin G, Kröber S, Schneider S, Margraf J, Pflug V, Gerlach AL, et al. Code for: Blended learning for diagnostic skills: a multicenter cluster randomized non-inferiority trial. PsychArchives. 2023:1. [doi: 10.23668/psycharchives.12494]

60. Suppiger A, In-Albon T, Hendriksen S, Hermann E, Margraf J, Schneider S. Acceptance of structured diagnostic interviews for mental disorders in clinical practice and research settings. Behav Ther. 2009;40(3):272-279. [doi: 10.1016/j.beth.2008.07.002] [Medline: 19647528]

61. van Leeuwen A, Bos N, van Ravenswaaij H, van Oostenrijk J. The role of temporal patterns in students' behavior for predicting course performance: a comparison of two blended learning courses. Brit J Educational Tech. 2018;50(2):921-933. [doi: 10.1111/bjet.12616]

62. Chen NS, Ko HC, Kinshuk *, Lin T. A model for synchronous learning using the Internet. Innov Educ Teach Int. 2006;42(2):181-194. [doi: 10.1080/14703290500062599]

63. Alnabelsi T, Al-Hussaini A, Owens D. Comparison of traditional face-to-face teaching with synchronous e-learning in otolaryngology emergencies teaching to medical undergraduates: a randomised controlled trial. Eur Arch Otorhinolaryngol. 2015;272(3):759-763. [doi: 10.1007/s00405-014-3326-6] [Medline: 25308244]

64. Cook DA. Web-based learning: pros, cons and controversies. Clin Med (Lond). 2007;7(1):37-42. [FREE Full text] [doi: 10.7861/clinmedicine.7-1-37] [Medline: 17348573]

65. Cook DA, Dupras DM, Thompson WG, Pankratz VS. Web-based learning in residents' continuity clinics: a randomized, controlled trial. Acad Med. 2005;80(1):90-97. [doi: 10.1097/00001888-200501000-00022] [Medline: 15618102]

66. Porter WW, Graham CR. Institutional drivers and barriers to faculty adoption of blended learning in higher education. Brit J Educational Tech. 2015;47(4):748-762. [doi: 10.1111/bjet.12269]

67. Cooper Z, Bailey-Straebler S, Morgan KE, O'Connor ME, Caddy C, Hamadi L, et al. Using the internet to train therapists: randomized comparison of two scalable methods. J Med Internet Res. 2017;19(10):e355. [FREE Full text] [doi: 10.2196/jmir.8336] [Medline: 29046265]

Abbreviations

DIPS-OA1.2: Diagnostic Interview for Mental Disorders – Open Access 1.2
(Kinder-)DIPS-OA: Diagnostic Interview for Mental Disorders (in Children and Adolescents) – Open Access
OR: odds ratio