Edukasiana: Jurnal Inovasi Pendidikan
ISSN 2809-6770
Volume 3, Issue 4, 2024, pp. 405-420

A Comparative Study of Fashion Illustration Project Assessment among Teachers and Students

Emmanuel Amos, George Kwame Fobiri*, Timothy Crentsil, Diana Oppong, Ama Kour Timpabi
Department of Fashion Design and Textiles Studies, Kumasi Technical University, Kumasi, Ghana
*Corresponding Author: kfobiri@gmail.com

Article History:
Received 2024-05-04
Accepted 2024-07-28

DOI:
10.56916/ejip.v3i4.799

ABSTRACT

This paper explores the assessment of student projects in fashion design by comparing the evaluations of teachers and students (peers). Peer assessment (PA) is increasingly recognized as a valuable tool for enhancing students' assessment skills and providing alternative perspectives on project quality. However, concerns persist regarding the reliability and validity of peer assessments compared to those of teachers. To address this gap, this study aims to compare the assessments of teachers and students, focusing on variables such as Self, FIT, COLOUR, DRAWN, and BEAUTY in fashion design projects. The study used simple random sampling to sample 10 designs of Higher National Diploma (HND) two students enrolled in the fashion department of the Creative Arts Faculty. Teh descriptitve and correlation analysis methods were used to analyse the sample gathered. The Descriptives analysis revealed differences in the mean scores and variability between the Student and Teaching Staff groups for the variables Self, COLOUR, FIT, DRAWP, and BEAUTY. Again, a correlation analysis revealed strong positive correlations between various self-perception categories. This results practically implies that peer assessment is a requisite assessment mehtod in higher education, especially for practical courses such as fashion design. Collaborative assessment, especially among peers is therefore recommended to help students build confidence and ensure a sense of ownership of their projects coupled with sharpening their practical skills.

Keywords: *Peer assessment; Teacher assessment; fashion design projects; Independent sample t-test.*

1. INTRODUCTION

For a long time, educators have accepted the system of including extra hands in the assessment of students' projects. In many cases, colleague lectures, teaching support staff, technicians, and students themselves have been employed to suggest opinions on students' projects. Peer Assessment (PA) is considered to be a particularly useful tool by some educators since students need to develop their assessment skills in order to recognize quality, comprehend assessment standards, and evaluate their own performance (Iglesias et al, 2022). However, little attention has been paid to comparing the assessments of facilitators, teaching support staff, and students in an effort to optimize the acquisition of skills in project-based learning. The reason for such research is that, while we appreciate the advantages of collaborative and peer assessment, there are other school of thought who believe that peers are sometimes perceived as less competent in assessment and therefore provide inexperience feedback (Usher & Barak, 2018). Others on the contrary, also believe that students participation in PA provide a good alternative to teacher assessment, particularly in large classes when the teacher is unable to examine each student individually (Misiejuk & Wasson, 2021). The study aims to compare and analyze the advantages and disadvantages of teacher assessment versus peer assessment of students' fashion design projects. Additionally, the research will examine the relationship between teacher and peer assessments to

determine whether a positive correlation exists between the two groups. Some educational research efforts have been made to compare classroom teaching methods, considering gender differences and variations among different schools, with the aim of enhancing pedagogical approaches in diverse ways (McKinney & Dong, 2022; Benini & de Oliveira Santos, 2018; Amos et al., 2023). However, there is limited research in the field of fashion education that examines and analyzes the significance of peer assessment as a human resource to support facilitation in large classes in higher education.

Teachers Assessment of Student Project Works

Assessing the quality of student project work is an essential aspect of teaching and learning. Teachers are responsible for ensuring that students have acquired the necessary knowledge and skills to succeed in their academic endeavors (Grek et al 2020). Therefore, it is crucial for teachers to assess the quality of student project work accurately. The researchers at this section seeks to explore the various methods of assessing the quality of student project work and the role of teachers in this process. Assessment of student project work can be done through a variety of methods such as self-assessment, peer assessment, and teacher assessment. Self-assessment involves students evaluating their own work and reflecting on their learning. Peer assessment involves students evaluating the work of their peers. Teacher assessment, on the other hand, involves teachers evaluating the quality of student project work.

Research has shown that teacher assessment is the most reliable method of assessing the quality of student project work (Pastore & Andrade, 2019: Dochy et al., 1999). Teachers have the necessary expertise and experience to assess the quality of student work accurately. However, it is important to note that the quality of teacher assessment depends on their understanding of the assessment criteria and their ability to provide constructive feedback (Ahyani et al, 2024: Black & Wiliam, 1998).

The role of teachers in assessing the quality of student project work

Teachers play a critical role in assessing the quality of student project work. They provide feedback to students on their work, which helps students to understand where they need to improve. Teachers also use assessment to identify gaps in their teaching and to improve their teaching methods (Kulal, 2024: Deluca & Klinger, 2010). Effective assessment requires teachers to have a good understanding of the assessment criteria and to provide constructive feedback to students. Research has shown that feedback that is specific, timely, and actionable can lead to significant improvements in student learning (Kinash, et al., 2015; Haughney, et al., 2020). Assessing the quality of student project work is an important aspect of teaching and learning. Teachers play a critical role in this process, and their expertise and experience are essential in ensuring that students receive accurate feedback on their work. Effective assessment requires teachers to have a good understanding of the assessment criteria and to provide constructive feedback to students.

The need for peer assessment

Systems that involve peer in assessment is becoming more prevalent due to the huge numbers of students enrolment in high education and the kind of complexities that facilitators are challenged with. A lot of facilitators monitor performance patterns of students in small class, but in large classes it is practically difficult (Cheng, et al., 2019), and so researchers are researching into alternative ways to assess students work and behavior. Again, receiving criticism from peers and providing feedback on their work can be beneficial to students. Additionally, some studies have showed that providing feedback is just as helpful as receiving it at enhancing ones abilities. According to Hornsby & Osman (2014), higher education is experiencing a significant increase in the number of students, which has led to an increase in the number of teaching and learning environments, which adversely affects student performance, motivation, and engagement as well as the quality of the educational experience. Again, the higher

institutions in the contemporary world are becoming more complex. With such complexity and huge number of enrollments today, most educators are researching into best managed human resource strategies to provide the needed creative solutions.

The Role of Peers in Assessing the Quality of Student Project Work

Student project work plays a crucial role in educational settings, allowing students to develop critical thinking, problem-solving, and collaboration skills. While teachers traditionally evaluate student projects, there is growing recognition of the value of incorporating peer assessment as an additional evaluation method (Alt & Raichel, 2020). Peer assessment involves students assessing the quality of their peers' work based on predetermined criteria.

Peer Assessment and its Benefits

Peer assessment provides students with the opportunity to engage actively in the evaluation process and develop a deeper understanding of the subject matter. It enhances students' metacognitive skills as they critically evaluate and compare their own work to that of their peers. Furthermore, peer assessment promotes a sense of ownership and responsibility among students, as they actively participate in shaping the learning environment.

1. Improvement of Critical Evaluation Skills

By evaluating their peers' work, students gain exposure to a variety of perspectives and approaches, which broadens their understanding of the subject matter. Engaging in critical evaluation helps students develop analytical skills, allowing them to identify strengths and weaknesses in project work. Through constructive feedback, students learn to communicate effectively and provide suggestions for improvement, fostering a culture of continuous learning (FasterCapital .com 2023).

2. Enhancement of Self-Reflection and Self-Assessment

Peer assessment encourages self-reflection and self-assessment among students. When students assess the work of their peers, they develop a better understanding of the assessment criteria and how it applies to their own work. This process prompts students to reflect on their own strengths and weaknesses and take steps to improve their performance. Peer assessment thus serves as a formative evaluation tool that empowers students to take ownership of their learning.

3. Promotion of Collaboration and Communication Skills

Incorporating peer assessment cultivates a collaborative learning environment. Students engage in discussions, share perspectives, and negotiate their assessments, fostering meaningful interactions. The process necessitates effective communication and argumentation skills, enabling students to articulate their thoughts and provide constructive feedback. Collaborative peer assessment not only enhances project outcomes but also prepares students for future professional environments where teamwork is essential.

4. Considerations and Challenges

While peer assessment offers numerous benefits, some considerations and challenges should be acknowledged. Students may initially lack confidence in assessing their peers' work or struggle to provide constructive feedback (Guardado & Shi, 2007; Liu & Sadler, 2003; Storch, 2019). Therefore, clear guidelines and training sessions should be provided to familiarize students with the assessment criteria and expectations. Additionally, ensuring fairness and preventing biases in the assessment process requires careful monitoring by teachers to maintain accuracy and reliability.

Finally, the role of peers in assessing the quality of student project work is multifaceted and impactful. Peer assessment fosters critical evaluation skills, enhances self-reflection and self-assessment, promotes collaboration, and develops effective communication skills. By actively engaging in the

evaluation process, students become active participants in their own learning journey. However, it is essential to address the challenges associated with peer assessment to ensure its successful implementation (Brookhart, 2017). As educators strive to provide holistic learning experiences, incorporating peer assessment into project evaluations proves to be a valuable approach. This paper therefore seeks to identify the need for peer assessment and scientifically compare the assessment of teachers and students (peers) to ascertain their relationship, differences and reliability for assessment in creative fashion design process.

2. METHODS

The study is aimed comparing the assessment of students and teaching staff to ascertain wether there are differences in the trend of assessment. The study used simple random sampling to sample 10 designs of Higher National Diploma (HND) two students enrolled in the fashion department of the Creative Arts Faculty. The rational for adopting this method was to ensures that each design has an equal chance of being chosen from the total group (Mweshi & Sakyi, 2020). The project was to design a creative suit with African print and the assessment were done by Peers (students) and Teaching staff. The designs were prepared on a google form with a scale of 2-10 for lecturers and students to assess on the internet. The system made use of the double-blind anonymity and was also convenient for assessors without the pressure and interreference of designers. The criteria for assessment were the same for all assessors which was adopted from the Functional Expressive and Aesthetic (FEA) model. The FEA model is a guide for students to design fashion project to meet customers demand within the context of design principles, functinability, self expression and culture.

The responses of both peers and the teaching staff were collected and prepared for analysis on google sheet. These were then subjected to Statistical analysis and the results reported for discussions. The study used Jamovi stats software to run an Analysis of frequencies and independent sample T-test to check the mean differences and statistically significant differences in the assessment of Peers and teaching staff.

The researchers presented 20 samples (Figure 1) of students' works to be assessed by peers and teaching stuff based on their perceived ratings of the works according to the following criteria: Self, FIT, COLOUR, DRAWING, and BEAUTY. And for the sake of time and space, ten designs were conveniently selected for discussion, which were designs 1-10.

Design 1 Design 2 Design 3 Design 4 Design 5

Design 6 Design 7 Design 8 Design 9 Design 10

Figure 1. Desigs 1-10: Samples of studet's work

Design 1

In design 1 the highest performance was colour, 8.67 and 6.16 for teaching staff and peers respectively. The lowest performance was Self Concept 5.78 and 4.55 respectively. The analysis indicates that for this design the ratings of the teaching stuff were higher than that of the students (peers) for all the criteria as shown in Figure 2.

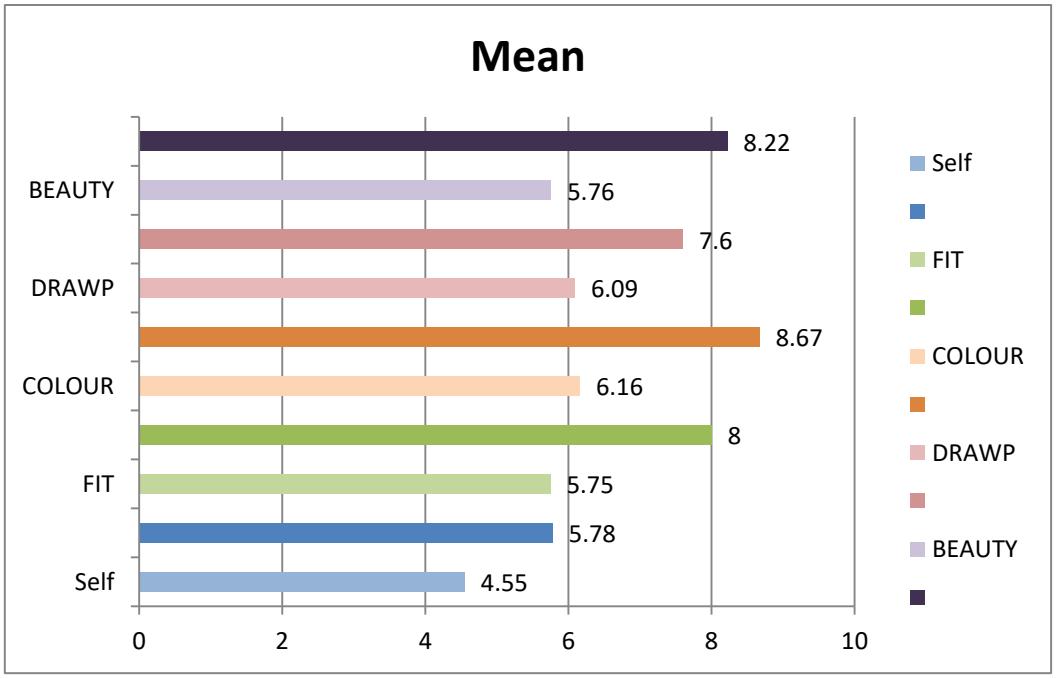

Figure 2. Rating of Design 1

Design 2

Design 2 shows that the highest performance was colour, 6.40 and 3.88 as rated by teaching staff and peers respectively. Also, the lowest performance was Drawing 5.6 for teaching staff and Self Concept 3.75 for peers. The analysis indicated that the ratings of the teaching staff were higher than that of the students (peers) for all the criteria as shown in figure 3.

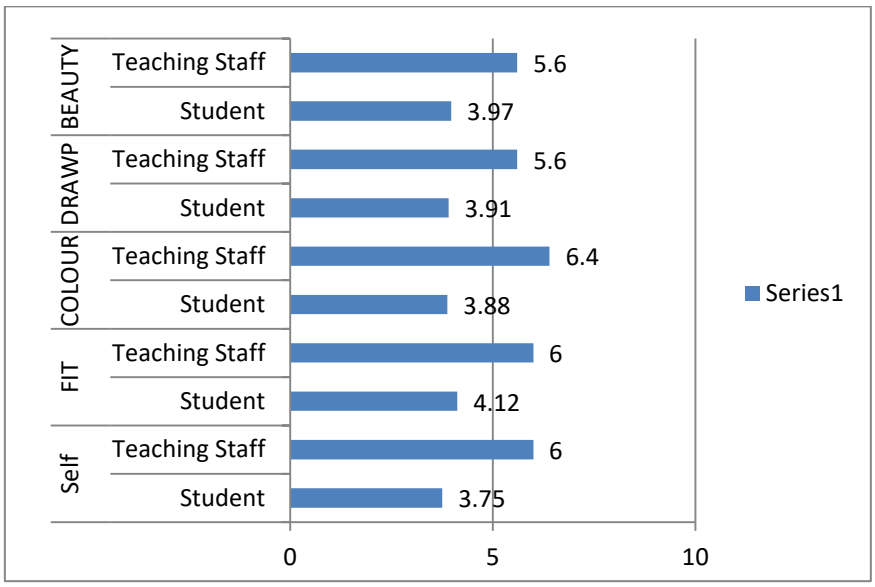

Figure 3. Rating of Design 2

Design 3

In Design 3, the Group Descriptive analysis provides valuable information about the central tendency and variability of the variables for each group. For the variable Self, both Students (peers) and Teaching Staff had mean scores above 4.0. The Teaching Staff had a slightly higher mean (5.00) compared to the Student (4.65). The median scores were similar in both groups, indicating a relatively balanced distribution. The standard deviation (SD) and standard error (SE) values were higher in the Teaching Staff group, suggesting greater variability in their ratings compared to Students.

Similarly, for the variables FIT, DRAWP, and BEAUTY, the mean scores were slightly higher for the Teaching Staff group compared to the Student group, although the differences were relatively small. The median scores were generally consistent across the groups, indicating a similar distribution of ratings. The standard deviation and standard error values again showed higher variability in the Teaching Staff group.

For the variable COLOUR, the Teaching Staff group had a notably higher mean (6.20) compared to the Student group (5.12). The median score for the Teaching Staff group was also higher (7.00) compared to the Student group (5.00), suggesting a skewed distribution. Additionally, the standard deviation and standard error values were larger for the Teaching Staff group, indicating greater variability in their ratings (Figure 4).

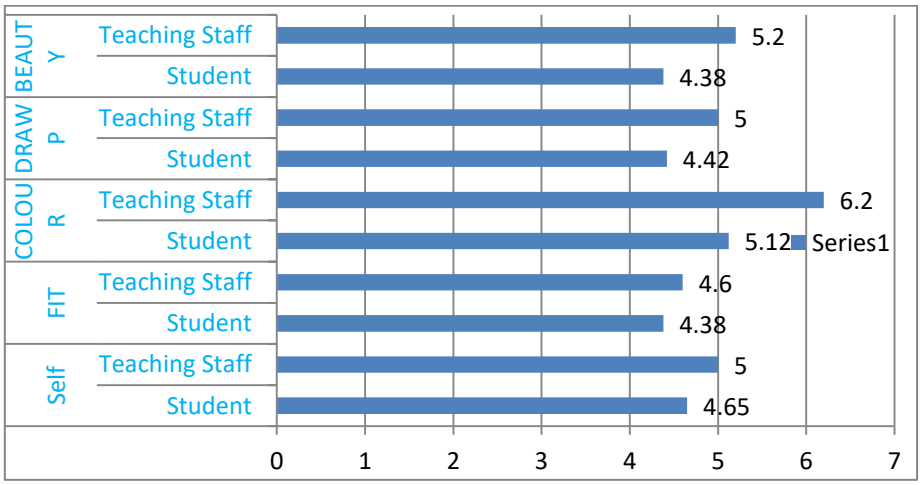

Figure 4. Rating of Design 3

Design 4

In design 4 Teaching staff consistently rate the designs higher than students (peers) across all categories, with the most significant difference observed in the COLOUR category (7.80 vs. 6.16). Student groups generally showed higher variability (larger SD) compared to teaching staff, suggesting less consistent self-ratings. Again, Student groups exhibit a central tendency towards a median of 6.00 across most categories, with some exceptions like BEAUTY (5.00).

Design 5

For the variable Self (Self concept) shown in figure 5, the Students' group (N=65) has a mean of 2.89, indicating a relatively low average score. The Teaching Staff group (N=10) has a slightly higher mean of 3.20. The median scores for both groups are lower than the respective means, suggesting some degree of skewness in the data. The standard deviations (SD) for both groups indicate moderate variability in the scores, with the Teaching Staff group showing a higher SD.

Similar patterns are observed for the variables FIT, COLOUR, DRAWING, and BEAUTY. The mean scores for the Student group are generally lower than those for the Teaching Staff group. The median scores and standard deviations also exhibit variations between the two groups, indicating differences in the central tendency and dispersion of the data.

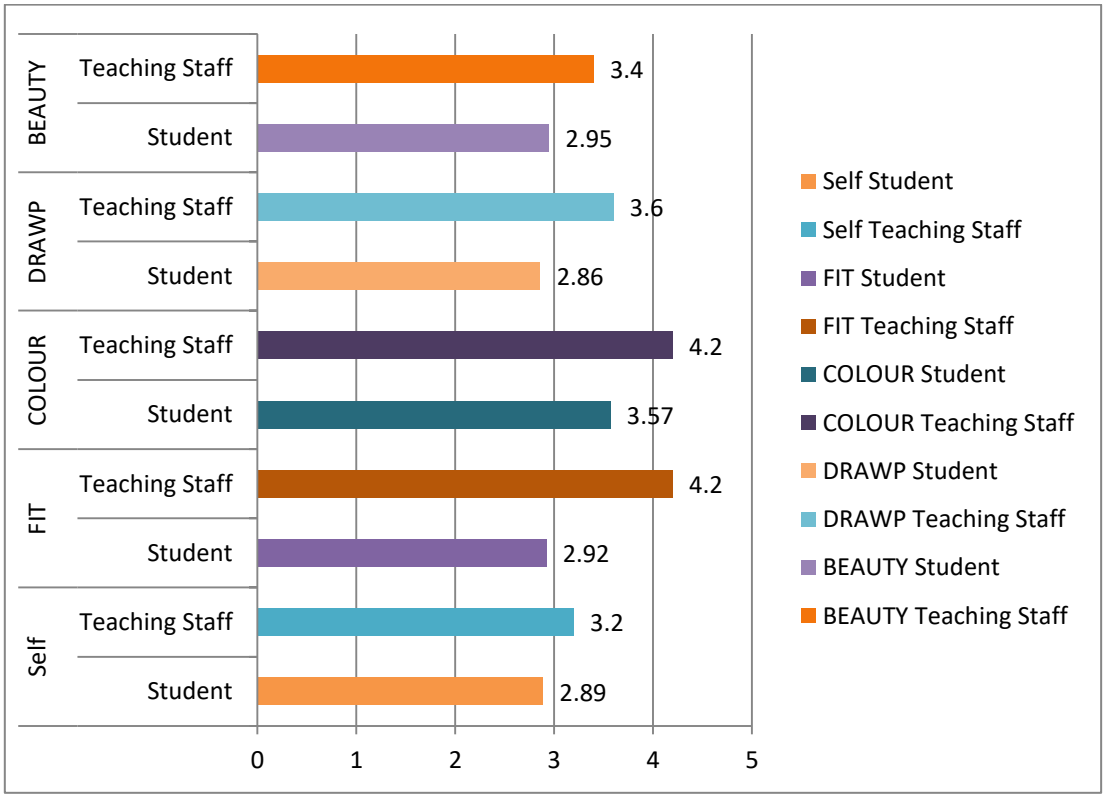

Figure 5. Rating of Design 5

Design 6

The findings from the Group Descriptives chart (figure 6) suggest that the teaching staff groups have a more positive assessment than the student groups, except for FIT. With COLOUR, the teaching Staff group has the highest mean score for all categories. This suggests that the teaching staff is particularly positive about their assessment. The variable Self has the lowest mean score for all categories. This suggests that both staff and students rated the design less than the other variables.

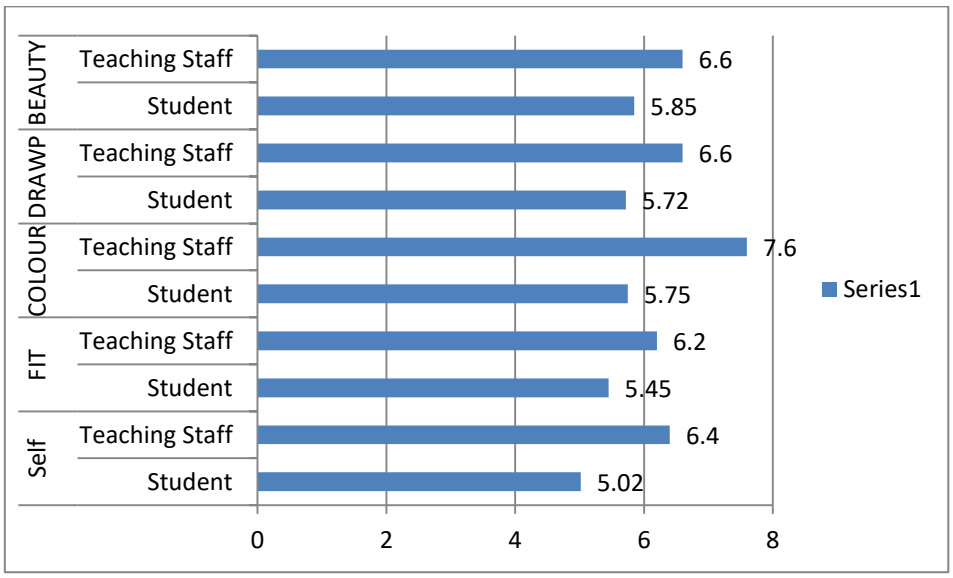

Figure 6. Rating of Design 6

Design 7

Design seven (figure 7) shows that, there is a lot of variability in the data. The means for the different groups range from 3.80 to 7.00, and the standard deviations range from 1.65 to 2.62. The medians are also quite variable, ranging from 4.00 to 8.00. It is also interesting to note that the teaching staff generally have higher means and medians than the students. This suggests that the teaching staff may have a better understanding of the concepts being measured. Overall, this data suggests that there is a lot of variability in the way that people perceive students' projects.

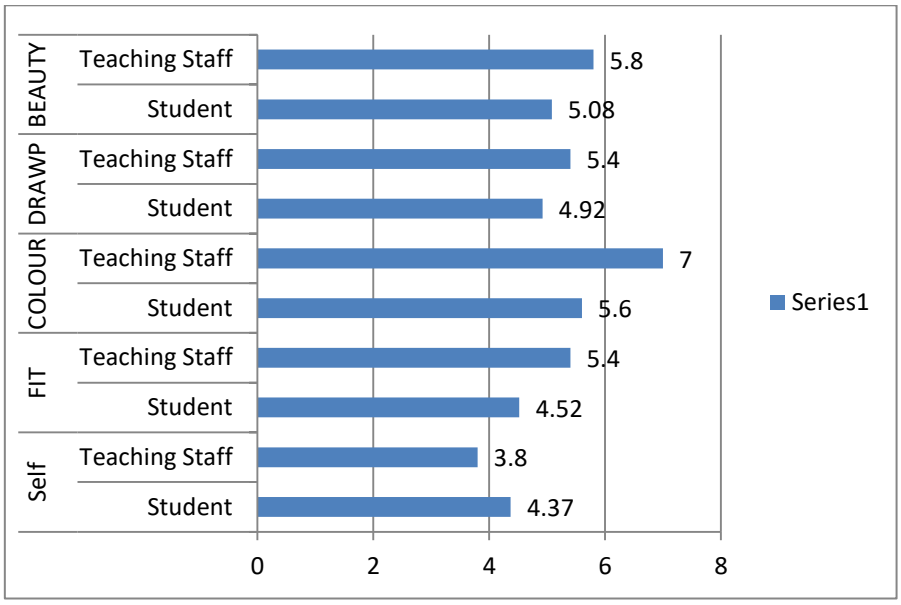

Figure 7. Rating of Design 7

Design 8

According to figure 8 the Teaching staff members rated the design slightly higher (4.00) than students (3.85) on average. Both groups have similar variability (SD around 2.0) and precision (SE around 0.25). For FIT vs. COLOUR vs. DRAWP vs. BEAUTY: Teaching staff consistently rated the design higher than students across all categories. The largest difference in means is observed in the COLOUR category, where

teaching staff (6.20) rated design significantly higher than students (4.43). Students in the FIT category have the lowest standard deviation (1.85) and standard error (0.229), indicating less variability and more precise ratings.

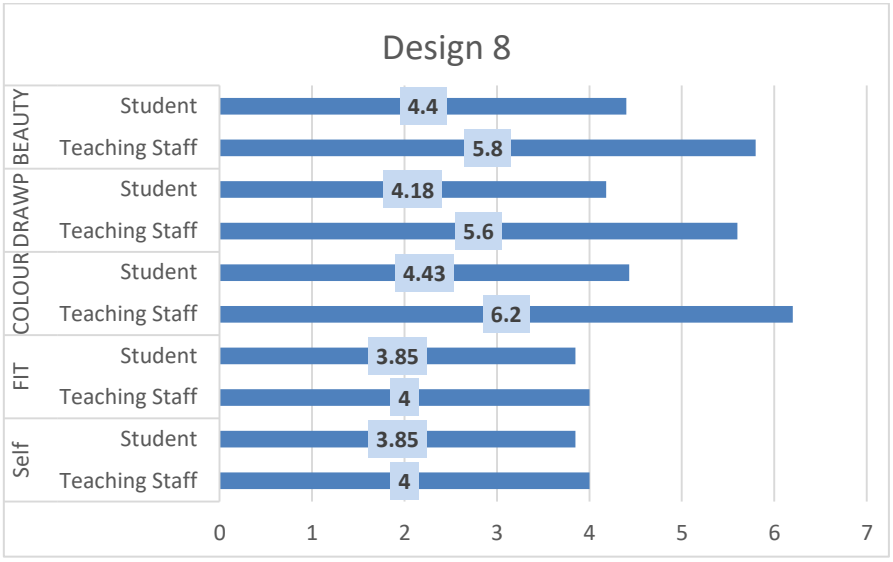

Figure 8. Rating of Design 8

Design 9

Figure 9 that teaching staff members again rated the design higher than students across all categories. The largest difference in means between teaching staff and students is observed in the FIT category (5.80 vs. 4.18). Students exhibit a more consistent pattern of self-rating across categories compared to teaching staff (probably because of the large number as compared to teaching staff numbers. Student groups generally have lower standard deviations and standard errors compared to teaching staff, indicating less variability and higher precision in their self-ratings. The FIT category for teaching staff shows the highest standard deviation (3.05), suggesting greater variability in their perception. Again, Student group assessment tend to have median values of 4.00 across all categories, suggesting a central tendency towards that value. Both groups show similar variability and precision. Teaching staff have a slightly higher mean rating than students (5.33 vs. 4.28) with similar variability and precision.

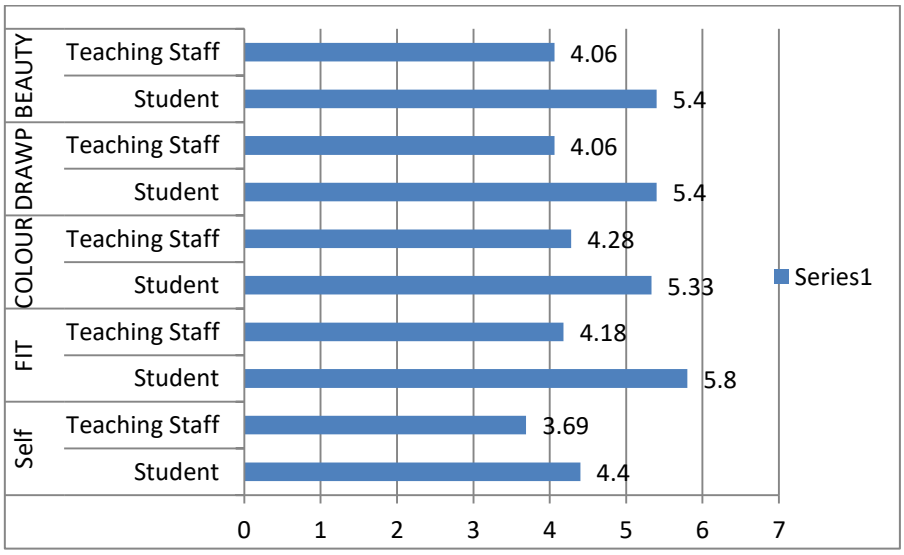

Figure 9. Rating of Design 9

Design 10

This report provides a summary of the descriptive statistics for different groups across several categories. From figure 10, teaching staff rated the designs considerably higher than students across all categories (7.40 vs. 5.82 on average). The largest difference in means between teaching staff and students is observed in the COLOUR category (8.60 vs. 6.68). Students showed a similar pattern of rating across categories, with median values consistently at 6.00. Both student and teaching staff groups showed moderate variability within each category, with standard deviations ranging from 1.75 to 2.64. The mean values for both groups are consistently higher than the median values, suggesting a positive skew in the data. This skewness is more pronounced for teaching staff, with mean values exceeding the median by as much as 1.20 points.

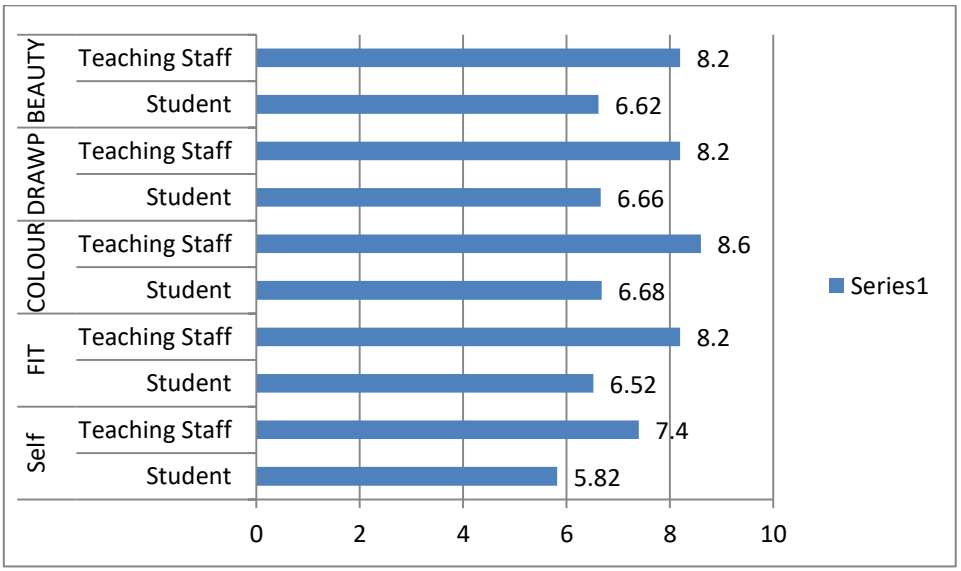

Figure 10. Rating of Design 10

Correlation test

The researchers sought to determine the relationship between the variables in the data to establish the reliability of the data to make a justified conclusions on the findings and results generated. According to Pearson's correlation co-efficient factor show inTable 1, all variables exhibit strong positive correlations with each other, with coefficients ranging from 0.674 to 0.883.

1. Highest Correlations
   a) DRAWN and BEAUTY have the highest correlation (0.883), indicating a very strong relationship between these two variables.
   b) FIT and DRAWN also show a strong positive correlation (0.810).
   c) Similarly, COLOUR and DRAWN are highly correlated (0.820).

2. Specific Findings
   a) FIT is strongly correlated with all other variables ($r > 0.790$), indicating that individuals who perceive the design as fitter tend to also rate designs higher in other categories like COLOUR, DRAWN, and BEAUTY.
   b) DRAWN and BEAUTY are very strongly correlated ($r = 0.883$), suggesting a close relationship between drawing skills and overall attractiveness.
   c) All correlations are statistically significant (p-value < 0.001), indicating a high level of confidence in the observed relationships.

Table 1. Correlation Matrix

Criteria		r	p-value
Self	Fit	0.72	< .001
Self	Colour	0.70	< .001
Self	Draw	0.70	< .001
Self	Beauty	0.70	< .001
Fit	Colour	0.80	< .001
Fit	Draw	0.81	< .001
Fit	Beauty	0.80	< .001
Colour	Draw	0.82	< .001
Colour	Beauty	0.82	< .001
Draw	Beauty	0.88	< .001

The Independent Samples T-Test

The independent samples t-test, also known as the two-sample t-test, is a statistical method used to compare the means of two independent groups (American Psychological Association 2020). According to Kline (2015) the Independent Samples T-Test is a statistical method used to determine if there are any significant differences between the means of two independent groups. It is a parametric test that assumes the data is normally distributed and that the variances of the two groups are equal. As indicated by Kline (2015) , the researchers wanted to find out by statistical and means if the variances of the two groups are equal. The Independent Samples T-Test was applied to five variables: SELF, FIT, COLOUR, DRAWN, and BEAUTY. Each variable was subjected to Student's t-test, with the respective statistics presented in Table 2.

Additionally, it was noted that for the FIT variable, Levene's test was significant ($p < .05$), indicating a violation of the assumption of equal variances. The results indicate that for all variables, there are statistically significant differences between the means of the two independent groups. The p-values for all variables are less than .001, indicating strong evidence against the null hypothesis, which suggests no difference between the means of the groups.

Table 2. Independent Samples T-Test

		Statistic	df	p
SELF	Student's t	-3.82	1501	< .001
FIT	Student's t	-6.33 [a]	1500	< .001
COLOUR	Student's t	-7.71	1482	< .001
DRAWN	Student's t	-6.58	1496	< .001
BEAUTY	Student's t	-6.78	1494	< .001

[a] Levene's test is significant ($p < .05$), suggesting a violation of the assumption of equal variances

3. GENERAL DISCUSSION

Descriptive analysis

The Descriptives analysis revealed differences in the mean scores and variability between the Student and Teaching Staff groups for the variables Self, COLOUR, FIT, DRAWP, and BEAUTY. The Teaching Staff group generally had slightly higher mean scores, while also exhibiting greater variability

compared to the students (peers) group and according to Kaufman & Schunn (2011) this may be attributed to inexperience. These findings provide insights into the rating patterns within the two groups and may have implications for understanding their perceptions or experiences related to the examined variables as noted by Kline (2015). The largest difference in means between staff and peers across all designs was observed in the COLOUR category. Again, the highest rated variable was COLOUR suggesting that students skill in colour combination seems to be high as compare to other variables. The second highest rated variable across all criteria was BEAUTY(Aesthetics).

The analysis showed differences in the mean scores and variability between the Student and Teaching Staff groups. However, there was a similar trend in the way the two groups assess the work in one criterion as compared to the other across all designs. Papanthymou and Darra (2018) found that students can equally assess themselves accurately if they are trained how to assess themselves. In connection with the above statement, the researchers have observed over the years that students need to be taken through the peer review processes to generate the needed results as prescribed in Calibrated Peer Review system. This also imply that the idea of involving students in the assesment processes can be harness well to support facilitators in high education. The study further revealed that involving students in the assessment process can significantly support facilitators in higher education by distributing the workload, fostering deeper learning, and promoting critical thinking and self-reflection among students. This approach leverages the collective expertise and diverse perspectives of students, making the assessment process more comprehensive and inclusive (Hundley & Keith, 2023). Students show a more consistent pattern in their ratings across categories for almost all design and this is probably due to the large number of students who assessed the works.

T-test analysis

The findings suggest that there are significant differences between the groups concerning the variables analysed. One of the main objectives of this paper was to evaluate an alternative assessment technique like peer assessment which affords collaboration, communication, or teamwork skills. The independent samples t-test allowed the researchers to compare the effectiveness of these different assessment techniques in capturing various aspects of student learning (Hundley & Keith, 2023; Falchikov & Boud, 2000). In the study, both teacher and peer assessments when subjected to the independent samples t-test helped assess inter-rater reliability, which refers to the consistency between different evaluators (Fink, 2010). Ultimately, this is contributing to a more comprehensive understanding of student learning and performance. The findings revealed a significant difference between the two sets of scores. In a related study, the researchers identified that these differences were attributed to several reasons: Students competence to assess peers work - Several peer assessment works have been done in combination with Teaching and learning analytics by the researchers and have revealed that students need to be trained well to assess peers works. Students confidence in assessing peers works - Most at times, students are quite skeptical when awarding marks to peers and tend to give low marks. In such situations it is importantant for further studies to be conducted to diagnose and comprehensively outline the causes of the disparities. These limitations could be attributed to time constrains. The researchers envisage further studies in the form of diagnostic analytics to carefully follow the trends in students assessment to make a more informed decision in the future.

Correlation analysis

This analysis reveals strong positive correlations between various self-perception categories, suggesting that both assessors tend to have a consistent overall self-perception across different aspects. McMillan (2018) emphasised that correlation plays a vital role in understanding the level of agreement

between teacher and peer assessments. Though the descriptive analysis shows that there are differences in the assessments of teachers and peers. However, the study shows that the high positive correlation indicates that both teachers and peers tend to assign similar scores, suggesting consistency in their evaluation criteria. On the other hand, a low correlation and negative correlation could have suggested significant discrepancies and bias in how teachers & peers judge student work (investopedia.com, 2023**;** Bangert et al., 2004; Brookhart, 2017).

The Analysis of the correlation between assessments informed instructional decisions (Yancey et al., 2009). In other words, it highlighted potential issues for the researchers to adjust their assessments or provide more opportunities for peer feedback in these areas (Topping, 2013). Additionally, a strong positive correlation between teacher and peer assessments lend validity to both methods, suggesting they capture similar aspects of student learning (Shuell, 1990). Finally, correlation was a valuable tool for comparing teacher and peer assessments. It helps us gauge agreement, identify potential biases, inform instructional decisions, validate assessment methods, and gain a richer understanding of student learning (Popham, 2008). The relationship of the two variables makes the data reliable and can be used as the bases for future studies in this area.

4. CONCLUSION

Assessment of fashion and textiles projects in higher education is crucial for effective output of students practical skill. The study sought to make a comparison of fashion project assessment among staff and students. With ten sampled projects, the views of both students and staff were gathered and analysed to acertain the relationship between the assessment level of the participants. A focus was placed on variables such as SELF, FIT, COLOUR, DRAWN, and BEAUTY. From the assessment, the teaching staff had a slightly higher mean score with greater inconsistency as compared to the assessmet given by the students. However, a large difference in means was observed on the variable 'COLOUR' across all designs. Findings from the T-test analysis suggest that there are significant differences between the groups concerning the variables analysed. The correlation analysis also revealed strong positive correlations between the various self-perception categories. Collaborative assessment, especially among peers is recommended to help students build confidence and ensure a sense of ownership of their projects coupled with sharpening their practical skill.

It is concluded that, peer assessment (PA) can significantly enhance students' assessment skills and offer diverse perspectives on project quality within fashion design education. By understanding the differences in assessment between teachers and peers, overall evaluation practices and grading fairness can be improved. Collaborative assessment methods can also foster student confidence and ownership of their projects. Therefore, it is recommended that educators integrate peer assessment alongside traditional teacher assessment methods to provide comprehensive feedback. Additionally, providing training and guidelines for both students and teachers is essential to optimize the effectiveness of peer assessment. Lastly, facilitators should emphasize the importance of consistency and calibration in assessment criteria to ensure reliability across all assessors.

Ethical consideration

Before the research, participants were informed about the purpose, procedures and benefits of the study. They were assured that participation in the research is purely voluntary and data gathered from them will be kept strictly confidential.

5. REFERENCES

Ahyani, E., Rifky, S., Muharam, S., Murnaka, N. P., Kadiyo, K., & Dhuhani, E. M. (2024). Implementation of Teacher Performance Management to Improve Education Quality. *Cendikia: Media Jurnal Ilmiah Pendidikan*, *14*(4), 391-398.

Alt, D., & Raichel, N. (2020). Higher education students' perceptions of and attitudes towards peer assessment in multicultural classrooms. *The Asia-Pacific Education Researcher*, 29, 567-580.

American Psychological Association (2020). Publication Manual of the American Psychological Association (7th ed.). https://apastyle.apa.org/.

Amos, E., Essel, H. B., Fobiri, G. K., & Ibrahim, M. (2023). Analysis on Students' Performance to Promote Gender Equality in Creative Fashion Design. *Edukasiana: Jurnal Inovasi Pendidikan*, *2*(4), 260-271.

Bangert, A. R., Kulik, J. A., & Kulik, C.-L. C. (2004). Peer assessment in elementary and secondary education: A meta-analysis. *Educational Psychologist*, 39(1), 3–14.

Benini, A. L., & de Oliveira Santos, M. (2018). Application of a different approach in ergonomics studies within fashion design education, results comparison between two different institutions. In *EDULEARN18 Proceedings* (pp. 8728-8733). IATED.

Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability (formerly: Journal of personnel evaluation in education)*, 21, 5-31.

Brookhart, S. M. (2017). Improving teacher grading: Using feedback to promote student learning. Alexandria, VA: ASCD.

Cheng, L., Ritzhaupt, A. D., & Antonenko, P. (2019). Effects of the flipped classroom instructional strategy on students' learning outcomes: A meta-analysis. *Educational Technology Research and Development*, 67, 793-824.

Danielson, C. (2007). Enhancing professional practice: A framework for teaching. AsCD.

DeLuca, C., & Klinger, D. A. (2010). Assessment literacy development: Identifying gaps in teacher candidates' learning. *Assessment in Education: Principles, Policy & Practice,* 17(4), 419-438.

Dochy, F., Segers, M., & Sluijsmans, D. (1999). The use of self-, peer and co-assessment in higher education: a review. *Studies in Higher Education*, 24 (3), 331-350.

Falchikov, N., & Boud, D. (2000). Peer learning and assessment. Routledge.

FasterCapital .com (2023). Peer Assessment: Strength in Numbers: The Impact of Peer Assessment, Dubai. https://fastercapital.com/content/Peer-Assessment--Strength-in-Numbers--The-Impact-of-Peer-Assessment.html.

Fink, A. (2010) International Encyclopedia of Education (Third Edition), Sconce Direct.

Grek, S., Maroy, C., & Verger, A. (Eds.). (2020). *World yearbook of education 2021: Accountability and datafication in the governance of education*. Routledge.

Guardado, M., & Shi, L. (2007). ESL students' experiences of online peer feedback. *Computers and Composition*, 24, 443–461.

Haughney, K., Wakeman, S., & Hart, L. (2020). Quality of feedback in higher education: A review of literature. *Education Sciences*, 10(3), 60.

Hornsby, D. J., & Osman, R. (2014). Massification in higher education: Large classes and student learning. *Higher Education*, 67, 711–719. https://doi.org/10.1007/s10734-014-9733-1.

Iglesias Pérez, M. C., Vidal-Puga, J., & Pino Juste, M. R. (2022). The role of self and peer assessment in Higher Education. *Studies in Higher Education*, *47*(3), 683-692.

Kaufman, J. H., & Schunn, C. D. (2011). Students' perceptions about peer assessment for writing: Their origin and impact on revision work. *Instructional science*, 39, 387-406.

Kinash, S., Naidu, V., Knight, D., Judd, M. M., Nair, C. S., Booth, S., ... & Tulloch, M. (2015). Student feedback: a learning and teaching performance indicator. *Quality Assurance in Education*, 23(4), 410-428.

Kline, R. B. (2015). Thinking statistically in the social sciences (5th ed.). Routledge.

Kulal, A., Nanjundaswamy, A., Dinesh, S., Suraj, N., & Mallika, N. (2024). Advancing teacher competencies: assessing the influence of faculty development programs (FDP) on embracing modern innovations in teaching. *Journal of Applied Research in Higher Education*

Liu, J., & Sadler, R. W. (2003). The effect and affect of peer review in electronic versus traditional modes on L2 writing. *Journal of English for academic Purposes*, 2(3), 193-227.

McKinney, E., & Dong, H. (2022). Fashion design students' self-evaluation of usage, benefits, and attitudes toward using design research sketchbook skills: a scaffolded pedagogical approach. *International Journal of Fashion Design, Technology and Education*, *15*(3), 407-417.

McMillan, J. H. (2018). Classroom assessment: Principles and practice for effective teaching (7th ed.). Pearson.

Misiejuk, K., & Wasson, B. (2021). Backward evaluation in peer assessment: A scoping review. *Computers & Education*, 175, 104319.

Papanthymou, A., & Darra, M. (2018). Student Self-Assessment in Higher Education: The International Experience and The Greek Example. *World journal of education*, 8(6), 130-146.

Popham, W. J. (2008). Comprehensive classroom assessment: A reader in theory and practice (2nd ed.). Allyn & Bacon.

Pastore, S., & Andrade, H. L. (2019). Teacher assessment literacy: A three-dimensional model. Teaching and Teacher Education, 84, 128–138

Sadler, D. R. (2014). Beyond feedback: Developing student capability in complex appraisal. In Approaches to assessment that enhance learning in higher education (pp. 45-60). Routledge.

Sadler, D.R. (2009). Transforming Holistic Assessment and Grading into a Vehicle for Complex Learning. In: Joughin, G. (eds) Assessment, Learning and Judgement in Higher Education. Springer, Dordrecht. https://doi.org/10.1007/978-1-4020-8905-3_4.

Shuell, T. J. (1990). Peers as assessors: A review of research on peer evaluation in higher education. *Studies in Educational Evaluation*, 16(1), 17–37.

Storch, N. (2019). Collaborative writing as peer feedback. Feedback in second language writing: Contexts and issues, 143-161.

Topping, K. J. (2013). Peer assessment and learning. In J. H. McMillan (Ed.), Classroom assessment: Principles and practice for effective teaching (7th ed.) (pp. 475–504). Pearson.

Usher, M., & Barak, M. (2018). Peer assessment in a project-based engineering course: comparing between on-campus and online learning environments. *Assessment & Evaluation in Higher Education*, *43*(5), 745-759.

Yancey, K. B., Knight, T. R., & Rife, M. (2009). Four lenses on student writing: Balancing perspectives for assessment. Jossey-Bass.