**KOLONI**: Jurnal Multidisiplin Ilmu, 5 (4), Tahun 2026
e-ISSN: 2828-6863

# Preclinical Educational Environment at an Indonesian Medical School: A Cross-Sectional Survey Using an Indonesian Translation of the JHLES

**Daniyah Umar Anis[1], Rahadian Indarto Susilo[1,2,3*], Erica Medina Ramadhaningtyas[2], Hermanto Tri Joewono[1,4]**

[1]undergraduate Student of Medicine Education Program, Faculty of Medicine, Airlangga University, Surabaya, Indonesia.
[2]Department of Neurosurgery, Faculty of Medicine, Faculty of Medicine, Airlangga University, Surabaya, Indonesia.
[3]Faculty of Health, Medicine, and Life Sciences, Airlangga University, Banyuwangi, Indonesia.
[4]Department of Obstetrics and Gynecology, Faculty of Medicine, Airlangga University, Airlangga University Academic Hospital, Surabaya, Indonesia.
*Corresponding author: rahadian-i-s@fk.unair.ac.id

**Abstract:** The educational environment is a key determinant of medical education quality, particularly during the preclinical stage. Evidence regarding the use of the Indonesian translation of the Johns Hopkins Learning Environment Scale (JHLES) among preclinical medical students remains limited. This study aimed to assess students' perceptions of the preclinical educational environment using the Indonesian translation of the JHLES and examine its association with the Global Satisfaction Score (GSS), and to explore differences by gender and cohort. A cross-sectional study was conducted among 210 preclinical medical students from the 2023 and 2024 cohorts at Universitas Airlangga. Data were collected using the 28-item Indonesian translation of the JHLES and analyzed using descriptive statistics, Spearman's rho correlation, and Mann–Whitney U test. The mean total JHLES score was 80.70 ± 11.48 (72.05% of the maximum possible score), above the scale midpoint. Community of Peers had the highest mean domain score (78.61%), whereas Inclusion and Safety had the lowest (55.59%). Total JHLES scores were significantly correlated with the GSS ($\rho = 0.379$, $p < 0.001$). No significant gender differences were observed in total JHLES score, domain score, or GSS. Students in the 2023 cohort reported significantly higher Mentoring scores than those in the 2024 cohort (median: 2023 = 6.00, 2024 = 5.00; $U = 3133.5$, $p < 0.001$). Preclinical medical students' perceptions of the educational environment varied across JHLES domains, with reliability also varying across subscales. Subgroup comparisons were exploratory. These findings are preliminary pending further psychometric validation.

**Keywords:** educational environment, medical students, preclinical education, JHLES, student perception.

***Abstrak:*** *Lingkungan pendidikan merupakan penentu utama kualitas pendidikan kedokteran, khususnya pada tahap praklinis. Bukti mengenai penggunaan terjemahan bahasa Indonesia dari Johns Hopkins Learning Environment Scale (JHLES) di kalangan mahasiswa kedokteran praklinis masih terbatas. Studi ini bertujuan untuk menilai persepsi mahasiswa terhadap lingkungan pendidikan praklinis menggunakan terjemahan bahasa Indonesia dari JHLES dan meneliti hubungannya dengan Global Satisfaction Score (GSS), serta untuk mengeksplorasi perbedaan berdasarkan jenis kelamin dan kohort. Studi potong lintang dilakukan pada 210 mahasiswa kedokteran praklinis dari kohort 2023 dan 2024 di Universitas Airlangga. Data dikumpulkan menggunakan terjemahan bahasa Indonesia dari JHLES yang terdiri dari 28 item dan dianalisis menggunakan statistik deskriptif, korelasi Spearman's rho, dan uji Mann–Whitney U. Skor total JHLES rata-rata adalah 80,70 ± 11,48 (72,05% dari skor maksimum yang mungkin), di atas titik tengah skala. Komunitas Rekan Sejawat memiliki skor domain rata-rata tertinggi (78,61%), sedangkan Inklusi dan Keamanan memiliki skor terendah (55,59%). Skor total JHLES berkorelasi signifikan dengan GSS ($\rho = 0{,}379$, $p < 0{,}001$). Tidak ditemukan perbedaan gender yang signifikan pada skor total JHLES, skor domain, atau GSS. Mahasiswa angkatan 2023 melaporkan skor Mentoring yang secara signifikan lebih tinggi daripada mahasiswa angkatan 2024 (median: 2023 = 6,00, 2024 = 5,00; $U = 3133{,}5$, $p < 0{,}001$). Persepsi mahasiswa kedokteran praklinis tentang lingkungan pendidikan bervariasi di seluruh domain JHLES, dengan reliabilitas juga bervariasi di seluruh subskala. Perbandingan subkelompok bersifat eksploratif. Temuan ini bersifat pendahuluan dan masih menunggu validasi psikometrik lebih lanjut.*

***Kata kunci:*** *lingkungan pembelajaran, mahasiswa kedokteran, pendidikan pra-klinik, JHLES, persepsi mahasiswa.*

## Introduction

The educational environment is one of the key determinants of quality in medical

education. It encompasses not only the physical learning setting but also interactions between students and faculty members, peer relationships, academic culture, mentoring systems, and the psychological climate that collectively shape students' learning experiences. Previous studies have demonstrated that a positive educational environment is associated with greater learning motivation, student engagement, psychological well-being, professional identity formation, and improved academic performance (Kassab, Rathan, Taylor, & Hamdy, 2024; Orsini, Binnie, & Wilson, 2016). Conversely, an unsupportive educational environment may reduce learning satisfaction, increase academic stress, and contribute to burnout among medical students (L. Gruppen, Irby, Durning, & Maggio, 2018). Therefore, evaluation of the educational environment has become an important indicator of quality assurance in medical education.

The educational environment is not static but continually evolves alongside changes in curricula, teaching methods, student characteristics, and educational systems. The implementation of competency-based curricula, student-centered learning, digital technologies, and increasing attention to student well-being has highlighted the need for regular evaluation of the educational environment. Findings from such evaluations may inform academic policy development, faculty development programs, improvements in academic mentoring systems, and continuous enhancement of students' learning experiences (Genn, 2001; L. Gruppen et al., 2018).

Several instruments have been developed to evaluate the educational environment in medical education, including the Dundee Ready Educational Environment Measure (DREEM) (S. U. E. Roff et al., 1997), the Postgraduate Hospital Educational Environment Measure (PHEEM) (S. Roff, McAleer, & Skinner, 2005), and the Johns Hopkins Learning Environment Scale (JHLES) (Shochet, Colbert-Getz, & Wright, 2015). Compared to other instruments, the JHLES was developed based on a contemporary conceptual framework of the educational environment that extends beyond academic aspects to encompass interpersonal relationships, student engagement, mentoring, inclusion and safety, and the quality of the physical learning environment. The instrument consists of 28 items grouped into seven domains: Community of Peers, Faculty Relationships, Academic Climate, Meaningful Engagement, Mentoring, Inclusion and Safety, and Physical Space (Shochet et al., 2015). Previous studies have demonstrated that the JHLES possesses good construct validity and provides a more comprehensive assessment of students' perceptions of the educational environment (Adamgbe et al., 2024; Shochet et al., 2015).

Previous studies conducted in different countries have shown that students' perceptions of the educational environment vary according to institutional characteristics, educational culture, and learning systems (Hee et al., 2019). Almansour et al. reported that medical students in Saudi Arabia rated peer relationships highly but identified weaknesses in inclusion and safety

(Almansour et al., 2024). Similarly, Liu et al. found that medical students with higher academic achievement in China tended to perceive the educational environment more positively than those with lower academic achievement (Liu et al., 2025). These findings suggest that evaluating the educational environment not only reflects students' satisfaction but also identifies areas requiring improvement, thereby providing valuable evidence for curriculum development and quality improvement initiatives. Taken together, these studies indicate that perceptions of the educational environment may differ across countries because of variations in institutional characteristics, academic culture, and educational systems. Therefore, findings from one country cannot be directly generalized to other educational settings, highlighting the importance of local evaluation (L. D. Gruppen & Stansfield, 2016).

In Indonesia, studies evaluating the educational environment in medical education have predominantly employed the DREEM instrument (Soemantri, 2017; Tejoyuwono, Armyanti, & Nugraha, 2015), whereas the application of the JHLES remains limited. Furthermore, few studies have examined Indonesian medical students' perceptions of the educational environment using the JHLES, particularly during the preclinical phase (Indriana & Novianto, 2020). As medical schools differ in curriculum design, academic culture, and educational practices, findings from other countries cannot be directly generalized to the Indonesian context. Information regarding students' perceptions of the educational environment at the Faculty of Medicine, Universitas Airlangga also remains limited, highlighting the need for a systematic evaluation to identify institutional strengths and areas requiring improvement.

Therefore, this study aimed to assess medical students' perceptions of the preclinical educational environment at Universitas Airlangga using the Indonesian translation of the Johns Hopkins Learning Environment Scale (JHLES). In addition to describing students' perceptions across the seven JHLES domains, this study examined the association between JHLES scores and students' overall satisfaction with the educational environment, as measured by the Global Satisfaction Score (GSS), and compared perceptions according to respondents' demographic characteristics. The findings are expected to assist the faculty in identifying institutional strengths and priority areas for improvement while supporting evidence-based strategies for the continuous enhancement of the educational environment.

## Methods

This study employed a quantitative cross-sectional design to assess medical students' perceptions of the preclinical educational environment at the Faculty of Medicine, Universitas Airlangga, using the Indonesian translation of the Johns Hopkins Learning Environment Scale (JHLES). The study was conducted between September 2025 and February 2026. The study population consisted of preclinical medical students from the 2023 and 2024 cohorts who had

completed at least one semester of study. Convenience sampling was used, whereby questionnaires were distributed to eligible students who were present during scheduled data collection sessions. Of the 722 eligible students, 210 participated in the study, comprising 61 students from the 2023 cohort and 149 from the 2024 cohort, representing an overall population coverage of 29.1%. All 210 students who were present and invited agreed to participate and completed the questionnaire. The inclusion criteria were active undergraduate medical students in the preclinical phase from the 2023 or 2024 cohorts who provided written informed consent. Incomplete questionnaires and participants who withdrew consent before data analysis were defined as exclusion criteria. Since convenience sampling was used and only a proportion of the eligible population was reached, selection bias could not be excluded, and the findings should therefore be interpreted within the context of the participating sample.

The study employed the Indonesian translation of the Johns Hopkins Learning Environment Scale (JHLES), consisting of 28 items across seven domains: Community of Peers, Faculty Relationships, Academic Climate, Meaningful Engagement, Mentoring, Inclusion and Safety, and Physical Space. The Indonesian translation was developed through a preliminary forward–backward translation procedure. Two certified translators independently translated the original instrument into Indonesian, followed by reconciliation into a single preliminary version. Subsequently, two independent translators who were blinded to the original instrument performed back-translation into English. The back-translated version was sent to the original instrument developer, Dr. Robert B. Shochet, to obtain feedback regarding semantic equivalence; however, no comments or suggested revisions were received. The translated instrument was therefore used as a preliminary Indonesian version, without interpreting the absence of feedback as formal authorization for use. Items 1–23 and 27–28 were scored from 0 (strongly disagree) to 4 (strongly agree), while negatively worded items (Items 24–26) were reverse-scored using the formula 4 minus the original score. In addition to the JHLES, participants completed the Global Satisfaction Score (GSS), a single-item measure assessing their overall perception of the learning environment, with response options ranging from Very Bad (0) to Very Good (4). Higher scores indicated more favorable perceptions of the educational environment.

Data were collected during scheduled classroom sessions when eligible students were present. Before completing the questionnaire, participants received an information sheet and a paper-based informed consent form explaining the study objectives, confidentiality, voluntary participation, and their right to withdraw without academic consequences. The questionnaire was administered online through Google Forms and accessed via a QR code displayed on the classroom monitor. The online questionnaire was configured to allow each respondent to submit only one response to minimize duplicate submissions. The number of completed paper-based informed consent forms was checked against the number of submitted online responses before

participants left the classroom to ensure consistency and monitor potential duplicate or unmatched responses. No personally identifying information was collected, and all responses were recorded without direct identifiers. Before statistical analysis, data completeness and response consistency were examined. The three negatively worded items (Items 24, 25, and 26) were reverse-scored, and missing data were assessed. No questionnaires met the predefined exclusion criteria, and all 210 responses were included in the final analysis.

Statistical analyses were performed using IBM SPSS Statistics version 31. Participant characteristics were summarized using frequencies, percentages, means, standard deviations, medians, and ranges, as appropriate. Total JHLES scores and domain scores were calculated by summing the corresponding item scores, with the total score ranging from 0 to 112. The percentage of the maximum possible score was calculated by dividing the observed score by the maximum possible score and multiplying by 100. Internal consistency was assessed using Cronbach's alpha, with values of 0.70 or higher considered acceptable. Corrected item–total correlations, Cronbach's alpha if item deleted, and inter-item correlations for two-item domains were also examined. Descriptive statistics were used to summarize total and domain scores, while item-level mean percentages and 95% confidence intervals were calculated to facilitate comparisons. Associations between JHLES scores and GSS were examined using Spearman's rank correlation coefficient. Differences according to gender and cohort were analyzed using the Mann–Whitney U test, with median scores and interquartile ranges reported. Effect sizes for statistically significant comparisons were calculated using rank-biserial correlation. Item-level comparisons were considered exploratory and were not adjusted for multiple testing. A p-value of less than 0.05 was considered statistically significant. This study received ethical approval from the Health Research Ethics Committee, Faculty of Medicine, Universitas Airlangga (Approval No. 238/EC/KEPK/FKUA/2025). Written informed consent was obtained from all participants, and the study was conducted in accordance with the principles of the Declaration of Helsinki.

## Results & Discussion

### Participant characteristics

A total of 210 preclinical medical students from the Undergraduate Medical Education Program at Universitas Airlangga participated in the study and completed all questionnaires; therefore, all responses were included in the analysis. Participant characteristics are presented in Table 1. Most respondents were female (74.8%) and belonged to the 2024 cohort (71.0%), while nearly all participants were of Asian ethnicity (99.5%). Most students reported a good Global Satisfaction Score (68.6%), followed by very good (17.1%) and fair (14.3%).

**Tabel 1 Demographic characteristics of the preclinical medical students**

| Variable | Category | Frequency (n) | Percentage (%) |
|---|---|---|---|

| | | | |
|---|---|---|---|
| **Gender** | Male | 53 | 25.2 |
| | Female | 157 | 74.8 |
| **Cohort** | 2023 | 61 | 29.0 |
| | 2024 | 149 | 71.0 |
| **Ethnicity** | Asian | 209 | 99.5 |
| | Caucasian | 1 | 0.5 |
| | Other | 0 | 0 |
| **Global Satisfaction** | Very good | 36 | 17.1 |
| | Good | 144 | 68.6 |
| | Fair | 30 | 14.3 |
| | Poor | 0 | 0 |
| | Very poor | 0 | 0 |

Internal consistency of the Indonesian translation of the JHLES

Cronbach's alpha analysis was performed to evaluate the internal consistency of the Indonesian translation of the Johns Hopkins Learning Environment Scale (JHLES). The 28-item JHLES yielded a Cronbach's alpha of 0.893 in this sample; given its multidomain structure, this coefficient was not interpreted as evidence of unidimensionality.

The Community of Peers domain also demonstrated good internal consistency (Cronbach's α = 0.853; 6 items). Examination of the Cronbach's alpha if item deleted statistics showed that removal of any individual item did not meaningfully improve the alpha coefficient; therefore, all six items were retained.

The Faculty Relationships domain demonstrated good internal consistency (Cronbach's α = 0.815; 6 items), and deletion of any individual item did not substantially improve the reliability of the domain. Therefore, all six items were retained.

The Academic Climate domain demonstrated good internal consistency (Cronbach's α = 0.816; 5 items). Examination of the Cronbach's alpha if item deleted statistics indicated that removal of any individual item would not meaningfully improve the alpha coefficient, and all five items were retained.

The Meaningful Engagement domain yielded a Cronbach's alpha of 0.649 (4 items), indicating lower internal consistency than the conventional threshold of 0.70. Checking the corrected item–total correlations alongside the alpha-if-item-deleted statistics showed values of 0.602, 0.488, 0.155, and 0.621 for Items 18, 19, 20, and 21, respectively, confirming that Item 20 demonstrated the lowest correlation within the domain. Removing Item 20 would increase Cronbach's alpha to 0.814. This substantial change indicates that Item 20 behaves differently from the other items in this domain, and interpretation of the Meaningful Engagement domain as a unified construct should therefore be made with caution pending further psychometric evaluation. However, because the present study did not

include cognitive interviewing or construct validity analyses, the reasons underlying this item's weaker performance remain speculative and should be clarified through cognitive testing in future research. All items in this domain were retained to preserve the original structure of the JHLES, while further studies are needed to evaluate its semantic and construct equivalence in the Indonesian context.(Adamgbe et al., 2024)

Among the seven domains, Mentoring had the highest Cronbach's alpha (α = 0.903; 2 items). For a two-item domain, Cronbach's alpha is mathematically equivalent to the Spearman-Brown coefficient applied to the inter-item correlation, corresponding to an inter-item correlation of r = 0.826 between Items 22 and 23. This strong inter-item correlation indicates high consistency between the two items in this sample; however, because the domain contains only two items, this coefficient reflects the correlation between a single item pair rather than the reliability of a broader subscale, and should be interpreted cautiously.

The Inclusion and Safety domain demonstrated good internal consistency (Cronbach's α = 0.830; 3 items). Examination of the Cronbach's alpha if item deleted statistics showed that deleting item 24 resulted in only a minimal increase in the alpha coefficient (0.833), whereas deletion of items 25 or 26 reduced the alpha coefficient to 0.755 and 0.698, respectively. Because removal of item 24 provided only a minimal improvement and all three items represent important aspects of inclusion and safety, all items were retained.

The Physical Space domain showed the lowest internal consistency (Cronbach's α = 0.561; 2 items), corresponding to an inter-item correlation of r = 0.394 between Items 27 and 28. This weak inter-item correlation may not be attributable solely to the small number of items in this domain. Item 27 concerns the preclinical building, whereas Item 28 concerns spaces where clinical teaching occurs; because participants in this study were preclinical students, these two items may not represent environments that were equally or directly experienced, which could account for their limited correlation. Given that this domain consisted of only two items, the alpha coefficient should also be interpreted cautiously because Cronbach's alpha is sensitive to the number of items in a scale. Both items were retained to preserve the original JHLES structure.

Overall, the domain-level analyses showed variation in internal consistency across the seven JHLES domains. The differences in domain-level coefficients may reflect variation in the number of items and the content represented by each domain. These findings provide preliminary evidence of internal consistency in the present sample, but do not establish the construct validity or broader applicability of the instrument. Further studies incorporating construct-validity assessment, cognitive testing, and factor-structure analyses are needed to evaluate the measurement properties of the Indonesian JHLES.

**Overall educational environment score**

The mean total JHLES score was 80.70 ± 11.48, with individual scores ranging from 35 to 104 out of a maximum possible score of 112, corresponding to 72.05% of the maximum attainable score.

Figure 1 illustrates the distribution of the total JHLES score and the seven domain scores. Overall, the domain scores showed similar distributions, with some variation across domains. Community of Peers had the highest mean percentage score (78.61%), followed by Meaningful Engagement (74.97%), Academic Climate (73.45%), Faculty Relationships (73.27%), and Physical Space (71.55%). Mentoring (64.52%) and Inclusion and Safety (55.59%) had the lowest scores. The overall JHLES mean score was 72.05%, indicating generally positive perceptions of the educational environment.

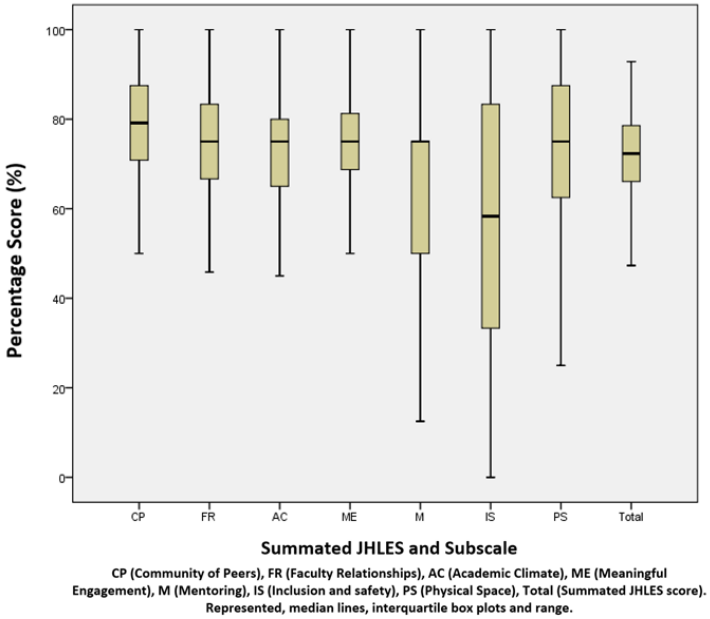



**Figure 1 Boxplot of Percentage Scores for the Total JHLES and Its Seven Domains**

Community of Peers and Meaningful Engagement achieved the highest domain scores, whereas Mentoring and Inclusion and Safety received comparatively lower scores. These findings suggest that students perceived peer support and engagement in learning more positively than mentoring and aspects related to inclusion and safety.

The present study demonstrated that preclinical medical students at Universitas Airlangga generally perceived their educational environment positively, as reflected by a mean JHLES score of 72.05% of the maximum attainable score and a high level of overall satisfaction. These findings indicate that the educational environment generally provides adequate support for students' learning experiences during the preclinical phase.

Our findings are consistent with previous studies using the JHLES and other educational environment instruments, which have generally reported positive perceptions among medical students when institutions foster supportive student–faculty relationships and cultivate a positive academic culture.(Adamgbe et al., 2024; Shochet et al., 2015) These findings reinforce the value of educational environment assessment as an important indicator of educational quality and a useful tool for guiding continuous quality improvement in medical education.

**Domain-specific analysis**

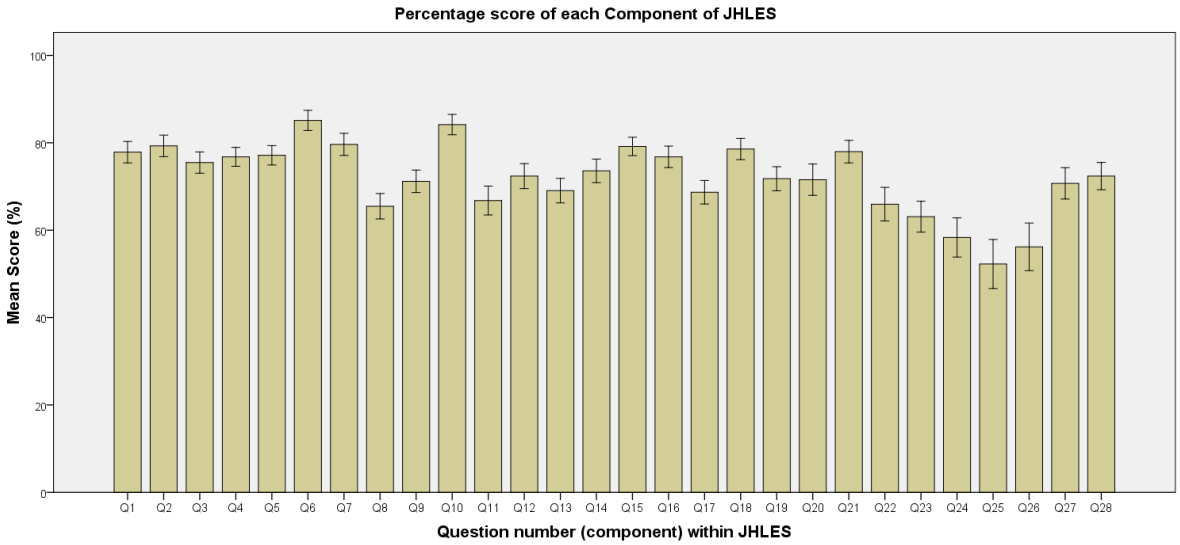



**Figure 2 Mean Percentage Scores for Each JHLES Items. (Q (Reference to each question number within JHLES). Represented averages and associated error bars set at 95% CI)**

The mean percentage scores and corresponding 95% confidence intervals (95% CIs) for all JHLES items are presented in Figure 2. Most items achieved mean percentage scores above 70%, indicating generally positive perceptions across most aspects of the preclinical educational environment. The highest mean percentage scores were observed for Item 6, which assesses the presence of positive and inspiring role models among fellow students, and Item 10, which assesses the presence of positive faculty role models. Other highly rated items included Items 1-5, reflecting positive perceptions of peer relationships and a sense of community among fellow students, and Items 14-19, which relate to academic climate and meaningful engagement during the preclinical learning process.

In contrast, lower mean percentage scores were observed for Items 8 and 11, within the Faculty Relationships domain, which assessed the extent to which faculty members take time to get to know students and are approachable for confiding important concerns. A relatively lower mean percentage score was also observed for Item 13, within the Academic Climate domain, which assesses the extent to which the curriculum accommodates students' preferred learning styles. The lowest scores were found for Items 24–26, which assessed concerns about mistreatment, discrimination, and personal safety. Although these negatively worded items were reverse-scored before analysis, their comparatively lower scores indicate that students perceived these aspects less positively than other components of the educational environment. Items 22 and 23, which assess research mentoring and mentoring in a clinical specialty or discipline of interest, also received relatively modest scores, consistent with the lower overall score observed for the Mentoring domain. Overall, these findings suggest that while students perceived strong peer relationships, supportive faculty, and meaningful engagement in learning, opportunities to strengthen mentoring, inclusion, and personal safety remain important priorities for further evaluation and improvement.

The high score observed in the Community of Peers domain was reflected by consistently positive responses across all six items (Items 1–6), which assess peer support, collaboration, belonging, and the quality of interactions among students The median score for the Community of Peers domain was 19 in both male and female students, with no significant difference between genders ($p = 0.836$). Similarly, the median score was 19 among students from the 2023 cohort and 18 among those from the 2024 cohort, with no significant difference between cohorts ($p = 0.187$). These findings suggest that students generally perceived a collaborative learning environment in which peers provided both academic and social support, with similar perceptions across gender and cohort. Such peer interactions are fundamental in competency-based medical education, where learning commonly occurs through collaborative activities such as problem-based learning (PBL), team-based learning (TBL), and small-group discussions. These learning approaches encourage students to exchange knowledge, provide constructive feedback, and develop communication and teamwork skills that are essential for future clinical practice.(Adamgbe et al., 2024)

In contrast, lower scores were observed for several items within the Faculty Relationships domain (Items 8 and 11), indicating that some aspects of students' interactions with faculty members were perceived less positively than peer relationships and academic engagement. The lowest individual item scores were observed for Items 24–26, which belong to the Inclusion and Safety domain. These findings indicate that students perceived inclusion, respectful treatment, and personal safety less positively than other components of the educational environment, despite these items being reverse-scored before analysis. In addition, Items 22 and 23, which form the Mentoring domain, also demonstrated relatively lower scores than most other items, consistent with the lower overall Mentoring domain score. These findings suggest that while peer relationships, academic engagement, and the overall learning climate were perceived positively, mentoring, inclusion, and personal safety represent areas that may benefit from further attention.

The Inclusion and Safety domain contained the lowest-scoring items in the present study (Items 24–26), indicating that students perceived aspects related to respectful treatment, discrimination, and personal safety less positively than other components of the educational environment. Although these negatively worded items were reverse-scored prior to analysis, their comparatively lower scores suggest that inclusion and safety remain areas requiring attention. Because this study did not investigate the factors underlying these perceptions, no conclusions can be drawn regarding their causes. Future qualitative and longitudinal studies are warranted to explore students' experiences related to inclusion, respectful treatment, and personal safety and to identify strategies for fostering a more inclusive and supportive educational environment.

**Association between JHLES and Global Satisfaction Score**

The mean Global Satisfaction Score (GSS) was $3.03 \pm 0.56$. As illustrated in Figure 3, higher overall JHLES scores were associated with higher GSS. Spearman's rank correlation analysis ($n = 210$) demonstrated a statistically significant weak positive correlation between the total JHLES score and

the GSS (ρ = 0.379, p < 0.001). According to Schober et al., correlation coefficients of 0.10–0.39 indicate weak correlations, whereas coefficients of 0.40–0.69 indicate moderate correlations.(Schober et al., 2018) Among the seven domains, Faculty Relationships (ρ = 0.433, p < 0.001) and Academic Climate (ρ = 0.428, p < 0.001) demonstrated moderate positive correlations with the GSS. Community of Peers (ρ = 0.344, p < 0.001), Meaningful Engagement (ρ = 0.334, p < 0.001), Mentoring (ρ = 0.224, p = 0.001), and Physical Space (ρ = 0.194, p = 0.005) demonstrated weak positive correlations. In contrast, the Inclusion and Safety domain demonstrated a weak negative correlation with the GSS (ρ = −0.293, p < 0.001). As described in the Methods, Items 24–26 were reverse-scored using the formula 4-score (0–4 scale) prior to computing the Inclusion and Safety domain score, so that higher values consistently represented more favorable perceptions across all items, including this domain. However, because item-level responses were reverse-coded at the point of data entry and the pre-reversal raw scores were not retained in a separately auditable form, a retrospective audit comparing raw and reverse-scored values could not be performed. Consequently, the possibility of an undetected coding or direction error affecting this specific domain cannot be fully excluded on the basis of the available data. The wording of Items 24–26 was confirmed against the original questionnaire (Additional file 1) to be consistently negative in framing, which supports but does not independently verify the correctness of the reverse-scoring procedure applied to this domain. Furthermore, the Inclusion and Safety domain demonstrated negative correlations with all other JHLES domains (ρ ranging from −0.19 to −0.33, all p < 0.01); while this systematic pattern is consistent with either a genuine, domain-wide association or a systematic scoring issue specific to this domain, it does not on its own resolve the question of data provenance. Given this limitation, the negative correlation between Inclusion and Safety and the GSS is reported here descriptively and should not be treated as a validated substantive finding. The restricted distribution of GSS responses, in which no students reported "Poor" or "Very poor" satisfaction, may have further limited the stability of correlation estimates across domains, including this negative association. This pattern should be regarded as provisional pending independent verification using auditable source-level data in future studies, and should not be used as a basis for institutional intervention or firm scientific conclusions until it can be corroborated.

These findings suggest that Faculty Relationships showed the strongest observed association with GSS among the domains. Supportive interactions with faculty, including constructive feedback, approachability, and genuine concern for students' well-being, were most strongly associated with higher GSS. A positive academic climate also demonstrated a moderate positive correlation with the GSS, consistent with a learning environment that promotes fairness, clear educational goals, and meaningful learning experiences. Although Community of Peers, Meaningful Engagement, Mentoring, and Physical Space were also positively associated with the GSS, these relationships were comparatively weaker.

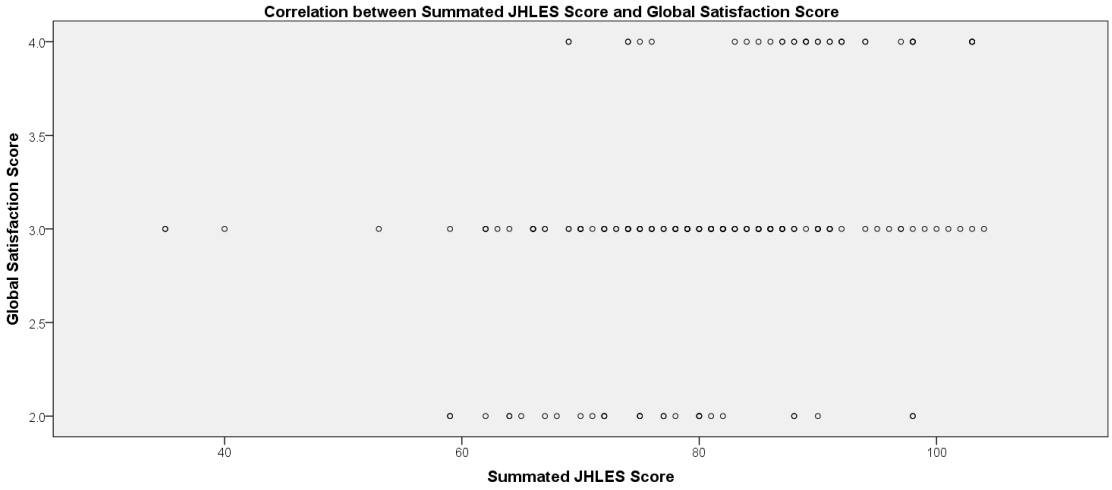



**Figure 3 Correlation between summated JHLES score and Global Satisfaction Score (ρ = 0.379, p < 0.001), indicating a weak positive correlation**

**Subgroup analysis**

The Mann–Whitney U test was used for subgroup analysis, and detailed results, including IQR and rank-biserial correlation effect size (r), are presented in Additional file 2. There were no significant differences between male and female students in the total JHLES score, domain scores, or the Global Satisfaction Score (GSS) (all p>0.05; r=0.02–0.12) with effect sizes indicating negligible-to-small differences. At the item level, several nominal differences were observed between male and female students (Items 10, 13, and 15); however, these analyses were exploratory and were not adjusted for multiple comparisons. The observed effect sizes were small (r=0.17–0.20), and these findings should therefore be interpreted cautiously.sa

At the item level, nominal, unadjusted differences (not corrected for multiple comparisons) were observed between male and female students for Items 10, 13, and 15, all favoring female students. Item 10 (assessing perceptions of positive and inspiring faculty role models) showed the largest of these differences (U = 3322.0, p = 0.015, r = 0.20; median 3.00 [IQR 3.00–4.00] for male vs. 4.00 [IQR 3.00–4.00] for female students). Items 13 and 15 showed smaller nominal differences (U = 3459.5, p = 0.046, r = 0.17; and U = 3459.0, p = 0.032, r = 0.17, respectively); for both items the median scores were identical between groups, but mean ranks were consistently higher among female students (Item 13: 109.96 vs. 92.27; Item 15: 109.97 vs. 92.26), suggesting a modestly more favorable response distribution. Given the number of item-level comparisons conducted without correction for multiplicity, these differences are unlikely to remain significant under a conservative multiplicity correction and should be interpreted as exploratory rather than confirmatory.

Given the large number of subgroup comparisons conducted across the total score, domain scores, and individual items without correction for multiple testing, these item-level findings should be interpreted with caution. Furthermore, the observed effect sizes were generally small (r≈0.17–0.20).

Comparison between cohorts showed no significant differences in the total JHLES score, most domain scores, or the Global Satisfaction Score (GSS). A statistically significant difference was

observed only in the Mentoring domain. The median score was 6.00 (IQR 5.00–8.00) for the 2023 cohort and 5.00 (IQR 4.00–6.00) for the 2024 cohort ($U = 3133.5$, $p<0.001$, $r = 0.31$). Students from the 2023 cohort therefore reported higher mentoring scores.

At the item level, Item 6 showed a nominal, unadjusted difference between cohorts ($U = 3765.5$, $p = 0.029$, $r = 0.17$). The median score was 4.00 (IQR 3.00–4.00) in the 2023 cohort and 3.00 (IQR 3.00–4.00) in the 2024 cohort. This item assesses the presence of positive and inspiring role models among fellow students, with responses favoring the 2023 cohort. Item 22 also showed a nominal, unadjusted difference between cohorts ($U = 2786.0$, $p<0.001$, $r = 0.39$). The median score was 3.00 (IQR 3.00–4.00) in the 2023 cohort and 3.00 (IQR 2.00–3.00) in the 2024 cohort. This item assesses mentoring in a research field of interest, with responses favoring the 2023 cohort. Item 23 showed a nominal, unadjusted difference as well ($U = 3637.5$, $p = 0.018$, $r = 0.20$). The median score was 3.00 (IQR 2.00–4.00) in the 2023 cohort and 2.00 (IQR 2.00–3.00) in the 2024 cohort. This item assesses mentoring in a clinical specialty or discipline of interest, with responses favoring the 2023 cohort.

The effect sizes for the cohort-level differences ranged from small for Items 6 and 23 to moderate for Item 22 and the Mentoring domain. These findings suggest that the mentoring-related disparity between cohorts may be more substantial than the gender-based differences. The observed cohort difference was primarily reflected in mentoring-related items. However, the substantially lower response coverage among the 2023 cohort (16.7%) compared with the 2024 cohort (41.7%), together with the attendance-based sampling approach, limits the interpretation of these findings. Therefore, the cohort comparisons should be considered applicable only to the students who participated in this study and should not be interpreted as representative differences between the 2023 and 2024 cohort populations as a whole.

Because this study employed a cross-sectional design, the observed difference should be interpreted as an association between cohort and mentoring perceptions rather than evidence of changes over time within individual students. The difference may reflect cohort characteristics, differences in exposure to the educational environment, or other unmeasured factors.(Levin, 2006; Setia, 2016) Therefore, longitudinal studies following the same students throughout the preclinical phase are needed to determine whether perceptions of mentoring change over time and to identify factors contributing to these differences.

These findings indicate that while students' overall perceptions of the educational environment remained consistent across demographic groups, students' perceptions of mentoring differed by cohort, with students from the 2023 cohort reporting higher mentoring scores than those from the 2024 cohort. Because this study employed a cross-sectional design, this difference should not be interpreted as evidence of improvement over time or increasing mentoring quality with duration of study. Further longitudinal research is needed to determine whether mentoring perceptions change throughout the preclinical period.

**Implications, strengths, and limitations**

This study yielded four principal findings: the total JHLES score was above the scale midpoint; Community of Peers emerged as the highest-scoring domain, whereas Inclusion and Safety and Mentoring received the lowest relative domain scores; the total JHLES score was significantly associated with the GSS; and students in the 2023 cohort reported higher Mentoring scores than those in the 2024 cohort within the participating sample.

A major strength of this study is its ability to identify both strengths and areas requiring improvement within the educational environment. These findings may support faculty quality assurance initiatives and provide baseline data for continuous monitoring of educational quality. Periodic use of the Indonesian translation of the JHLES may facilitate evaluation of curriculum changes, educational innovations, and faculty development programs aimed at enhancing students' learning experiences.

This study has several limitations. First, its cross-sectional design precludes assessment of changes in students' perceptions over time. Second, the study was conducted at a single institution, which may limit the generalizability of the findings to other medical schools. Third, although the Indonesian translation of the JHLES demonstrated acceptable internal consistency in this sample, reliability estimates varied across domains. The Mentoring and Physical Space domains each comprised only two items, and their Cronbach's alpha coefficients ($\alpha = 0.903$ and $\alpha = 0.561$, respectively) should therefore be interpreted cautiously because Cronbach's alpha is sensitive to the number of items in a scale. In addition, Item 28 refers to clinical-teaching workspaces, whereas the participants in this study were preclinical students. The extent to which participants had sufficient direct exposure to the setting described in this item was not specifically assessed. This contextual difference may have influenced responses to the Physical Space domain and should be considered when interpreting its low internal consistency (Cronbach's $\alpha = 0.561$). This study did not evaluate other psychometric properties, such as content validity, construct validity, expert committee review, cognitive pretesting, cognitive interviewing, or conceptual equivalence. Therefore, the findings should be interpreted as preliminary evidence of reliability, and the conceptual equivalence of the translated items to the original instrument could not be established in the present study. Further validation studies are required before broader application of the instrument. Fourth, because item responses for Items 24–26 were reverse-coded at the point of data entry rather than stored separately in their original form, a full raw-to-reversed scoring audit could not be retrospectively performed. As a result, the negative correlation observed between the Inclusion and Safety domain and the GSS cannot be independently confirmed as a genuine substantive finding, and the possibility of an undetected scoring or coding artifact affecting this specific domain cannot be excluded. This represents a data provenance limitation rather than a resolved analytical finding. Future studies should retain both raw and reverse-coded item responses to allow independent verification of scoring procedures. Fifth, the distribution of GSS responses was notably restricted, with no students selecting the "Poor" or "Very poor" categories and the majority reporting "Good" or "Very good" satisfaction. This restricted range

and ceiling tendency may have attenuated the variability needed to detect stronger or more stable correlations between GSS and JHLES domain scores, and should be considered when interpreting the magnitude and direction of these associations. Finally, the use of self-reported questionnaires may have introduced response bias and social desirability bias. The unequal response coverage between the 2023 and 2024 cohorts (16.7% vs. 41.7%) may have introduced selection or non-response bias and may limit the comparability of cohort-level findings. Therefore, differences between cohorts should be interpreted as applying to the students represented in the study sample rather than to the respective cohort populations. Future multicenter longitudinal studies incorporating comprehensive psychometric evaluation are warranted to provide a more robust assessment of the educational environment across medical schools in Indonesia.

## Conclusion

This study demonstrated that preclinical medical students at Universitas Airlangga reported a total JHLES score above the scale midpoint based on the Indonesian translation of the Johns Hopkins Learning Environment Scale (JHLES), with Community of Peers showing the highest relative domain score. Domain-level scores varied, with relatively higher scores observed for peer relationships, academic climate, and faculty interactions, and relatively lower scores observed for Inclusion and Safety and Mentoring.

Despite these findings, Inclusion and Safety emerged as the lowest-scoring domain. The negative correlation between this domain and the GSS could not be fully verified against pre-reversal raw data, owing to a data provenance limitation described in the Methods and Limitations; this pattern should therefore be treated as provisional rather than a confirmed substantive finding. Before this pattern is investigated further or used to inform institutional action, independent verification using auditable source-level data is needed to rule out a scoring artifact. If corroborated in future studies, further qualitative assessment of students' experiences of respectful treatment, discrimination, and personal safety may help clarify the nature and potential drivers of this finding. Institutional interventions based on this domain are not currently warranted. Mentoring also received a relatively low score and warrants further evaluation. Potential institutional responses may include strengthening mentoring structures and evaluating their effectiveness in future studies.

Higher JHLES scores were associated with higher GSS values, indicating that more favorable perceptions of specific domains of the educational environment were ssociated with more favorable overall perceptions of that environment, consistent with the global assessment construct measured by this item.

Overall, this study provides preliminary evidence that the Indonesian translation of the JHLES demonstrates acceptable internal consistency within this sample of preclinical medical

students. Although further psychometric evaluation, particularly construct validity testing using confirmatory factor analysis (CFA), is required before broader application, the instrument appears feasible for describing students' perceptions of the preclinical educational environment in this setting. The findings may serve as an initial basis for identifying aspects of the educational environment that warrant further evaluation and improvement, while additional validation studies are needed to establish the instrument's broader applicability.

## Acknowledgement

The authors gratefully acknowledge all preclinical medical students of the Faculty of Medicine, Universitas Airlangga, for their participation in this study. We also thank the Medical Doctor Study Program, Faculty of Medicine, Universitas Airlangga, for its administrative support and assistance during the data collection process.

## References

Adamgbe, M. A., Okpanachi, J. A., Inalegwu, A. J., Firima, E., Bakwa, N. D., Gani, S. A., … Anuligo, G. I. (2024). Perception of learning environment among undergraduate medical students at university of Jos using Johns Hopkins learning environment scale. *Nigerian Medical Journal*, *65*(6), 934–945.

Almansour, M., Abouammoh, N., Idris, R. Bin, Alsuliman, O. A., Alhomaidi, R. A., Alhumud, M. H., & Alghamdi, H. A. (2024). Exploring medical students' experience of the learning environment: a mixed methods study in Saudi medical college. *BMC Medical Education*, *24*(1), 723. https://doi.org/10.1186/s12909-024-05716-4

Genn, J. M. (2001). AMEE Medical Education Guide No. 23 (Part 2): Curriculum, environment, climate, quality and change in medical education – a unifying perspective. *Medical Teacher*, *23*(5), 445–454. https://doi.org/10.1080/01421590120075661

Gruppen, L. D., & Stansfield, R. B. (2016). Individual and Institutional Components of the Medical School Educational Environment. *Academic Medicine*, *91*(11), S53–S57. https://doi.org/10.1097/ACM.0000000000001361

Gruppen, L., Irby, D. M., Durning, S. J., & Maggio, L. A. (2018). Interventions Designed to Improve the Learning Environment in the Health Professions: A Scoping Review. *MedEdPublish*, *7*, 211. https://doi.org/10.15694/mep.2018.0000211.1

Hee, J. M., Yap, H. W., Ong, Z. X., Quek, S. Q. M., Toh, Y. P., Mason, S., & Krishna, L. K. R. (2019). Understanding the Mentoring Environment Through Thematic Analysis of the Learning Environment in Medical Education: a Systematic Review. *Journal of General Internal Medicine*, *34*(10), 2190–2199. https://doi.org/10.1007/s11606-019-05000-y

Indriana, N., & Novianto, M. R. (2020). Hubungan Persepsi Terhadap Lingkungan Pembelajaran Dengan Hasil Belajar Mahasiswa Pendidikan Dokter Fakultas Kedokteran Ilmu Kesehatan Universitas Islam Negeri Maulana Malik Ibrahim Malang. *Journal of Islamic Medicine*, *4*(1), 21–29.

Kassab, S. E., Rathan, R., Taylor, D. C. M., & Hamdy, H. (2024). The impact of the educational environment on student engagement and academic performance in health professions education. *BMC Medical Education*, *24*(1), 1278. https://doi.org/10.1186/s12909-024-06270-9

Levin, K. A. (2006). Study design III: Cross-sectional studies. *Evidence-Based Dentistry*, *7*(1), 24–25.

Liu, Y., Lyu, D., Xie, S., Yao, Y., Liu, J., Lu, B., … Pan, X. (2025). The impact of grade point average on medical students' perception of the learning environment: a multicenter cross-sectional study across 12 Chinese medical schools. *BMC Medical Education*, *25*(1), 448. https://doi.org/10.1186/s12909-025-06977-3

Orsini, C., Binnie, V. I., & Wilson, S. L. (2016). Determinants and outcomes of motivation in health

professions education: a systematic review based on self-determination theory. *Journal of Educational Evaluation for Health Professions*, *13*, 19. https://doi.org/10.3352/jeehp.2016.13.19

Raharjanti, N. W., Wiguna, T., Purwadianto, A., Soemantri, D., Indriatmi, W., Poerwandari, E. K., … Levania, M. K. (2023). Translation, adaptation, validity and reliability of Multidimensional Emotion Questionnaire for Indonesian forensic psychiatry context. *Heliyon*, *9*(3), e13787. https://doi.org/10.1016/j.heliyon.2023.e13787

Roff, S., McAleer, S., & Skinner, A. (2005). Development and validation of an instrument to measure the postgraduate clinical learning and teaching educational environment for hospital-based junior doctors in the UK. *Medical Teacher*, *27*(4), 326–331. https://doi.org/10.1080/01421590500150874

Roff, S. U. E., McAleer, S., Harden, R. M., Al-Qahtani, M., Ahmed, A. U., Deza, H., … Primparyon, P. (1997). Development and validation of the Dundee ready education environment measure (DREEM). *Medical Teacher*, *19*(4), 295–299.

Rupani, N., Evans, A., & Iqbal, M. (2022). A quantitative cross-sectional study assessing the surgical trainee perception of the operating room educational environment. *BMC Medical Education*, *22*(1), 764. https://doi.org/10.1186/s12909-022-03825-6

Schober, P., Boer, C., & Schwarte, L. A. (2018). Correlation coefficients: appropriate use and interpretation. *Anesthesia & Analgesia*, *126*(5), 1763–1768.

Setia, M. S. (2016). Methodology series module 3: Cross-sectional studies. *Indian Journal of Dermatology*, *61*(3), 261.

Shochet, R. B., Colbert-Getz, J. M., & Wright, S. M. (2015). The Johns Hopkins Learning Environment Scale. *Academic Medicine*, *90*(6), 810–818. https://doi.org/10.1097/ACM.0000000000000706

Soemantri, D. (2017). Medical Students' Perceptions of Educational Environment: A Study on DREEM in the Faculty of Medicine Universitas Indonesia in Year 2012–2015. *Advanced Science Letters*, *23*(7), 6774–6777. https://doi.org/10.1166/asl.2017.9394

Susilo, R. I., Ramadhaningtyas, E. M., Utomo, B., Dachlan, E. G., Soemantri, D., & Groen, R. J. M. (2026). Cross-cultural adaptation and validation of the Indonesian version of the operating room educational environment measure with Indonesian residents. *BMC Medical Education*. https://doi.org/10.1186/s12909-026-09673-y

Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's alpha. *International Journal of Medical Education*, *2*, 53–55. https://doi.org/10.5116/ijme.4dfb.8dfd

Tejoyuwono, A. A. T., Armyanti, I., & Nugraha, R. P. (2015). Gambaran Evaluasi Penilaian Mahasiswa Program Studi Pendidikan Dokter terhadap Lingkungan Pembelajarannya. *Jurnal Pendidikan Kedokteran Indonesia: The Indonesian Journal of Medical Education*, *4*(3), 109. https://doi.org/10.22146/jpki.25279

Wang, X., & Cheng, Z. (2020). Cross-Sectional Studies: Strengths, Weaknesses, and Recommendations. *CHEST*, *158*(1), S65–S71. https://doi.org/10.1016/j.chest.2020.03.012