Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Impact of Flipped Classroom Combined with Mini-CEX Formative Assessment on Core Competencies in Undergraduate Endocrinology Education: A Randomized Controlled Trial

Langen Zhuang
The First Affiliated Hospital of Bengbu Medical University
Qianmin Ma
The First Affiliated Hospital of Bengbu Medical University
Shengqun Jiang
The First Affiliated Hospital of Bengbu Medical University
Xiaoyan Pei
The First Affiliated Hospital of Bengbu Medical University
Guoxi Jin
The First Affiliated Hospital of Bengbu Medical University
Qiong Xu
x035540163.com

The First Affiliated Hospital of Bengbu Medical University

Research Article

Keywords: flipped classroom, Mini-CEX, formative assessment, core competency, endocrinology, undergraduate medical education, randomized controlled trial

Posted Date: June 28th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9840288/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Abstract

Background

Traditional lecture-based instruction in undergraduate endocrinology education often falls short of cultivating the multidimensional core competencies required by modern medical curricula. Whether integrating a flipped classroom with Mini-Clinical Evaluation Exercise (Mini-CEX) formative assessment can address this gap remains unclear.

Objective

To evaluate the effect of a flipped classroom combined with Mini-CEX formative assessment on core competencies in undergraduate endocrinology education.

Methods

This single-center, prospective randomized controlled trial enrolled 120 third-year clinical medicine undergraduates between June 2024 and March 2026. Participants were randomly allocated to the intervention group (n = 60) or the control group (n = 60). The intervention group received a structured "pre-class online learning, in-class case discussion, post-class Mini-CEX assessment and feedback" cycle, while the control group received conventional lectures. Primary outcomes included theoretical examination and Objective Structured Clinical Examination (OSCE) scores. Secondary outcomes comprised Mini-CEX dimensional scores, Self-Directed Learning Readiness Scale (SDLRS) [Guglielmino, L.M. 1977. Development of the Self-Directed Learning Readiness Scale. Unpublished Doctoral Dissertation, University of Georgia. Dissertation Abstracts International, vol. 38 (11a), p. 6467] scores, Critical Thinking Disposition Inventory (CTDI-CV) scores, and teaching satisfaction.

Results

The intervention group achieved significantly higher theoretical examination scores (84.7 ± 7.3 vs. 79.8 ± 8.1, $P < 0.01$) and OSCE scores (85.9 ± 6.8 vs. 80.5 ± 7.6, $P < 0.01$). Significant improvements were also observed in SDLRS total scores ($P < 0.01$), CTDI-CV total scores ($P < 0.05$), and most Mini-CEX dimensions ($P < 0.05$). However, the intervention group perceived a heavier study workload ($P < 0.01$).

Conclusions

The flipped classroom integrated with Mini-CEX formative assessment significantly enhances multiple core competencies in undergraduate endocrinology education, though increased workload warrants optimization in future implementations.

Introduction

Endocrinology is a clinically complex discipline characterized by multisystem hormonal interactions, overlapping disease phenotypes, and rapidly evolving diagnostic and therapeutic paradigms [1]. Undergraduate students frequently find endocrine content abstract and difficult to integrate into clinical reasoning frameworks, and surveys of preclinical endocrine curricula have revealed considerable heterogeneity in teaching approaches and assessment methods across medical schools [2]. As the global burden of endocrine disorders, particularly diabetes mellitus and thyroid diseases, continues to increase, the demand for medical graduates who possess not only factual knowledge but also robust clinical competencies has become increasingly urgent [3]. These challenges underscore the need for evidence-based pedagogical innovations that go beyond information transfer and actively cultivate core competencies during the undergraduate training stage.

The transition from teacher-centered didactic instruction to student-centered active learning has been identified as a central priority in contemporary medical education reform [4, 5]. Competency-based medical education (CBME) frameworks emphasize that graduates should demonstrate measurable abilities in clinical reasoning, communication, self-directed learning, and professionalism, rather than merely accumulating credit hours [6]. Active learning strategies, including the flipped classroom, have emerged as promising vehicles for achieving these competency-based outcomes [7]. The flipped classroom model reverses the traditional instructional sequence by delivering foundational content through pre-class materials and devoting in-class time to interactive, higher-order learning activities such as case discussion and collaborative problem solving [8]. Systematic reviews in health professions education have confirmed that this approach is associated with improved knowledge acquisition, enhanced learner engagement, and greater student satisfaction relative to conventional lectures [9, 10].

However, knowledge acquisition alone does not equate to clinical competence. Workplace-based assessment (WBA) tools have been developed to evaluate the practical application of knowledge in authentic or simulated clinical encounters [11]. Among WBA instruments, the Mini-Clinical Evaluation Exercise (Mini-CEX) has gained widespread recognition as a structured formative assessment tool that enables direct observation of a learner's clinical performance followed by immediate, individualized feedback [12]. Originally designed for postgraduate training, the Mini-CEX evaluates multiple competency dimensions, including history taking, physical examination, clinical judgment, and communication skills, within a single brief encounter [13]. Recent studies have highlighted the potential of Mini-CEX to promote self-directed learning behaviors and professional identity formation when applied in undergraduate settings [14]. Nevertheless, most existing evidence on Mini-CEX pertains to postgraduate and clinical clerkship contexts, and its integration with other pedagogical strategies at the preclinical undergraduate level remains underexplored [15].

Despite the individual merits of the flipped classroom and Mini-CEX, few studies have investigated the synergistic effect of combining these two approaches into a unified instructional model for undergraduate medical education. Furthermore, prior research has predominantly focused on single-

dimension outcomes such as examination scores or satisfaction, without comprehensively assessing higher-order competencies including self-directed learning readiness and critical thinking disposition. To address these gaps, the present study designed a randomized controlled trial to evaluate the impact of a flipped classroom integrated with Mini-CEX formative assessment on core competencies in undergraduate endocrinology education, hypothesizing that this combined model would produce superior outcomes compared with traditional lecture-based instruction.

Materials and Methods

Study Design and Ethical Approval

This study was designed as a single-center, two-arm, prospective randomized controlled trial conducted at the Department of Endocrinology, [Hospital Name], affiliated with [University Name], between June 2024 and March 2026. The study protocol was approved by the Institutional Ethics Committee of [Hospital Name] (approval number: 2024-EC-ENDO-087; approval date: May 20, 2024) and was prospectively registered at the Chinese Clinical Trial Registry (registration number: ChiCTR2400081456). All procedures were performed in accordance with the Declaration of Helsinki (2013 revision). Written informed consent was obtained from each participant before enrollment.

Participants and Eligibility Criteria

The study population consisted of third-year clinical medicine undergraduates who were scheduled to complete the endocrinology curriculum during the study period. The inclusion criteria were as follows: (1) full-time undergraduate students enrolled in the five-year clinical medicine program; (2) first-time enrollment in the endocrinology course with no prior formal endocrinology training; (3) voluntary participation with signed informed consent. The exclusion criteria included: (1) students who had already completed an endocrinology rotation or elective; (2) students who were absent for more than 20% of the total scheduled sessions; (3) students who withdrew consent at any point during the study. A total of 128 eligible students were initially screened, of whom 120 met the inclusion criteria and were enrolled in the final analysis.

Randomization and Allocation Concealment

Enrolled participants were randomly allocated to the intervention group (n = 60) or the control group (n = 60) at a 1:1 ratio. The randomization sequence was generated by an independent statistician using a computer-based random number generator (SPSS version 26.0, IBM Corp., Armonk, NY, USA) with a block size of four. Allocation concealment was maintained through the use of sequentially numbered, opaque, sealed envelopes, which were opened only after the participant had completed baseline assessment. Due to the nature of the educational intervention, blinding of participants and instructors

was not feasible. However, outcome assessors and the data analyst were blinded to group allocation throughout the study.

Teaching Intervention

Control Group

Students assigned to the control group received conventional lecture-based instruction delivered by the same teaching faculty. Each teaching unit consisted of a 90-minute didactic lecture supplemented with projected slides, followed by a 30-minute question-and-answer session. The lecture content covered the same endocrinology topics as the intervention group, including diabetes mellitus, thyroid disorders, adrenal diseases, pituitary disorders, and metabolic bone diseases. Students were provided with textbook chapters and supplementary reading materials after each session. No structured formative assessment or individualized feedback was provided during the course.

Intervention Group

Students assigned to the intervention group underwent a blended teaching model integrating the flipped classroom approach with Mini-CEX formative assessment. The instructional cycle for each teaching unit comprised three sequential phases.

In the pre-class phase, students were required to complete self-directed learning tasks before each scheduled session. These tasks included watching pre-recorded video lectures (15 to 20 minutes each) uploaded to an online learning platform, reading assigned textbook sections and selected guideline excerpts, and completing a brief online quiz designed to assess baseline comprehension. Instructors reviewed quiz results before class to identify common misconceptions and knowledge gaps.

In the in-class phase, face-to-face sessions (90 minutes) were restructured to prioritize interactive learning. Each session began with a 15-minute instructor-led summary addressing the knowledge gaps identified from the pre-class quiz. This was followed by a 45-minute small-group case-based discussion, during which students collaboratively analyzed clinical scenarios involving real or standardized patient data relevant to the topic. The remaining 30 minutes were devoted to group presentations, peer questioning, and instructor-guided debriefing. Throughout this phase, instructors served as facilitators rather than primary lecturers.

In the post-class assessment phase, each student participated in a one-on-one bedside or simulated patient encounter evaluated using the Mini-CEX instrument. The Mini-CEX assessment was conducted once per teaching unit, covering seven predefined competency dimensions: medical interviewing, physical examination, clinical judgment, counseling, organization and efficiency, humanistic qualities and professionalism, and overall clinical competence. Each dimension was scored on a nine-point Likert scale (1 to 3, below expectations; 4 to 6, meets expectations; 7 to 9, exceeds expectations). Immediately after each encounter, the assessor provided structured verbal feedback lasting approximately five

minutes, highlighting specific strengths and areas requiring improvement. Students were encouraged to reflect on the feedback and integrate it into subsequent learning activities, thereby creating a continuous "learning, practice, assessment, feedback" cycle.

All teaching content, clinical case materials, and assessment rubrics were standardized across both groups to minimize content-related confounding. The same panel of five attending endocrinologists delivered instruction to both groups, and each instructor received a two-hour training workshop on the flipped classroom methodology and Mini-CEX scoring calibration before the study commenced.

Outcome Measures

Primary Outcomes

The primary outcomes were the end-of-course theoretical examination score and the Objective Structured Clinical Examination (OSCE) score. The theoretical examination consisted of 100 multiple-choice and short-answer questions covering the five core endocrinology topics, with a maximum score of 100 points. The OSCE comprised four standardized stations (history taking, physical examination, clinical reasoning, and patient counseling), each scored on a 100-point scale, and the final OSCE score was calculated as the mean of the four station scores.

Secondary Outcomes

Secondary outcomes included Mini-CEX dimensional scores, self-directed learning ability, critical thinking disposition, and teaching satisfaction. Self-directed learning ability was measured using the Self-Directed Learning Readiness Scale (SDLRS) [Guglielmino, L.M. 1977. Development of the Self-Directed Learning Readiness Scale. Unpublished Doctoral Dissertation, University of Georgia. Dissertation Abstracts International, vol. 38 (11a), p. 6467], a validated 40-item instrument scored on a five-point Likert scale with total scores ranging from 40 to 200, where higher scores indicate greater readiness for self-directed learning. Critical thinking disposition was assessed using the Chinese version of the Critical Thinking Disposition Inventory (CTDI-CV), a 70-item questionnaire with seven subscales, each scored on a six-point Likert scale, yielding total scores from 70 to 420. A total score above 280 is generally considered indicative of a positive critical thinking disposition. Both the SDLRS and CTDI-CV were administered at baseline (before the first teaching session) and at the end of the course (within one week after the final session). Teaching satisfaction was evaluated using a self-designed questionnaire consisting of 10 items rated on a five-point Likert scale, covering dimensions such as learning motivation, knowledge acquisition, clinical skill development, teaching format preference, and overall satisfaction. The questionnaire was distributed at the end of the course to all participants in both groups.

Sample Size Calculation

The sample size was estimated based on the primary outcome of theoretical examination scores. According to preliminary pilot data and published literature, a clinically meaningful difference of 5 points between groups was assumed, with an expected standard deviation of 5.5 points. Using a two-sided alpha of 0.05 and a statistical power of 0.90, a minimum of 52 students per group was required by the independent-samples t-test formula. Accounting for an anticipated dropout rate of 15%, the final target enrollment was set at 60 students per group, yielding a total of 120 participants.

Statistical Analysis

All statistical analyses were performed using SPSS version 26.0 (IBM Corp., Armonk, NY, USA). Continuous variables were expressed as mean ± standard deviation for normally distributed data or as median with interquartile range for non-normally distributed data. The normality of data distribution was examined using the Shapiro-Wilk test. Between-group comparisons of continuous variables were conducted using independent-samples t-tests for normally distributed data or Mann-Whitney U tests for non-normally distributed data. Within-group comparisons of pre- and post-intervention scores (SDLRS and CTDI-CV) were performed using paired-samples t-tests. Repeated-measures analysis of variance was applied to evaluate the interaction effect between group allocation and time point on SDLRS and CTDI-CV total scores. Categorical variables were presented as frequencies and percentages and compared using the chi-square test or Fisher exact test as appropriate. All statistical tests were two-sided, and a P value of less than 0.05 was considered statistically significant.

Results

Participant Flow and Baseline Characteristics

Of the 128 students initially screened, 120 met the eligibility criteria and were enrolled in the study. They were randomly allocated to the intervention group (n = 60) and the control group (n = 60). During the study period, three students in the intervention group and two students in the control group were excluded due to attendance rates falling below 80%. Consequently, 57 students in the intervention group and 58 students in the control group completed all assessments and were included in the final analysis (Fig. 1). The participant flow is illustrated in the CONSORT diagram.

Baseline demographic and academic characteristics of the two groups are presented in Table 1. There were no statistically significant differences between the groups in terms of age, sex distribution, prior grade point average, baseline SDLRS total score, or baseline CTDI-CV total score (all P > 0.05), indicating adequate comparability at enrollment.

Table 1
Baseline characteristics of participants in the intervention and control groups.

Characteristic	Intervention group (n = 57)	Control group (n = 58)	t / χ^2	P value
Age (years), mean ± SD	21.3 ± 1.1	21.5 ± 1.0	1.02	0.311
Sex, male/female	24/33	27/31	0.33	0.567
Prior GPA (out of 4.0), mean ± SD	3.12 ± 0.38	3.08 ± 0.41	0.55	0.582
Baseline SDLRS total score, mean ± SD	128.6 ± 16.3	130.2 ± 15.8	0.54	0.588
Baseline CTDI-CV total score, mean ± SD	274.1 ± 28.7	277.3 ± 27.4	0.63	0.532

Primary Outcomes

The end-of-course theoretical examination scores and OSCE scores are summarized in Table 2. Students in the intervention group achieved a mean theoretical examination score of 84.7 ± 7.3, which was significantly higher than the control group mean of 79.8 ± 8.1 (t = 3.49, P < 0.001). For the OSCE, the intervention group scored 85.9 ± 6.8 compared with 80.5 ± 7.6 in the control group (t = 4.10, P < 0.001). When OSCE station scores were analyzed individually, significant between-group differences were observed in the history taking station (86.2 ± 8.4 vs. 81.7 ± 9.1, P = 0.007), the clinical reasoning station (85.1 ± 8.9 vs. 79.3 ± 9.7, P < 0.001), and the patient counseling station (87.4 ± 7.5 vs. 82.1 ± 8.8, P = 0.001). The physical examination station also showed a numerical advantage for the intervention group; however, the difference did not reach statistical significance (84.8 ± 9.6 vs. 81.4 ± 10.2, P = 0.069).

Table 2
Comparison of primary outcome scores between the intervention and control groups.

Outcome	Intervention group (n = 57), mean ± SD	Control group (n = 58), mean ± SD	t value	P value
Theoretical examination (total)	84.7 ± 7.3	79.8 ± 8.1	3.49	< 0.001
OSCE total score	85.9 ± 6.8	80.5 ± 7.6	4.10	< 0.001
OSCE: History taking	86.2 ± 8.4	81.7 ± 9.1	2.76	0.007
OSCE: Physical examination	84.8 ± 9.6	81.4 ± 10.2	1.84	0.069
OSCE: Clinical reasoning	85.1 ± 8.9	79.3 ± 9.7	3.41	< 0.001
OSCE: Patient counseling	87.4 ± 7.5	82.1 ± 8.8	3.56	0.001

Mini-CEX Formative Assessment Scores

The Mini-CEX assessment was conducted only in the intervention group. The mean scores across the seven competency dimensions obtained from the final assessment are presented in Table 3. Among all dimensions, counseling (7.1 ± 1.0) and medical interviewing (6.9 ± 1.1) received the highest ratings, whereas humanistic qualities and professionalism (6.0 ± 1.4) and organization and efficiency (6.2 ± 1.3) received the lowest ratings. To evaluate the trajectory of improvement, scores from the first and last Mini-CEX encounters were compared using paired-samples t-tests. Significant improvements were observed in medical interviewing (5.4 ± 1.3 to 6.9 ± 1.1, P < 0.001), clinical judgment (5.1 ± 1.4 to 6.6 ± 1.2, P < 0.001), counseling (5.6 ± 1.2 to 7.1 ± 1.0, P < 0.001), physical examination (5.3 ± 1.5 to 6.5 ± 1.3, P < 0.001), and organization and efficiency (5.0 ± 1.4 to 6.2 ± 1.3, P < 0.001). The improvement in humanistic qualities and professionalism was modest and did not reach statistical significance (5.5 ± 1.5 to 6.0 ± 1.4, P = 0.072). Overall clinical competence also improved significantly (5.2 ± 1.3 to 6.7 ± 1.1, P < 0.001). The dimensional profile of the intervention group at the first and final Mini-CEX encounters is depicted in Fig. 2.

Table 3
Mini-CEX dimension scores at the first and final assessment in the intervention group (n = 57).

Dimension	First encounter, mean ± SD	Final encounter, mean ± SD	t value	P value
Medical interviewing	5.4 ± 1.3	6.9 ± 1.1	6.72	< 0.001
Physical examination	5.3 ± 1.5	6.5 ± 1.3	4.63	< 0.001
Clinical judgment	5.1 ± 1.4	6.6 ± 1.2	6.18	< 0.001
Counseling	5.6 ± 1.2	7.1 ± 1.0	7.24	< 0.001
Organization and efficiency	5.0 ± 1.4	6.2 ± 1.3	4.81	< 0.001
Humanistic qualities and professionalism	5.5 ± 1.5	6.0 ± 1.4	1.82	0.072
Overall clinical competence	5.2 ± 1.3	6.7 ± 1.1	6.55	< 0.001

Self-Directed Learning Readiness

The SDLRS total scores and subscale scores measured at baseline and post-intervention are presented in Table 4. At baseline, SDLRS total scores were comparable between the two groups (128.6 ± 16.3 vs. 130.2 ± 15.8, P = 0.588). After the intervention, the intervention group showed a significantly higher SDLRS total score than the control group (152.3 ± 14.7 vs. 138.6 ± 16.1, P < 0.001). Repeated-measures analysis of variance revealed a significant group-by-time interaction effect on SDLRS total scores (F = 18.42, P < 0.001), indicating that the intervention group experienced a greater increase in self-directed

learning readiness over time. At the subscale level, the intervention group demonstrated significantly greater improvements in self-management ($P < 0.001$), desire for learning ($P = 0.003$), and self-control ($P = 0.008$) compared with the control group.

Critical Thinking Disposition

The CTDI-CV total scores and subscale scores are also presented in Table 4. Baseline CTDI-CV total scores did not differ significantly between the groups (274.1 ± 28.7 vs. 277.3 ± 27.4, $P = 0.532$). After course completion, the intervention group attained a mean CTDI-CV total score of 298.5 ± 26.3, compared with 285.7 ± 27.8 in the control group ($P = 0.011$). The group-by-time interaction was statistically significant ($F = 7.86$, $P = 0.006$). Subscale analysis indicated that the intervention group showed significantly greater improvement in analyticity ($P = 0.004$), inquisitiveness ($P = 0.012$), and self-confidence ($P = 0.031$). However, the differences in open-mindedness ($P = 0.187$), systematicity ($P = 0.093$), truth-seeking ($P = 0.214$), and cognitive maturity ($P = 0.146$) did not reach statistical significance, suggesting that these higher-order dispositions may require a longer intervention period to develop measurably.

Table 4
Comparison of SDLRS and CTDI-CV scores at baseline and post-intervention between groups.

Variable	Time point	Intervention group (n = 57), mean ± SD	Control group (n = 58), mean ± SD	P value (between groups)	Group × Time interaction F	P value (interaction)
SDLRS total	Baseline	128.6 ± 16.3	130.2 ± 15.8	0.588	18.42	< 0.001
	Post-intervention	152.3 ± 14.7	138.6 ± 16.1	< 0.001		
Self-management	Baseline	42.8 ± 6.7	43.5 ± 6.4	0.571		
	Post-intervention	52.1 ± 5.9	46.4 ± 6.8	< 0.001		
Desire for learning	Baseline	44.1 ± 7.2	44.8 ± 6.9	0.594		
	Post-intervention	51.7 ± 6.3	47.6 ± 7.1	0.003		
Self-control	Baseline	41.7 ± 6.5	41.9 ± 6.3	0.864		
	Post-intervention	48.5 ± 5.8	44.6 ± 6.7	0.008		
CTDI-CV total	Baseline	274.1 ± 28.7	277.3 ± 27.4	0.532	7.86	0.006
	Post-intervention	298.5 ± 26.3	285.7 ± 27.8	0.011		
Analyticity	Baseline	40.2 ± 5.8	41.0 ± 5.5	0.443		
	Post-intervention	45.6 ± 5.1	42.3 ± 5.7	0.004		
Inquisitiveness	Baseline	39.8 ± 6.1	40.3 ± 5.9	0.653		
	Post-intervention	44.7 ± 5.4	41.8 ± 6.2	0.012		
Self-confidence	Baseline	38.5 ± 6.4	39.1 ± 6.0	0.601		
	Post-intervention	42.8 ± 5.7	40.5 ± 6.3	0.031		

Variable	Time point	Intervention group (n = 57), mean ± SD	Control group (n = 58), mean ± SD	P value (between groups)	Group × Time interaction F	P value (interaction)
Open-mindedness	Baseline	40.7 ± 5.6	41.2 ± 5.4	0.619		
	Post-intervention	43.1 ± 5.3	41.9 ± 5.8	0.187		
Systematicity	Baseline	38.1 ± 5.9	38.6 ± 5.7	0.639		
	Post-intervention	41.2 ± 5.5	39.5 ± 5.9	0.093		
Truth-seeking	Baseline	37.9 ± 6.2	38.4 ± 6.0	0.659		
	Post-intervention	40.3 ± 5.8	39.2 ± 6.1	0.214		
Cognitive maturity	Baseline	38.9 ± 5.7	38.7 ± 5.6	0.845		
	Post-intervention	41.5 ± 5.4	40.1 ± 5.9	0.146		

The longitudinal change trajectories of SDLRS and CTDI-CV total scores are shown in Fig. 3.

Teaching Satisfaction

The results of the teaching satisfaction questionnaire are shown in Table 5. The intervention group reported significantly higher satisfaction scores than the control group in eight of the ten surveyed items, including stimulation of learning interest (4.3 ± 0.7 vs. 3.5 ± 0.9, P < 0.001), improvement of clinical thinking (4.2 ± 0.8 vs. 3.4 ± 0.9, P < 0.001), enhancement of doctor-patient communication skills (4.1 ± 0.8 vs. 3.6 ± 0.9, P = 0.002), usefulness of feedback (4.4 ± 0.6 vs. 3.3 ± 1.0, P < 0.001), facilitation of teamwork (4.0 ± 0.8 vs. 3.4 ± 0.9, P < 0.001), consolidation of theoretical knowledge (4.1 ± 0.7 vs. 3.6 ± 0.8, P = 0.001), willingness to recommend the teaching format (4.3 ± 0.7 vs. 3.5 ± 1.0, P < 0.001), and overall teaching satisfaction (4.2 ± 0.7 vs. 3.5 ± 0.9, P < 0.001). However, two items showed no significant between-group difference. The item assessing perceived study workload appropriateness scored lower in the intervention group than in the control group (3.2 ± 1.1 vs. 3.7 ± 0.9, P = 0.009), indicating that students in the intervention group perceived a heavier workload. Additionally, the item on sufficiency of learning resources did not differ significantly between the groups (3.9 ± 0.8 vs. 3.7 ± 0.9, P = 0.198). The distribution of satisfaction responses is illustrated in Fig. 4.

Table 5
Teaching satisfaction questionnaire scores in the intervention and control groups.

Item	Intervention group (n = 57), mean ± SD	Control group (n = 58), mean ± SD	t value	P value
Stimulation of learning interest	4.3 ± 0.7	3.5 ± 0.9	5.41	< 0.001
Improvement of clinical thinking	4.2 ± 0.8	3.4 ± 0.9	5.12	< 0.001
Enhancement of communication skills	4.1 ± 0.8	3.6 ± 0.9	3.18	0.002
Usefulness of feedback	4.4 ± 0.6	3.3 ± 1.0	7.20	< 0.001
Facilitation of teamwork	4.0 ± 0.8	3.4 ± 0.9	3.85	< 0.001
Consolidation of theoretical knowledge	4.1 ± 0.7	3.6 ± 0.8	3.63	0.001
Appropriateness of study workload	3.2 ± 1.1	3.7 ± 0.9	2.65	0.009
Sufficiency of learning resources	3.9 ± 0.8	3.7 ± 0.9	1.30	0.198
Willingness to recommend	4.3 ± 0.7	3.5 ± 1.0	5.03	< 0.001
Overall satisfaction	4.2 ± 0.7	3.5 ± 0.9	4.72	< 0.001

Discussion

The present study represents one of the first randomized controlled trials to systematically evaluate the combined effect of a flipped classroom and Mini-CEX formative assessment on multidimensional core competencies in undergraduate endocrinology education. The findings demonstrated that this integrated pedagogical model significantly improved theoretical knowledge, clinical reasoning, self-directed learning readiness, and critical thinking disposition compared with conventional lecture-based instruction, while also revealing challenges related to increased student workload. These results provide new evidence for the application of blended teaching strategies in preclinical medical curricula and extend prior research that has primarily focused on either the flipped classroom or Mini-CEX as standalone interventions in clinical clerkship settings.

Enhancement of Theoretical Knowledge and Clinical Skills

The intervention group achieved significantly higher scores on both the theoretical examination and the OSCE, which is consistent with the growing body of evidence supporting the superiority of active learning approaches over passive didactic instruction [16, 17]. A recent systematic review and meta-analysis confirmed that flipped classroom models yield higher knowledge scores than traditional lectures across diverse medical disciplines, with pooled effect sizes remaining significant in both observational and randomized controlled subgroup analyses [16]. In the specific context of endocrinology, earlier work has shown that combining the flipped classroom with problem-based learning improved clinical case analysis scores among medical interns [18]. The present study extends these observations by demonstrating comparable advantages in a preclinical undergraduate cohort, suggesting that the benefits of the flipped classroom are not confined to advanced clinical training. Notably, the OSCE station assessing physical examination skills did not show a statistically significant between-group difference. This finding aligns with reports from Bhai and Poustinchian [19], who found that the flipped classroom approach yielded only modest improvements in physical examination performance among osteopathic medical students, and with the observation by Zhang et al. [20] that the advantage of flipped classroom instruction tends to be more pronounced in cognitive domains than in psychomotor skill acquisition. Physical examination proficiency relies heavily on repeated hands-on practice and direct tactile experience, which may not be sufficiently enhanced through video-based pre-class learning or case discussion alone.

Role of Mini-CEX in Competency Development

A distinctive feature of the current study is the integration of Mini-CEX as a structured formative assessment component within the flipped classroom cycle. The intervention group demonstrated significant improvements across most Mini-CEX dimensions from the first to the final encounter, including medical interviewing, clinical judgment, physical examination, counseling, and organization and efficiency. These findings are in agreement with recent evidence supporting the formative value of Mini-CEX in undergraduate medical education [21, 22]. Iqbal et al. [21] reported that repeated Mini-CEX encounters with structured feedback significantly enhanced clinical skills among fourth-year medical students, and a bibliometric analysis by Sharma et al. [23] has highlighted the growing international interest in Mini-CEX as a tool for competency-based assessment. However, the present study found that improvement in the dimension of humanistic qualities and professionalism did not reach statistical significance, which echoes previous observations that affective and attitudinal competencies are inherently difficult to modify through short-term educational interventions [24]. Professionalism and humanistic qualities are shaped by long-term socialization processes, role modeling, and reflective practice, and a single-semester intervention may be insufficient to produce measurable change in these domains.

The integration of Mini-CEX into the post-class phase created a structured feedback loop that reinforced the cyclical learning process. Lee and Chiu [25] have emphasized that formative assessment in competency-based medical education should function as a developmental tool rather than a grading mechanism, and the present model aligns well with this principle. Each Mini-CEX encounter provided

individualized, timely, and specific feedback that encouraged students to identify their own learning gaps and adjust subsequent study strategies, thereby supporting the transition from passive knowledge reception to active self-regulation.

Promotion of Self-Directed Learning and Critical Thinking

One of the most notable findings of this study was the significant improvement in SDLRS scores observed in the intervention group, with particularly pronounced gains in the subscales of self-management, desire for learning, and self-control. Lu et al. [26] similarly demonstrated that blended teaching methods significantly enhanced self-directed learning abilities among medical students, and the present results extend this evidence by identifying the specific SDLRS domains most responsive to a flipped classroom intervention. The pre-class self-learning tasks, combined with the accountability provided by online quizzes and subsequent Mini-CEX encounters, may have created an environment in which students were compelled to take ownership of their learning process. Ji et al. [27] further observed that the self-directed learning skills fostered by flipped classroom teaching persisted in long-term follow-up assessments, suggesting that the benefits observed in the current study may extend beyond the immediate intervention period.

Regarding critical thinking, the intervention group showed a significant increase in CTDI-CV total scores, with the subscales of analyticity, inquisitiveness, and self-confidence demonstrating the greatest gains. This pattern is consistent with the findings of Huang et al. [28], who reported that Chinese medical students generally possess positive critical thinking dispositions but show uneven development across subscales. The absence of significant improvement in open-mindedness, systematicity, truth-seeking, and cognitive maturity in the present study deserves particular attention. These higher-order dispositions are deeply embedded in cognitive habits and epistemological beliefs that develop gradually over years of education and professional socialization [29]. Nichat et al. [30] similarly noted that while the flipped classroom promotes active engagement and critical thinking, its effect on deeply rooted cognitive dispositions may require sustained exposure over multiple courses or academic years. The case-based discussions and peer questioning activities in the in-class phase of the present model likely stimulated analytical reasoning and intellectual curiosity, but a single-semester intervention may have been insufficient to reshape more stable traits such as truth-seeking or cognitive maturity.

Student Workload and Satisfaction

The teaching satisfaction results revealed a paradox commonly observed in flipped classroom implementations: while students in the intervention group reported significantly higher satisfaction across most dimensions, they also perceived a significantly heavier workload compared with the control group. This finding is consistent with the observations of multiple previous studies [18, 31]. A recent meta-analysis by Li et al. [31] confirmed that increased out-of-class time commitment remains one of the most frequently cited disadvantages of the flipped classroom model, even when students acknowledge its educational benefits. In the present study, the additional time demands included watching pre-recorded video lectures, completing online quizzes, preparing for case discussions, and

participating in Mini-CEX encounters, all of which cumulatively exceeded the time investment required in a traditional lecture format. The item assessing sufficiency of learning resources did not differ between groups, suggesting that the perceived burden was related to time and effort rather than to a lack of available materials. Future implementations should consider strategies to optimize the balance between learning depth and workload, such as shortening video lectures, reducing quiz frequency, or distributing Mini-CEX encounters more evenly across the semester [32].

Limitations and Future Directions

Several limitations should be acknowledged. First, this was a single-center study, and the findings may not be directly generalizable to institutions with different curricula, student populations, or resource availability. Second, although outcome assessors and the data analyst were blinded, the nature of the educational intervention precluded blinding of students and instructors, which may have introduced performance bias or a Hawthorne effect. Third, the study period spanned approximately 20 months, but the actual teaching intervention covered a single endocrinology course cycle, and the long-term retention of the observed improvements was not assessed. Fourth, the SDLRS and CTDI-CV are self-report instruments that may be subject to response bias, and the use of more objective measures of self-directed learning behavior (such as learning analytics data) would strengthen future studies. Fifth, the teaching satisfaction questionnaire was self-designed and has not undergone formal psychometric validation, which limits the interpretability of the satisfaction data. Future multicenter randomized controlled trials with larger sample sizes, longer follow-up periods, validated satisfaction instruments, and objective learning analytics are warranted to confirm and extend the present findings [33]. Additionally, exploring the differential effects of this model across various medical subspecialties and student subgroups (for example, students with different baseline learning styles or academic performance levels) would provide more nuanced guidance for curriculum design [34].

Conclusions

The flipped classroom combined with Mini-CEX formative assessment significantly improved theoretical knowledge, clinical skills, self-directed learning readiness, and critical thinking disposition in undergraduate endocrinology education. However, the increased workload associated with this model warrants attention in future implementations. Multicenter trials with long-term follow-up are needed to validate the sustainability and generalizability of these findings.

Declarations

Trial Registration

Chinese Clinical Trial Registry (registration number ChiCTR2400081456), Registered on May 28 2024.

Ethics approval and consent to participate

The study protocol was approved by the Institutional Ethics Committee of The First Affiliated Hospital of Bengbu Medical University ((approval number: 2024-EC-ENDO-087; approval date: May 20, 2024)). Written informed consent was obtained from all participants prior to enrollment. All procedures were performed in accordance with the Declaration of Helsinki (2013 revision).

Consent for publication

Not applicable. This manuscript does not contain any individual person's data in any form.

Competing interests

The authors declare that they have no competing interests.

Funding

None.

Availability of data and materials

The data used to support the findings of this study are available from the corresponding author upon reasonable request.

Author Contribution

Langen Zhuang: Conceptualization, Investigation, ValidationQianmin Ma: Methodology, Data Curation, SoftwareShengqun Jiang: Writing - Original Draft, Visualization, Formal AnalysisXiaoyan Pei: Resources, Writing - Review & EditingGuoxi Jin: Supervision, Funding AcquisitionQiong Xu: Formal Analysis, Investigation, Writing - Review & Editing

Acknowledgements

Not applicable.

References

1. Anas S, Kyrou I, Rand-Weaver M, et al. The effect of online and in-person team-based learning (TBL) on undergraduate endocrinology teaching during COVID-19 pandemic. BMC Med Educ. 2022;22(1):120.

2. Long C, Tsai K, Napolitano JD, Khawaja RA, Leung AM. The current status of preclinical endocrine education in US Medical Schools. Endocr Pract. 2022;28(8):744–8.
3. Kandra K, Vennam P. The active classroom: a narrative review of active teaching methods and the flipped classroom model in graduate medical education. Cureus. 2025;17(11):e98186.
4. Phillips J, Wiesbauer F. The flipped classroom in medical education: a new standard in teaching. Trends Anaesth Crit Care. 2022;42:4–8.
5. Ai Li E, Wilson CA, Davidson J, et al. Exploring perceptions of competency-based medical education in undergraduate medical students and faculty: a program evaluation. Adv Med Educ Pract. 2023;14:381–9.
6. Bhattacharya S. Competency-based medical education: an overview. Annals Med Sci Res. 2023;2(3):132–8.
7. Naing C, Whittaker MA, Aung HH, Chellappan DK, Riegelman A. (2023). The effects of flipped classrooms to improve learning outcomes in undergraduate health professional education: A systematic review. Campbell Syst reviews, 19(3), e1339.
8. Bhavsar MH, Javia HN, Mehta SJ. Flipped classroom versus traditional didactic classroom in medical teaching: a comparative study. Cureus. 2022;14(3):e23657.
9. Mishall PL, Meguid EMA, Elkhider IA, Khalil MK. The application of flipped classroom strategies in medical education: a review and recommendations. Med Sci Educ. 2025;35(1):531–40.
10. Lu C, Xu J, Cao Y, et al. Examining the effects of student-centered flipped classroom in physiology education. BMC Med Educ. 2023;23(1):233.
11. Batra P, Batra R, Verma N, et al. Mini clinical evaluation exercise (Mini-CEX): a tool for assessment of residents in department of surgery. J Educ Health Promot. 2022;11:253.
12. Zaki HA, Yigit Y, Shaban E, et al. The utility of the Mini-Clinical Evaluation Exercise (Mini-CEX) in the emergency department: a systematic review and meta-analysis evaluating the readability, feasibility, and acceptability of Mini-CEX utilization. Cureus. 2023;15(8):e44443.
13. Motefakker S, Shirinabadi Farahani A, Nourian M, et al. The impact of the evaluations made by Mini-CEX on the clinical competency of nursing students. BMC Med Educ. 2022;22(1):634.
14. Gupta SK, Srivastava T. Assessment in undergraduate competency-based medical education: a systematic review. Cureus. 2024;16(4):e58073.
15. Bashir K, Arshad W, Azad AM, et al. Acceptability and feasibility of mini clinical evaluation exercise (Mini-CEX) in the busy emergency department. Open Access Emerg Med. 2021;13:481–6.
16. Spaic D, Bukumiric Z, Rajovic N, Markovic K, Savic M, Milin-Lazovic J, Milic N. The flipped classroom in medical education: systematic review and meta-analysis. J Med Internet Res. 2025;27:e60757.
17. Shi XY, Yin Q, Wang QW, Lu BR, Li GX, Huang SH, Sun ZG. Is the flipped classroom more effective than the traditional classroom in clinical medical education: a systematic review and meta-analysis. Frontiers in Education. Volume 9. Frontiers Media SA; 2025, January. p. 1485540.

18. Hu X, Zhang H, Song Y, Wu C, Yang Q, Shi Z, Chen W. Implementation of flipped classroom combined with problem-based learning: an approach to promote learning about hyperthyroidism in the endocrinology internship. BMC Med Educ. 2019;19(1):290.
19. Bhai SA, Poustinchian B. The flipped classroom: a novel approach to physical examination skills for osteopathic medical students. J Osteopath Med. 2021;121(5):475–81.
20. Zhang W, Gu J, Li F, Feng F, Chen H, Xing X, Liu L. The effect of flipped classroom in multiple clinical skills training for clinical interns on Objective Structured Clinical Examinations (OSCE). Med Educ Online. 2022;27(1):2013405.
21. Durrani SF, Hameed A, Rani P, Ahmed S, Baqai S, Ali R. Effect of Mini-CEX as formative learning tool for clinical skills in undergraduate medical students in a private medical university in Karachi, Pakistan. BMC Med Educ. 2025;25(1):1356.
22. Sivaraman G, Lakshmanan J, Alexander A, Mahalakshmy T, Raja K, Sabharisan P, Zayapragassarazan Z. Use of Mini-CEX as formative assessment tool in the training of undergraduate medical students in ENT situation analysis and the way forward. Indian J Otolaryngol Head Neck Surg. 2024;76(3):2698–703.
23. Sharma R, Gupta T, Haidery TH, et al. Current trends in mini-clinical evaluation exercise in medical education: a bibliometric analysis. Cureus. 2022;14(12):e33121.
24. Lee GB, Chiu AM. Assessment and feedback methods in competency-based medical education. Ann Allergy Asthma Immunol. 2022;128(3):256–62.
25. Lee GB, Chiu AM. Assessment and feedback methods in competency-based medical education. Ann Allergy Asthma Immunol. 2022;128(3):256–62.
26. Lu SY, Ren XP, Xu H, et al. Improving self-directed learning ability of medical students using the blended teaching method: a quasi-experimental study. BMC Med Educ. 2023;23(1):614.
27. Ji M, Luo Z, Feng D, et al. Short- and long-term influences of flipped classroom teaching in physiology course on medical students' learning effectiveness. Front Public Health. 2022;10:835810.
28. Huang L, Fan APC, Su N, et al. Chinese medical students' disposition for critical thinking: a mixed methods exploration. BMC Med Educ. 2021;21(1):385.
29. Dang H, Liu Y, Chen J, et al. Critical thinking disposition and influencing factors among sophomore pediatric medical students. BMC Med Educ. 2024;24(1):1189.
30. Nichat A, Gajbe U, Bankar NJ, Singh BR, Badge AK, Badge A. (2023). Flipped classrooms in medical education: improving learning outcomes and engaging students in critical thinking skills. Cureus, *15*(11).
31. Shi XY, Yin Q, Wang QW, Lu BR, Li GX, Huang SH, Sun ZG. Is the flipped classroom more effective than the traditional classroom in clinical medical education: a systematic review and meta-analysis. Frontiers in Education. Volume 9. Frontiers Media SA; 2025, January. p. 1485540.
32. Martinsen SSS, Espeland T, Berg EAR, et al. Examining the educational impact of the mini-CEX: a randomised controlled study. BMC Med Educ. 2021;21(1):228.

33. Chikeme PC, Ihudiebube-Splendor CN, Ogbonnaya NP, Mbadugha CJ, Elodi LO. Flipped classroom model versus conventional teaching method: Effects on nursing students' self-directed learning readiness in a research methodology course. Pan Afr Med J. 2024;47:70.
34. Paul A, Leung D, Salas RME, Cruz TE, Abras C, Saylor D, Strowd RE. Comparative effectiveness study of flipped classroom versus online-only instruction of clinical reasoning for medical students. Med Educ Online. 2023;28(1):2142358.

Figures

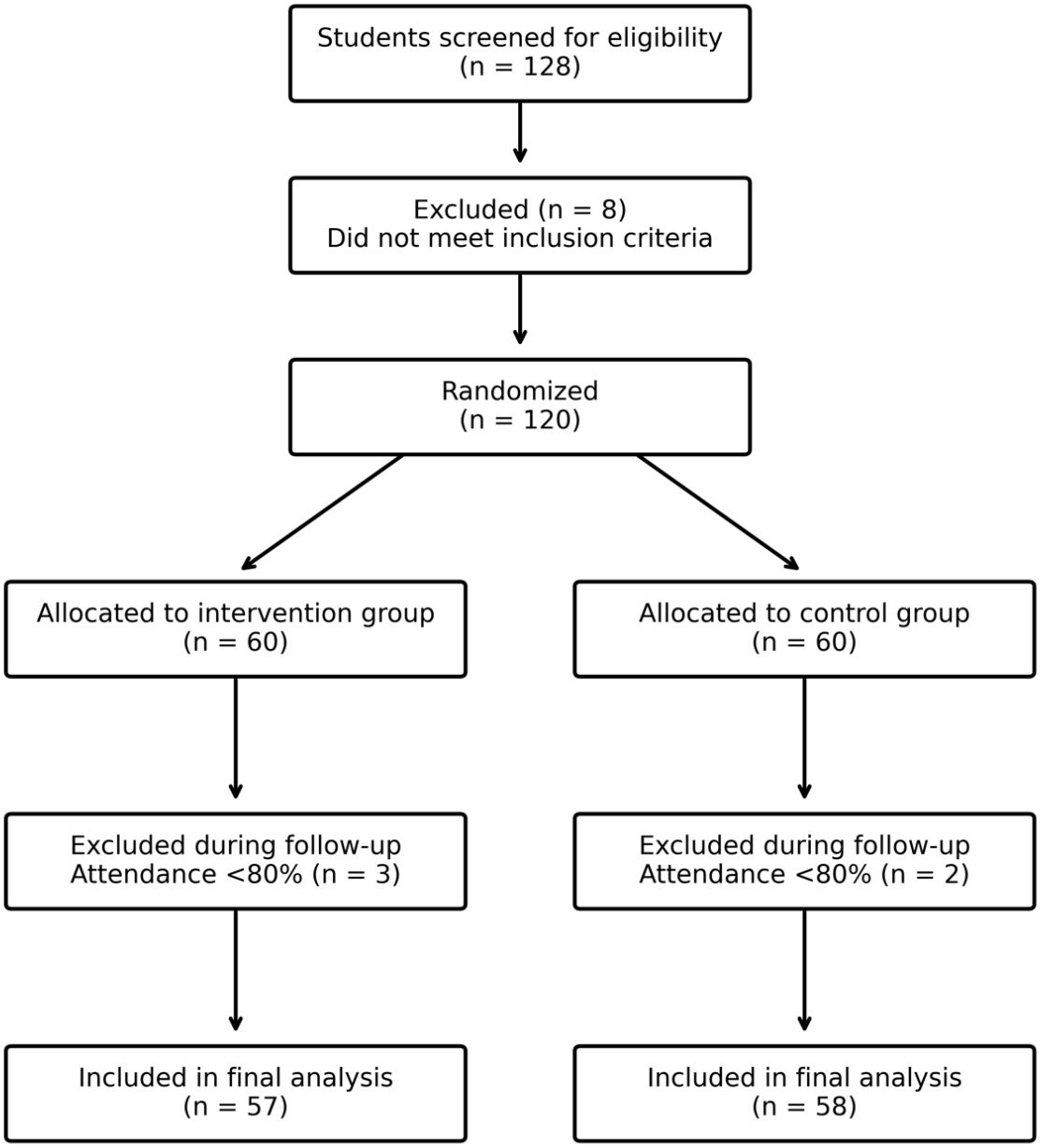


Figure 1

CONSORT flow diagram of participant enrollment, allocation, follow-up, and analysis.

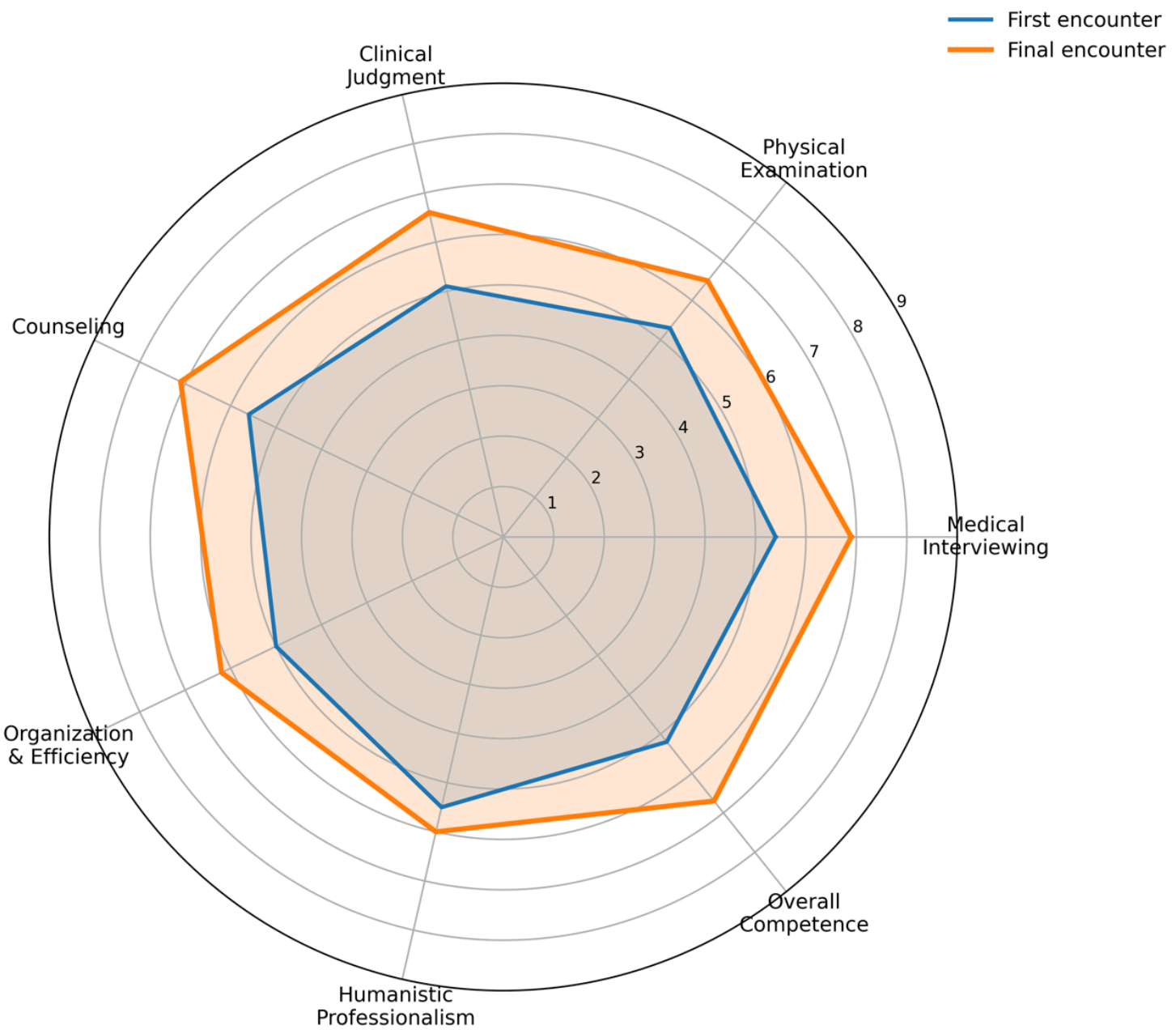


Figure 2

Radar chart illustrating the Mini-CEX dimension score profiles at the first and final encounters in the intervention group.

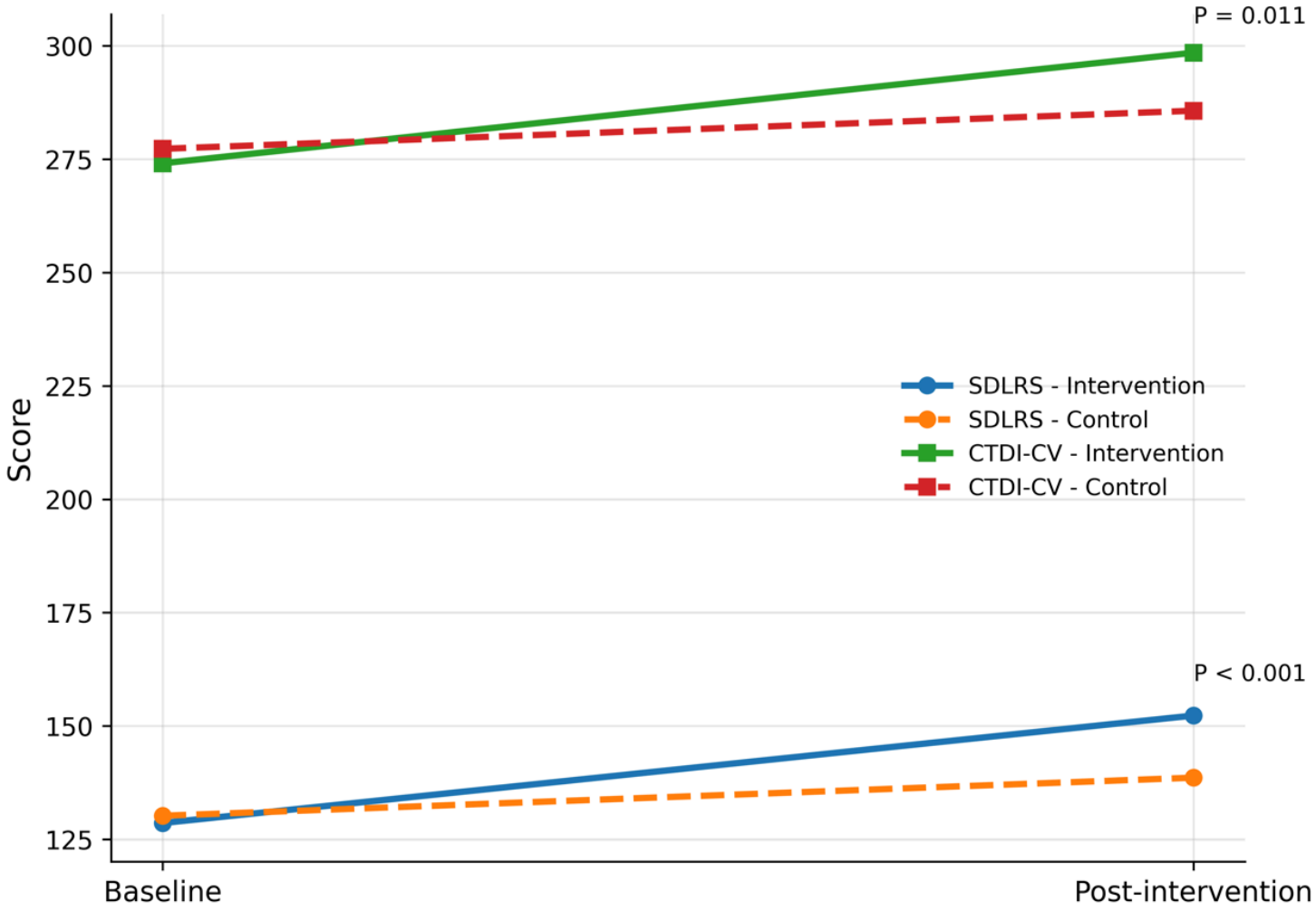


Figure 3

Line graph depicting the change in SDLRS and CTDI-CV total scores from baseline to post-intervention in the intervention and control groups. Error bars represent standard deviations.

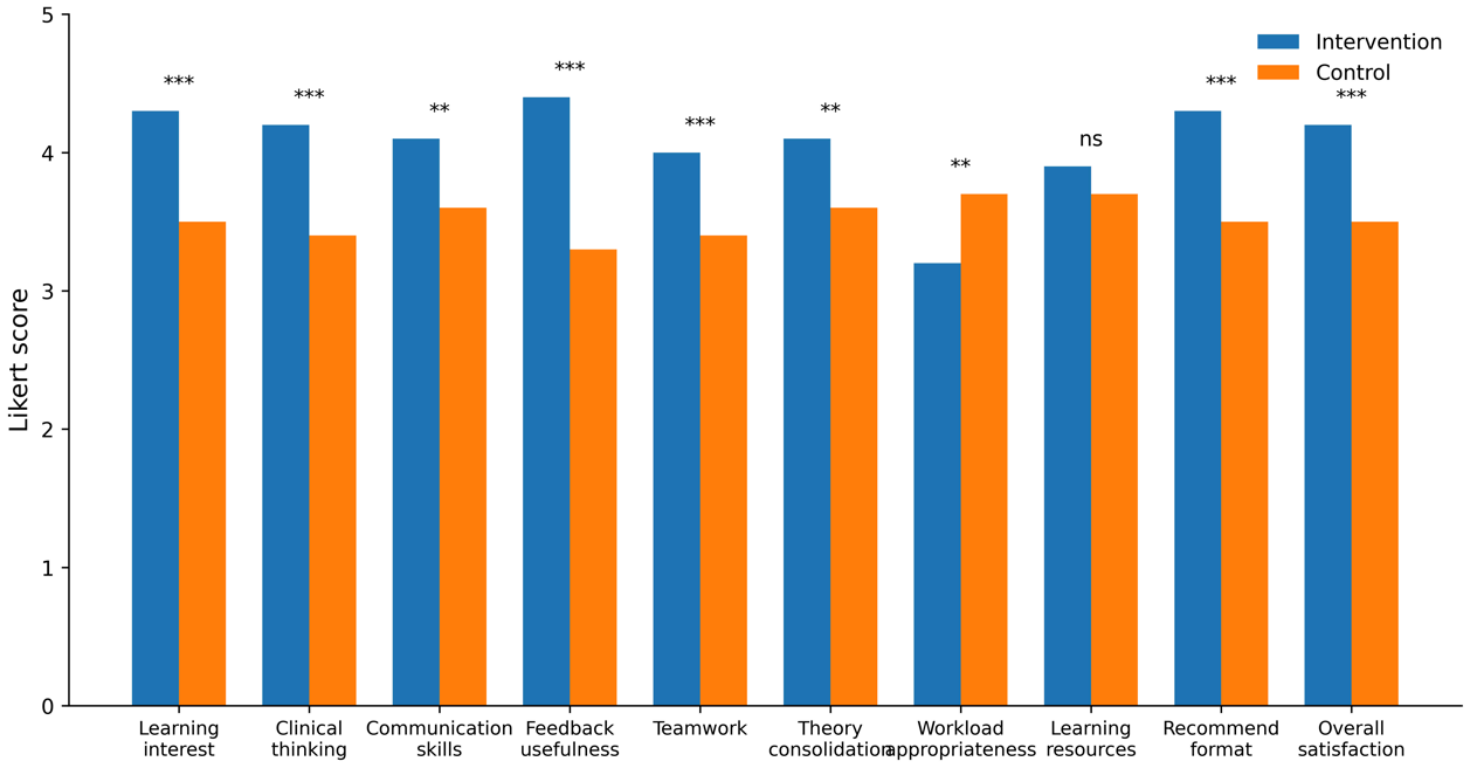


Figure 4

Stacked bar chart showing the distribution of Likert-scale responses (strongly disagree to strongly agree) for each teaching satisfaction item in the intervention and control groups.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- CONSORT2025editablechecklist.docx