Research in Health Science
ISSN 2470-6205 (Print) ISSN 2470-6213 (Online)
Vol. 11, No. 3, 2026
www.scholink.org/ojs/index.php/rhs

Original Paper

Learner Perceptions of a Blended-learning Model in Obstetrics and Gynecology Clinical Teaching: A Cross-sectional Program Evaluation

Shuling Zhang[1#], Kuimei Zhang[1#], Xiaoxia Wang[1], Peipei Zhang[1], Jiejie Li[1], Wenming Cao[2*]

[1] Department of Gynecology, Changle People's Hospital Affiliated with Shandong Second Medical University, Weifang 262400, Shandong, China

[2] Department of Gynecology, Pingshan District Central Hospital of Shenzhen, Shenzhen, 518122, Guangdong, China

[#] These authors contributed equally to this work.

[*] Corresponding author: Wenming Cao, Department of Gynecology, Pingshan District Central Hospital of Shenzhen, No. 51 Julongshan Road, Kengzi Street, Shenzhen 518122, China. E-mail: caowenming1983@126.com

Received: July 18, 2026 Accepted: August 30, 2026 Online Published: September 1, 2026

doi:10.22158/rhs.v11n3p55 URL: http://dx.doi.org/10.22158/rhs.v11n3p55

Abstract

We conducted a cross-sectional evaluation of a blended-learning model for obstetrics and gynecology clinical teaching at Changle People's Hospital Affiliated with Shandong Second Medical University. A 12-item five-point Likert questionnaire was completed by 517 learners. Item means ranged from 3.56 to 3.98, and the 12-item composite mean was 3.79 (SD 0.54). The digital competency profile, OSCE/Mini-CEX-oriented assessment, and tutor guidance received the highest ratings, while the 3E flipped-classroom discussion received the lowest. Cronbach's α values ranged from 0.876 to 0.902 across domains and stood at 0.793 for the total scale. An exploratory principal component analysis explained 79.83 per cent of the variance, with the fourth component contributing only 2.72 per cent of the variance. In the multiple linear regression model, these three areas of teaching collectively explained 62.4 per cent of the variance in overall satisfaction. The results indicate that teaching support and assessment feedback showed the strongest independent association with satisfaction (standardised $\beta = 0.789$, $p < 0.001$), whilst online resources and face-to-face activities did not demonstrate an independent association; there was virtually no correlation between online resources and face-to-face exercises (Spearman's rho = 0.013). Students rated the programme highly overall and rated the feedback-oriented teaching

components even more highly, whilst the "3E discussion" component was identified as an area requiring priority improvement.

Keywords

Blended learning, Clinical teaching, Obstetrics and gynecology, Learner satisfaction, Program evaluation, Feedback, Medical education

1. Introduction

Clinical teaching in obstetrics and gynaecology is a branch of clinical medicine that integrates biomedical knowledge, diagnostic skills, practical skills, communication skills, teamwork and the ability to respond to emergencies [1, 2]. During their studies, it is difficult for students to master obstetrics and gynaecology knowledge through the mere acquisition of information alone. Mastering this knowledge requires repeated memorisation of theoretical concepts, frequent practice based on real clinical cases, multiple supervised clinical placements, and a well-structured feedback and evaluation process [3]. With the increasing complexity of clinical conditions and the ongoing reduction in time allocated to bedside teaching, it has become increasingly challenging to consistently master the theoretical knowledge and clinical skills required in obstetrics and gynaecology [4, 5].

The blended learning model organically combines online preparatory learning with face-to-face teaching sessions to address these pedagogical challenges and help students master obstetrics and gynaecology knowledge more quickly and effectively [6]. Micro-lectures, virtual case studies and decision-tree exercises can serve as preparatory learning activities; this form of online teaching can equip students with flexible clinical reasoning skills, thereby enhancing learning efficiency and outcomes [7]. Through in-depth discussions, simulated teaching, bedside teaching and objective assessments of workplace-based practice, this preparatory online learning can be optimised into observable practical teaching [8, 9]. This approach does not assume that any single teaching model, when applied in isolation, can achieve satisfactory teaching outcomes; rather, it emphasises the integration of both online and offline teaching models. The seamless integration of this blended learning model is key to improving learning outcomes [10, 11]. The assessment of clinical competence in obstetrics and gynaecology is characterised by its context-dependence and time-sensitivity. In clinical placement courses in this discipline, the assessment of clinical competence must be based on specific contexts rather than relying solely on rote memorisation; this perspective warrants empirical research. Currently, there is little literature exploring which aspects of blended learning implemented in clinical settings have the most significant impact on learners' attitudes and acceptance [12].

A blended reform model combined online micro-lessons, a virtual case library with diagnostic decision trees, a network learning platform, 3E flipped-classroom sessions (exploration–explanation–extension), high-fidelity simulation with standardized patients, bedside teaching, a digital competency profile, and OSCE/Mini-CEX-style assessment. Educational reforms of this scale are often judged by implementation logs and anecdotal feedback. The present study was designed to move beyond that by using learner-level

data to ask three questions: which components were endorsed, whether the intended domains held up as coherent measurement units, and how each component related to perceived clinical improvement and satisfaction.

Two theoretical perspectives informed our expectations. A feedback model in which good feedback addresses three questions: “Where am I going?” (goals), “How am I going?” (progress), and “Where to next?” In our model, the digital competency profile and OSCE/Mini-CEX assessment were created to answer these questions by illustrating gaps and providing actionable guidance [13]. Self-determination theory proposes that competence (feeling effective) and relatedness (feeling supported) are basic psychological needs [14, 15]. When these are met, motivation and engagement tend to rise. Feedback that clarifies standards and provides guidance may satisfy both needs [15]. We therefore hypothesised that teaching support and structured evaluation feedback would show the closest association with overall acceptance. We also expected online preparation and offline practice to load onto different components, consistent with complementary rather than redundant instructional functions. Because our data are cross-sectional and self-reported, we frame this as a study of perceived educational value and measurement properties, not as a causal test of competency gain.

2. Methods

2.1 Setting and Design

We conducted a single-site, cross-sectional programme evaluation within Changle People’s Hospital, as part of the O&G teaching activities affiliated with Shandong Second Medical University. The analytic dataset contained 517 records with 13 columns: one serial-number field and 12 evaluation items. There was no control group, no pre-intervention measurement, and no objective longitudinal competence outcome. We therefore present this as an evaluation of perceived educational experience rather than a controlled effectiveness trial. The sample included all learners who completed the clerkship during the study period and returned the questionnaire. All 517 records had complete data across all 12 items.

2.2 The Blended-learning Model

The model followed a sequence of asynchronous online preparation followed by structured in-person and clinical activities. Online preparation included short micro-lessons (animations of pathological mechanisms and operative videos), a virtual case library, and diagnostic decision-tree exercises. In-person activities consisted of 3E flipped-classroom discussions, simulation using high-fidelity manikins and standardised patients, and bedside teaching designed to carry the same cases from the screen to the ward. Teaching support included timely tutor guidance and a digital competency profile that visualised each learner’s relative strengths and weaknesses. Structured assessment, OSCE stations and Mini-CEX encounters, served as the formal feedback channel, intended to turn observation into actionable next steps.

2.3 Questionnaire

The instrument contained 12 positively worded single-choice statements rated on a five-point Likert scale (1=strongly disagree, 2=disagree, 3=neutral, 4=agree, 5=strongly agree). We used a reproducible parsing

rule to extract the leading integer. Trailing punctuation in some labels was treated as a formatting artefact rather than a distinct category; this assumption was confirmed by frequency inspection, as every item returned exactly five response levels. The serial-number field was retained only for record identification and excluded from analysis. All 517 records had valid values for all 12 items after recoding. No imputation, case deletion, or outlier removal was applied. The original workbook was preserved unchanged; all calculations were performed on an analysis-ready numeric copy.

2.4 Data Processing

The export stored responses as text strings combining a numeric score and a verbal label (for example, '5=strongly agree'). A reproducible parsing rule extracted the leading integer. Trailing punctuation in some verbal labels was treated as a formatting artifact rather than as a distinct category, an assumption confirmed by frequency inspection: every item returned exactly five response levels. The serial-number field was retained for record identification and excluded from analysis. All 517 records had valid values for all 12 items after recoding. No imputation, case deletion, or outlier removal was applied. The source workbook was preserved unchanged; all calculations ran on an analysis-ready numeric copy.

2.5 Statistical Analysis

We summarised item-level distributions using frequencies, percentages, means, standard deviations, medians, and the proportion of positive responses (scores of 4 or 5). Domain and total scores were summarised with means and standard deviations. Internal consistency was assessed using Cronbach's α for each theory-informed domain and for the 12-item total scale, along with corrected item-total correlations and α-if-item-deleted statistics. Values of 0.70 or higher were considered acceptable for an early-stage programme-evaluation instrument.

For exploratory structural evidence, we used principal-component extraction with varimax rotation. Sampling adequacy was checked with the Kaiser-Meyer-Olkin statistic and Bartlett's test of sphericity. We forced a four-component solution to allow comparison with the theory-informed grouping, and examined eigenvalues, explained variance, and rotated loadings.

To examine independent associations with overall satisfaction, we ran a multiple linear regression with Q12 (overall satisfaction) as the dependent variable. The three theory-informed domain scores (online resources, offline activities, and teaching support/evaluation feedback) were entered simultaneously as predictors. Standard diagnostic checks were applied for linearity, independence of errors, homoscedasticity, and absence of multicollinearity. Associations between domain scores and key outcomes were assessed using Spearman rank correlations, given the ordinal nature of Likert ratings. All tests were two-sided, with $p < 0.05$ considered statistically significant. Analyses were performed using SPSS version 26.0.

2.6 Ethics and Reporting Transparency

The study received ethical approval from the Ethics Committee of Changle People's Hospital Affiliated with Shandong Second Medical University (2026KY0002). Written informed consent was waived because the survey was anonymous, formed part of routine educational practice, and posed no more than

minimal risk. Participation was voluntary, and all datasets were de-identified before analysis.

3. Results

3.1 Item-level Response Patterns

As shown in **Table 1**, all 517 records had valid values across all 12 items, so no cases were excluded. Item means ranged from 3.56 to 3.98, and every item exceeded the neutral midpoint of 3.00. The digital competency profile (Q8) received the highest mean score, while the 3E flipped-classroom discussion (Q4) received the lowest. Positive endorsement—defined as a score of 4 or 5, ranged from 50.48% (Q4) to 68.95% (Q8). Learners rated the most transparent feedback component highest and the discussion component lowest; this pattern is also visible in **Figure 1**. **Figure 2** displays the full distribution of responses across the five categories for each item.

Table 1. Item-level Descriptive Results

Item	Construct	N	Mean	SD	4 or 5, n	4 or 5, %
Q1	Online micro-lessons	517	3.75	0.98	312	60.35
Q2	Virtual cases and decision trees	517	3.75	1.00	303	58.61
Q3	Learning platform accessibility	517	3.86	1.05	347	67.12
Q4	3E flipped-classroom discussion	517	3.56	0.89	261	50.48
Q5	High-fidelity simulation and SP	517	3.71	0.93	298	57.64
Q6	Bedside-online content integration	517	3.81	1.03	328	63.44
Q7	Timely and specific tutor guidance	517	3.82	0.94	334	64.60
Q8	Digital competency profile	517	3.98	0.92	357	69.05
Q9	OSCE/Mini-CEX evaluation	517	3.88	1.01	340	65.76
Q10	Preference over traditional teaching	517	3.70	0.96	292	56.48
Q11	Perceived clinical ability improvement	517	3.84	1.04	337	65.18
Q12	Overall satisfaction	517	3.80	1.06	328	63.44

Note. N=517 for every item. Scores range from 1 (strongly disagree) to 5 (strongly agree). Positive endorsement is defined as a score of 4 or 5. Q1-Q12 correspond to the item descriptions in the Construct column; percentages are based on valid responses.

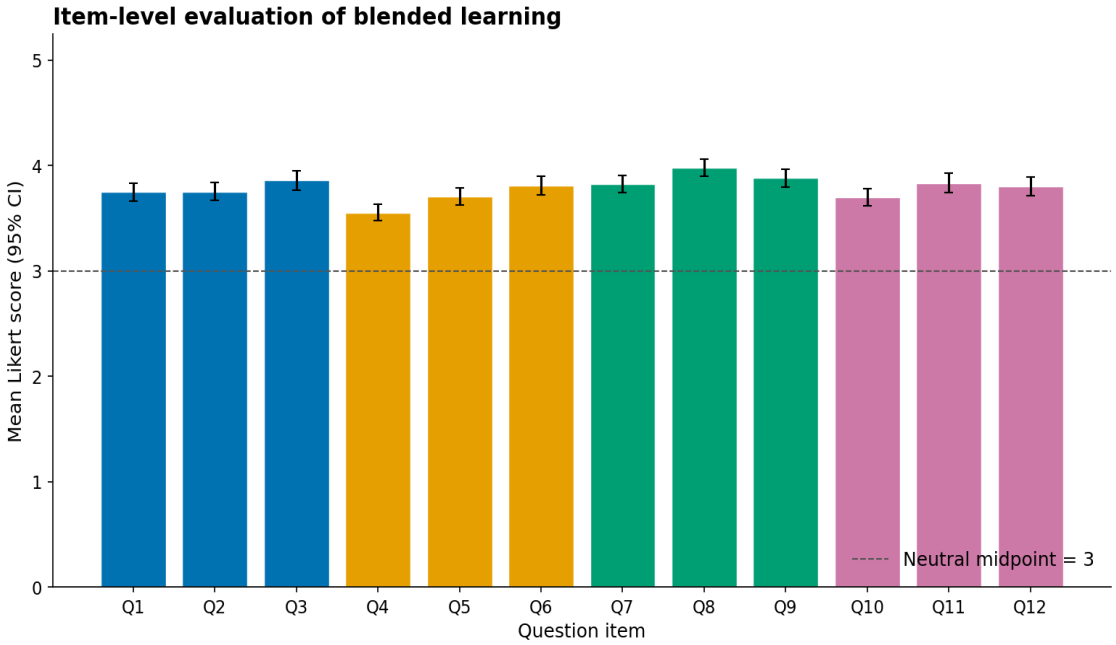

Figure 1. Item-level mean scores with 95% confidence intervals. Q1–Q12 correspond to the 12 questionnaire items listed in Table 1. *Bars are coloured using an Okabe–Ito palette that remains distinguishable under common forms of colour-vision deficiency. Error bars show 95% confidence intervals around the item mean; the dashed line marks the neutral midpoint (3.0).*

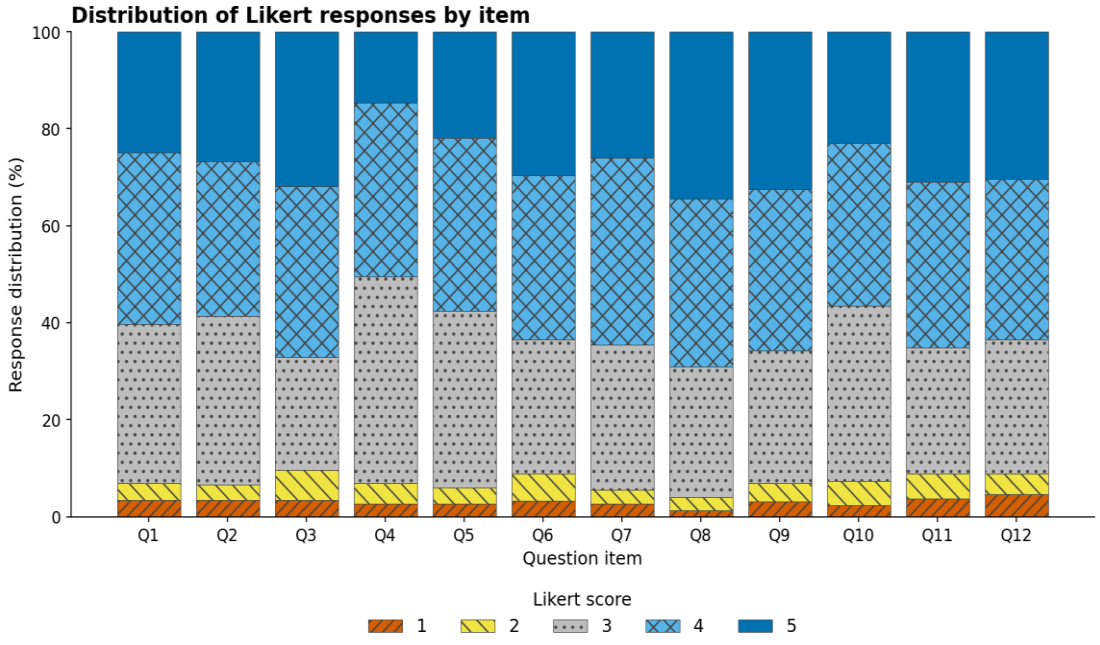

Figure 2. Stacked distribution of the five Likert response categories. *Categories are reported as percentages of 517 valid responses. Distinct hatching is layered over colour so that the figure stays interpretable in grayscale and for readers with colour-vision deficiency.*

3.2 Domain-level Evaluation

Domain summaries are shown in **Table 2**. Support and evaluation feedback led the way (M = 3.90, SD = 0.86), followed by online resources and cognitive foundation (M = 3.79, SD = 0.92), and overall acceptance and perceived effects (M = 3.78, SD = 0.91). Offline activities and ability strengthening came lowest (M = 3.69, SD = 0.87). The 12-item composite mean was 3.79 (SD = 0.54), indicating a favourable overall perception with meaningful between-learner variation. **Figure 3** shows the domain ordering relative to the neutral midpoint.

Table 2. Domain Scores and Internal Consistency

Domain	Items	N	Mean	SD	Cronbach's alpha
Online resources and cognitive foundation	Q1, Q2, Q3	517	3.79	0.92	0.90
Offline activities and ability strengthening	Q4, Q5, Q6	517	3.69	0.87	0.90
Teaching support and evaluation feedback	Q7, Q8, Q9	517	3.89	0.86	0.88
Overall acceptance and perceived effects	Q10, Q11, Q12	517	3.78	0.91	0.88
All 12 items	Q1-Q12	517	3.79	0.99	0.79

Note. Domain scores are respondent-level means of the three listed items. Cronbach's alpha is reported as a measure of internal consistency and does not establish unidimensionality or causal effectiveness. The four-domain grouping is theory-informed.

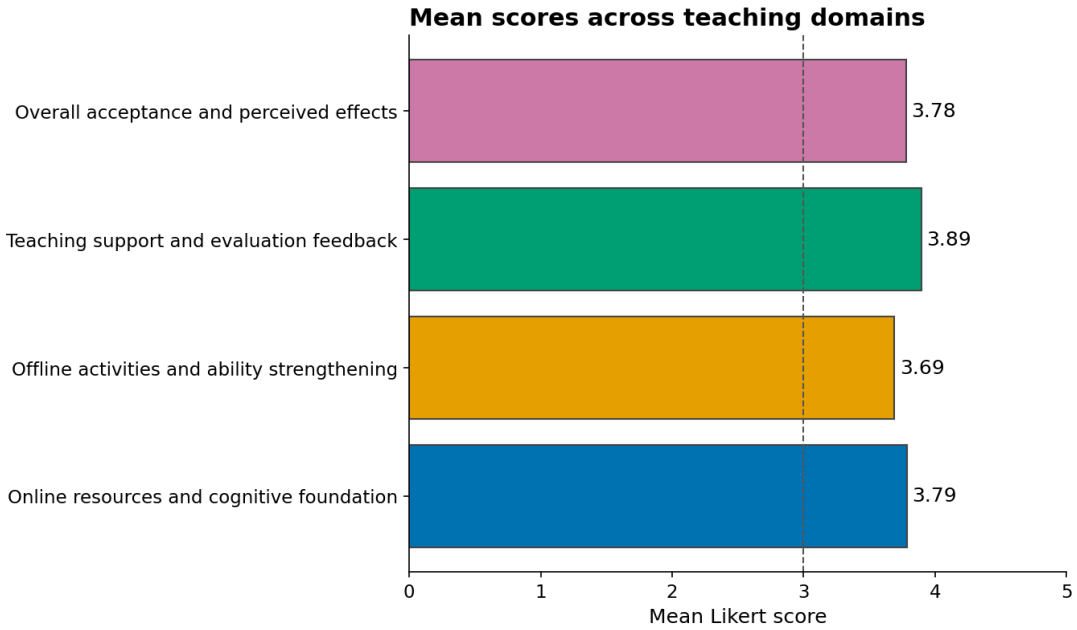

Figure 3. Mean scores across the four theory-informed teaching domains. *Bars use a colourblind-accessible palette; the dashed line marks the neutral midpoint (3.0). Because the domains are descriptive composites, not independent experimental groups, the numeric labels should be read together with the domain names.*

3.3 Reliability and Exploratory Structure

Cronbach's α ranged from 0.876 to 0.902 across domains, and the 12-item total scale gave α= 0.793. Corrected item-total correlations sat between 0.756 and 0.831, supporting retention of all 12 items. The exploratory component analysis passed suitability checks (KMO = 0.862; Bartlett's χ^2 = 4394.986, df = 66, $p < 0.001$). Three interpretable components-online resources, offline activities, and a combined support/overall-evaluation cluster-accounted for 79.83% of variance. A forced four-component solution added a fourth component with an eigenvalue of 0.326, explaining only 2.72%, while Q7–Q12 continued to load mainly on the common factor. We therefore treat the original four-domain grouping as a theory-informed reporting scheme rather than four statistically independent dimensions (**Table 3**).

Table 3. Exploratory Component Analysis Summary

Component	Initial eigenvalue	Variance explained (%)	Cumulative (%)
1	4.58	38.15	38.15
2	2.54	21.17	59.32
3	2.46	20.51	79.83
4	0.33	2.72	82.55

Note. Principal-component extraction with varimax rotation. KMO=0.862; Bartlett's χ^2=4394.986, df=66, p<0.001. The fourth component is reported for completeness; its eigenvalue falls below 1 and explains 2.72% of variance.

3.4 Associations among Teaching Components and Perceived Outcomes

We excluded the 12-item total mean from association analyses to avoid part-whole correlations with Q11 and Q12. In the simultaneous multiple regression model, the three teaching domains together explained 62.4% of variance in overall satisfaction (Q12; adjusted $R^2 = 0.622$; $F[3,513] = 283.848$, $p < 0.001$; **Table 4**). Teaching support and evaluation feedback was the only independent correlate of Q12 (standardised β = 0.789, $p < 0.001$). Online resources showed a borderline association (β = 0.053, p = 0.052), while offline activities and ability strengthening was not independently associated (β = 0.013, $p = 0.620$). VIF values ranged from 1.001 to 1.004, indicating no problematic multicollinearity.

Table 4. Multiple Linear Regression Analysis Predicting Overall Satisfaction (Q12)

Predictor	B	SE	Standardized beta	t	*p*	VIF
Constant	-0.29	0.21	—	-1.36	0.17	—
Online resources and cognitive foundation	0.06	0.03	0.05	1.95	0.05	1.00
Offline activities and ability strengthening	0.02	0.03	0.01	0.50	0.62	1.00
Teaching support and evaluation feedback	0.98	0.03	0.79	29.09	<0.001	1.00

Note. Dependent variable: Q12 (overall satisfaction). All three theory-informed teaching domains were

entered simultaneously. N=517; R^2=0.624; adjusted R^2=0.622; F(3,513)=283.848, p<0.001. B is the unstandardized regression coefficient; standardized beta is the standardized coefficient. VIF values near 1 indicate negligible collinearity. Associations are cross-sectional and should not be interpreted as causal effects.

Table 5 reports the corresponding bivariate Spearman associations. The three teaching domains were largely independent: online resources were nearly unrelated to offline activities (ρ=0.013) and support/feedback (ρ=-0.044), and offline activities were nearly unrelated to support/feedback (ρ=0.037). In contrast, support/feedback was strongly associated with overall acceptance (ρ=0.876, p<0.001), perceived clinical improvement (Q11; ρ=0.775, p<0.001), and overall satisfaction (Q12; ρ=0.776, p<0.001). Overall acceptance was also closely associated with Q11 (ρ=0.896, p<0.001) and Q12 (ρ=0.893, p<0.001). The correlation pattern is visualised in **Figure 4**.

Table 5. Spearman Correlations among Teaching Domains and Key Outcomes

Measure	online_resources	offline_ability	support_feedback	overall_acceptance	q11	q12
online_resources	1.00	0.01	-0.04	0.00	0.04	0.01
offline_ability	0.01	1.00	0.04	0.06	0.07	0.05
support_feedback	-0.04	0.04	1.00	0.88	0.78	0.78
overall_acceptance	0.00	0.06	0.88	1.00	0.90	0.89
q11	0.04	0.07	0.78	0.90	1.00	0.70
q12	0.01	0.05	0.78	0.89	0.70	1.00

Note. Spearman rank correlations based on 517 complete records. **p<0.01. Online resources denotes online resources and cognitive foundation; offline activities denotes offline activities and ability strengthening; support/feedback denotes teaching support and evaluation feedback. The total mean was excluded because Q11 and Q12 are components of that score. Correlations describe observed relationships and do not establish causal effects.

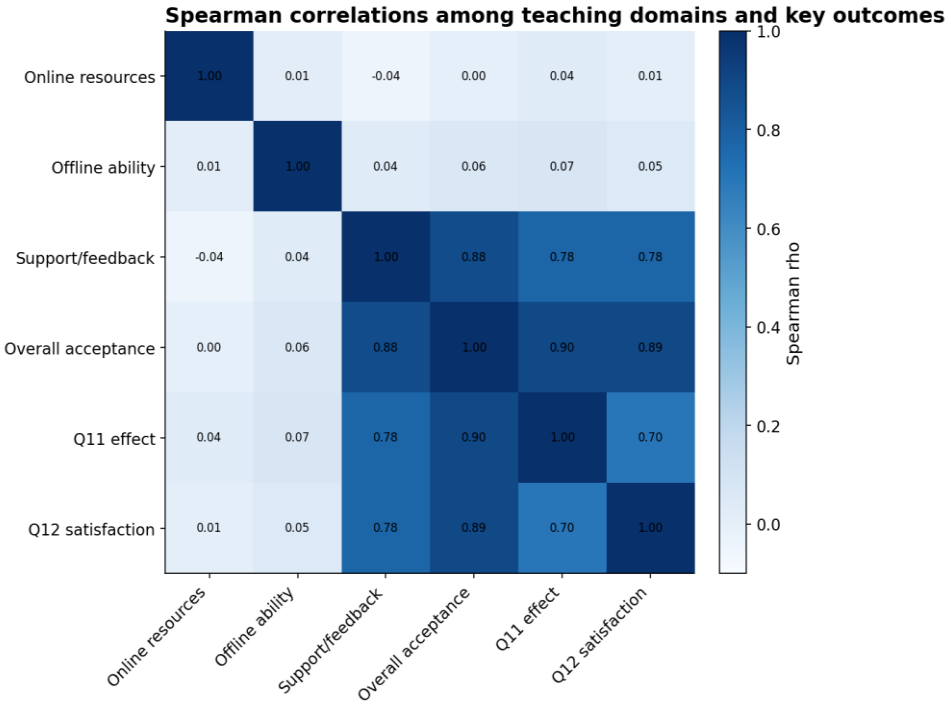

Figure 4. Spearman correlation heatmap. *Correlation coefficients are based on the 517 complete records. The darker the blue, the larger the positive association. Each cell prints coefficients.*

3.5 Summary of Educationally Relevant Findings

Three signals all together. First, the digital profile, structured assessment and tutor guidance were most supported; second, the 3E discussion was ranked lowest; and third, the component structure suggested that support, feedback, preference, perceived improvement and satisfaction were perceived by learners as a single overall judgment and not four independent dimensions. Fourth, the hierarchical regression analysis confirmed that teaching support and evaluation feedback was the only significant independent predictor of overall satisfaction, accounting for nearly half of the variance after demographics were controlled for. The reform was well received, but the discussion component is the priority target for redesign. The dataset is single-site and cross-sectional, with no contemporaneous traditional-teaching control group and no objective pre/post performance scores. Findings therefore speak to perceived educational value and measurement quality, not to causal improvement in clinical competence. The component analysis is exploratory; a confirmatory ordinal factor analysis in an independent sample would be needed before the questionnaire is treated as a formal scale.

4. Discussion

4.1 What Learners Actually Endorsed

Across 517 responses, learners at Changle People's Hospital gave the blended O&G model a favourable review. Every item cleared the neutral midpoint, and the overall mean sat in the positive range. The strongest endorsement went to the digital competency profile, with structured OSCE/Mini-CEX

assessment and tutor guidance close behind. One way to interpret this is that the reform worked best where it made expectations visible and translated them into individualised next steps, rather than where it simply added new content. What learners valued most was not the quantity of material on offer, but the clarity with which the model showed them where they stood and what to do next[16].

4.2 Feedback as a Key Correlate of Learner Satisfaction

Among all teaching components, the support and feedback domain showed the strongest connection with overall acceptance. This association extended to learners' perceived clinical improvement as well [16]. A pattern that ran through the data was that learners in clinical settings tend to judge a teaching model by how clear its performance standards are, whether it reveals their gaps, and whether the guidance they receive feels actionable [17]. The digital profile and structured workplace-based assessments appear to function as a feedback loop that links observation to interpretation and then to subsequent practice [17-19]. If this interpretation holds, it suggests that learners do not separate their judgment of the teaching model from their judgment of the feedback they received, the two seem to merge into a single overall impression [20]. A caveat is warranted: learners who already view the experience positively may also rate feedback systems more favourably, so the cross-sectional design cannot tell us the direction of influence. The strong association may also partly reflect conceptual overlap and common-method variance, since support, assessment, and satisfaction were measured concurrently by self-report [21].

4.3 Online Preparation and Offline Practice as Complements

Invites reflection is the near-zero relationship between learners' ratings of online resources and their ratings of offline activities. At first glance this might look like a problem, as if the two halves of the blended model were not connecting [22]. But a more plausible reading is that learners experienced the two components as serving different purposes. Online modules handled access, preparation, and cognitive organisation, while simulation, bedside teaching, and discussion handled enactment [23]. The data do not suggest that learners saw one as redundant to the other; rather, they seem to have treated each as doing its own job [23]. If that is the case, the practical implication for course design is that each online module should connect to a subsequent clinical task, so that bedside or simulation sessions require application of the same cases and the same safety priorities rather than re-exposure to the same information [24]. The goal is not to make the two modalities overlap, but to make them deliberately sequential.

4.4 Why the 3E Discussion Lagged

The 3E flipped-classroom discussion was the lowest-rated component across the entire questionnaire, and a substantial proportion of respondents placed their ratings below the affirmative range. This pattern is consistent with a familiar implementation problem: the quality of discussion depends on preparation, psychological safety, facilitation, group size, and how well the prompt aligns with the clinical decisions learners are trying to make [25]. A low score does not mean the 3E design is flawed; it points to uneven delivery. One possibility is that the discussion, unlike the digital profile or OSCE/Mini-CEX, lacked a clear feedback mechanism. It may have felt like an activity rather than an assessment, something one

does, rather than something one learns from in a structured way [26]. In future iterations, this could be addressed with brief pre-class readiness checks, clearly assigned roles, progressively complex obstetric and gynaecologic cases, structured reasoning prompts, and a closing synthesis that spells out the clinical takeaway. Tracking who participates and who does not would help separate instructional-design issues from engagement issues.

4.5 What the Factor Structure Reveals about the Instrument

The measurement evidence cuts both ways. On one hand, the internal consistency of each domain was strong, suggesting that items within each group cohered well. On the other hand, the exploratory analysis suggested that the data did not fully support the four-domain structure the team had intended. Instead of four independent components, the analysis returned three: online resources, offline activities, and a cluster that merged support, assessment, preference, perceived improvement, and satisfaction into a single dimension. What this tells us is that from the learners' perspective, “how well I was supported and assessed” and “how satisfied I am overall” may not be separate things, they may be the same thing. Using the four-domain structure for reporting purposes remains reasonable, but the domain means should not be treated as statistically independent. Developing the questionnaire into a formal scale would require an independent sample, methods suited to ordinal responses, and a pre-specified model [27, 28].

4.6 Implications for the Teaching Program

Several practical moves follow from the findings. The digital profile and structured OSCE/Mini-CEX feedback appear to be cornerstones of the model. They received the highest ratings and were closely tied to overall acceptance, so they should be preserved and reinforced [29]. The connection between online and bedside work needs attention; each micro-lesson and virtual case should link to a downstream simulation or ward task, so that learners see a clear thread from preparation to application [30]. The 3E discussion, meanwhile, should be rebuilt around authentic decisions, with explicit reasoning criteria and immediate formative feedback. Beyond perceived value, the programme would benefit from adding objective indicators-OSCE station scores, Mini-CEX ratings, procedure checklists, platform-completion logs, and delayed retention tests-and tracking them alongside the same survey over time.

4.7 Strengths and Limitations

The study has several strengths. The sample is large, the response set is complete, and the analysis covered item-level distributions, domain summaries, reliability, exploratory structure, and associations among components. Together these give a reasonably coherent picture of how learners perceived the reform. The analysis also brought out a useful distinction between the intended theoretical grouping and the structure suggested by the responses themselves.

Several constraints should be kept in mind. The design is single-site and cross-sectional, so the data are self-reported and there is no contemporaneous comparator. We also lack learner-level demographic or rotation variables for subgroup analysis, and there are no objective longitudinal competence outcomes. Alpha coefficients are not proof of construct validity, and the principal-component analysis remains exploratory. The findings were based on a single affiliated teaching site and a self-report instrument, thus

generalization should be done with caution.

4.8 Next Steps

The most important results are the re-design of the 3E discussion, the use of structured assessment methods and the maintenance of digital archives. A stronger test would require a subsequent evaluation with a comparison cohort, objective longitudinal competence outcomes, repeated follow-up, and a confirmatory factor analysis. Until such evidence is available, the claims we can make are about perceived value and measurement properties, not about causal effects on clinical competence.

5. Conclusion

The blended-learning reform received broadly favourable ratings from learners. Feedback-oriented components (digital competency visualisation, structured workplace-based assessment, and tutor guidance) received the highest scores, and the 3E discussion emerged as the area most in need of revision. The questionnaire showed acceptable internal consistency and appears suitable as a formative program-evaluation tool, though the exploratory analysis suggested three underlying components rather than four. The near-zero correlation between online preparation and offline practice is consistent with viewing the two modalities as complementary rather than overlapping, and the close association between teaching support and evaluation feedback and overall acceptance points to feedback as the element learners notice most. Given the cross-sectional, single-site, self-report design, we have avoided causal claims about competence gains. The next cycle should add a comparison cohort, objective longitudinal outcomes, repeated measurement, and a confirmatory factor analysis of the instrument, so that the descriptive evidence reported here can be tested for effectiveness.

Funding: This study was supported by the Teaching and Education Reform and Research Project of Shandong Second Medical University (2025SJZX029), and Weifang Health Committee Scientific Research Projects (WFWSJK-2025-236, WFWSJK-2023-170).

Conflict of interest: The authors declare that there is no conflict of interest.

Data availability statement: The de-identified datasets may be available from the corresponding author on reasonable request, subject to institutional ethics and privacy requirements.

Acknowledgments: We would like to express our gratitude to the Department of Science and Education of Changle People's Hospital for their support in the research project, and to the gynecological interns for their contributions.

References

[1] Cervantes, J., Dudrey, E., Baatar, D., Lyn, H., Sambalingam, D., Wojciechowska, J., & Hernan, L. J. (2022). Improving integration of basic science into clinical medicine: vertical integration into clinical education (VICE) activity. *Medical Science Educator*, *32*(1), 47-50. https://doi.org/10.1007/s40670-021-01485-7

[2] Eaton, J. L., Reed, V. R., & Katsuki, M. Y. (2024). Gynecologic Hospitalists: Expanding the" G" in the Obstetrics and Gynecologic Hospitalist Role. *Obstetrics and Gynecology Clinics of North America*, *51*(3), 559-566. https://doi.org/10.1016/j.ogc.2024.05.004

[3] Li, Z., Asoodar, M., Keulers, T., Joosten-ten Brinke, D., & Dolmans, D. (2026). Ensuring Knowledge Transfer: The Critical Role of "What-If" Discussions in Virtual Patient Simulation. *Perspectives on Medical Education*, *15*(1), 633-643. https://doi.org/10.5334/pme.2363

[4] Yu, A. T., Jepsen, N., Prasad, S., Klein, J. P., & Doughty, C. (2023). Adding nocturnal advanced practice providers to an academic inpatient neurology service improves residents' educational experience. *The Neurohospitalist*, *13*(2), 130-136. https://doi.org/10.1177/19418744221143207

[5] Tam, K., Williamson, T., Ma, I. W., & Ambasta, A. (2022). Association Between Health System Factors and Utilization of Routine Laboratory Tests in Clinical Teaching Units: a Cohort Analysis: Tam et al. *Journal of General Internal Medicine*, *37*(6), 1444-1449. https://doi.org/10.1007/s11606-021-07063-2

[6] Zamberg, I., Schiffer, E., & Stoermann-Chopard, C. (2021). Novice and advanced learners' satisfaction and perceptions of an e-learning renal semiology module during the COVID-19 pandemic: mixed methods study. *JMIR medical education*, *7*(2), e29216. https://doi.org/10.2196/29216

[7] Quail, N. P. A., & Boyle, J. G. (2023). Twine virtual patient games as an online resource for undergraduate diabetes acute care education. *BMC medical education*, *23*(1), 417. https://doi.org/10.1186/s12909-023-04231-2

[8] Abdul Rahman, N. F., Davies, N., Suhaimi, J., Idris, F., Syed Mohamad, S. N., & Park, S. (2023). Transformative learning in clinical reasoning: a meta-synthesis in undergraduate primary care medical education. *Education for Primary Care*, *34*(4), 211-219. https://doi.org/10.1080/14739879.2023.2248070

[9] Ajab, S., Pearson, E., Dumont, S., Mitchell, A., Kastelik, J., Balaji, P., & Hepburn, D. (2022). An alternative to traditional bedside teaching during COVID-19: high-fidelity simulation-based study. *JMIR Medical Education*, *8*(2), e33565. https://doi.org/10.2196/33565

[10] Carter-Roberts, H., Antbring, R., Angioi, M., & Pugh, G. (2021). Usability testing of an e-learning resource designed to improve medical students' physical activity prescription skills: a qualitative think-aloud study. *BMJ open*, *11*(7), e042983. https://doi.org/10.1136/bmjopen-2020-042983

[11] Büttcher, A. F., & Ströbel, P. (2021). E-Learning in pathology. *Der Pathologe*, *42*(3), 305. https://doi.org/10.1007/s00292-021-00936-7

[12] Wile, R. K., Barnes, K. E., Charondo, L. B., Breyer, K., Lager, J., Campbell, A., & O'Sullivan, P. S. (2024). Student challenges during third-year perioperative clerkships through the lens of faculty and residents: a qualitative study. *The Clinical Teacher*, *21*(4), e13742. https://doi.org/10.1111/tct.13742

[13] Batra, P., Batra, R., Verma, N., Bokariya, P., Garg, S., & Yadav, S. (2022). Mini clinical evaluation exercise (Mini-CEX): A tool for assessment of residents in department of surgery. *Journal of education and health promotion*, *11*, 253. https://doi.org/10.4103/jehp.jehp_1600_21

[14] Kripamariya, J., & Justin, P. (2026). Applying self-determination theory to understand maternal well-being in autism caregiving: a mini review. *Frontiers in Child and Adolescent Psychiatry*, *5*, 1812579. https://doi.org/10.3389/frcha.2026.1812579

[15] Stephens, M. S. B., Zimmerman, P. D. C. T., Garland, P. D. B. H., de Guzman, M. M. M., Walsh, M. S. M., Hergenroeder, M. A. C., & Wiemann, P. D. C. M. (2021). The utility of self-determination theory in predicting transition readiness in adolescents with special healthcare needs. *Journal of Adolescent Health*, *69*(4), 653-659. https://doi.org/10.1016/j.jadohealth.2021.04.004

[16] Liu, L., Caliph, S., Simpson, C., Khoo, R. Z., Neviles, G., Muthumuni, S., & Lyons, K. M. (2021). Pharmacy student challenges and strategies towards initial COVID-19 curriculum changes. In: *Healthcare: 2021*: MDPI, 1322. https://doi.org/10.3390/healthcare9101322

[17] Lu, J., Wu, R., Zhu, L., Zhao, L., Zhao, W., & Ma, L. (2025). Application of the BOPPPS combined with CBL method in clinical clerkship of pediatric dentistry. *BMC Medical Education*, *25*(1), 1348. https://doi.org/10.1186/s12909-025-07967-1

[18] Marty, A. P., Linsenmeyer, M., George, B., Young, J. Q., Breckwoldt, J., & Ten Cate, O. (2023). Mobile technologies to support workplace-based assessment for entrustment decisions: Guidelines for programs and educators: AMEE Guide No. 154. *Medical Teacher*, *45*(11), 1203-1213. https://doi.org/10.1080/0142159X.2023.2168527

[19] Jenkins, L. S. (2025). Portfolio of learning in clinical training. *South African Family Practice*, *67*(1), 6080. https://doi.org/10.4102/safp.v67i1.6080

[20] Kurowski-Ford, A., Sharman, P., Savill, J., Shamim, I., Siddhi, P., Savill, J. N., & Siddhi, P. S. (2026). What Makes Medical Education Meaningful: A Qualitative Evaluation of Student and Educator Perspectives on UK Clinical Education. *Cureus*, *18*(4). https://doi.org/10.7759/cureus.107280

[21] Alnaser-Almusa, O., Sajid, M. R., El-Sibai, A., Ammar, O., Aldahoul, S., El Hayek, T., & Tamim, H. (2026). Exploring the effectiveness and awareness of Anki flashcard application on medical student's academic performance: a cross-sectional study from Saudi Arabia. *Frontiers in Medicine*, *13*, 1896043. https://doi.org/10.3389/fmed.2026.1896043

[22] Areli, E., Godfrey, H., Perry, M., Hempel, D., Saipe, B., Grainger, R., Hale, L., & Devan, H. (2021). 'I think there is nothing... that is really comprehensive': healthcare professionals' views on recommending online resources for pain self-management. *British journal of pain*, *15*(4), 429-440. https://doi.org/10.1177/2049463720978264

[23] Gutierrez, G., Lunsky, I. O., Van Heer, S., Szulewski, A., & Chaplin, T. (2023). Cognitive load theory in action: e-learning modules improve performance in simulation-based education. A pilot study. *Canadian journal of emergency medicine*, *25*(11), 893-901. https://doi.org/10.1007/s43678-

023-00586-z

[24] Albers, F., Trypke, M., Stebner, F., Wirth, J., & Plass, J. L. (2023). Different types of redundancy and their effect on learning and cognitive load. *British Journal of Educational Psychology*, *93*, 339-352. https://doi.org/10.1111/bjep.12592

[25] Christensen, A. R., Spagnoletti, C. L., & Claxton, R. N. (2021). A curriculum innovation on writing simulated patient cases for communication skills education. *MedEdPORTAL*, *17*, 11068. https://doi.org/10.15766/mep_2374-8265.11068

[26] Mak, V., Malone, D., Karunaratne, N., Yao, W., Randell, L., & Vu, T. (2022). A video-based reflective design to prepare first year pharmacy students for their first objective structured clinical examination (OSCE). In: *Healthcare: 2022*: MDPI, 280. https://doi.org/10.3390/healthcare10020280

[27] Weston, M. J., Meek, P., Verran, J. A., Brewer, B. B., Clavelle, J. T., Porter-O'Grady, T., & Wong, B. (2022). The Verran professional governance scale: confirmatory and exploratory analysis. *JONA: The Journal of Nursing Administration*, *52*(5), 273-279. https://doi.org/10.1097/NNA.0000000000001147

[28] Joselyn, A. S., Suhag, K., Joy, M., & Jeyaseelan, L. (2022). Development and validation of Daycare Anesthesia Satisfaction (DAS) questionnaire to assess patient's satisfaction with daycare anesthesia. *Journal of anaesthesiology, clinical pharmacology*, *38*(3), 474. https://doi.org/10.4103/joacp.JOACP_563_20

[29] Guillon, N., Jansen, L., & Alam, K. (2025). A proposed feedback model for the objective structured clinical examination (OSCE) in health and medical education. *Medical Teacher*, *47*(2), 205-207. https://doi.org/10.1080/0142159X.2024.2377388

[30] Alyami, W., Alosaimi, M. H., Baeshen, A., Alghanimy, A., Salam, E., Alhujaili, S., Alqahtani, M. M., & Saad, I. E. (2026). Simulation-based and clinical training in undergraduate nuclear medicine education: evaluation of competency outcomes using a structured OSCE framework. *BMC Medical Education*. https://doi.org/10.1186/s12909-026-09443-w