DOI: 10.38027/ICCAUA2021120N2

An Analysis of Grading Discrepancy in Architectural Juries: The Case of Summer School Design Courses

Dr. **Milorad Pavlovic**
Alanya Hamdullah Emin Pasa University, Faculty of Architecture, Alanya, Turkey
E-mail: milorad.pavlovic@ahep.edu.tr
ORCID: https://orcid.org/0000-0003-4631-7735

Abstract
The architectural education process is fundamentally dominated by design studio courses. Their organization remains an open challenge for educators due to the complexity of related factors, such as teaching methods, design topics, assignments and assessments. In particular, the assessment of design studio courses is often a complex procedure, due to the involvement of several jury members with different expertise and their subjective interpretations. In this research, an analysis of the grading discrepancy in architectural juries is proposed. The study aims on quantifying the divergence between the final grades and the marks proposed by each juror. The percentage error was considered as a method for research. The procedure was introduced in the vertical architectural design studio at Alanya Hamdullah Emin Pasa University, Department of Architecture, during the summer semester of 2018. The paper presents the results of the analysis, author's observations, and proposes further developments of the research.
Keywords: Architecture; Architectural Design; Architectural Studio; Assessment Analysis; Grading Discrepancy.

1. Introduction

The Design Studio course is well-known for its importance in architectural education. Generally, this course devotes the majority of the students' weekly workload due to its learning-by-doing character and the highest number of credits that are currently due. Since the early development of architectural education, the master-apprentice relationship played the primary role in the atelier's atmosphere, enhancing the character of learning-by-doing (Dutton, 1987). The atelier remained the center of education at the beginning of the twentieth century, with the Bauhaus institution, whereas in the 1980s the present-day model of education has been established (Wayne & Mugerauer, 1991). Such a model, which has been widely adopted for decades, has become the symbol of architectural education (Tzonis, 2014).

Another peculiar characteristic is that the design studio courses can be considered as the meeting point of the knowledge acquired from different disciplines and applied to the solution of a practical problem. This heterogeneous character of the Design Studio course is also at the base of its complex social and cultural aspects (Datta, 2007; Goldschmidt et al 2010). Hence, the role of critiques and juries should not be ignored as a significant part of architectural education (Oh et al, 2013; De La Harpe et al, 2009).

However, in more recent years, several debates have been opened about the validity of the existing models in the present days' educational systems. For example, the condition of the web-based educational practices has been analyzed from the pedagogical (Kvan, 2001) and organizational (Sagun et al, 2008) points of view. Similarly, the objectives of the studio courses have been re-examined, as well as students' learning styles (Kvan & Jia, 2005; Charalambous & Christou, 2016). Likewise, the problem of assessing and grading has been furtherly examined, focusing the attention on differences in grading parameters (Utabeta et al, 2012; Alagbeet al 2015; Ragheb, 2016).

Starting from the considerations mentioned above and considering the framework of the studio course proposed in Figure 1 by the author (Pavlovic, 2020), the present research intends to focus on the problems of grading and its subjective interpretation. In particular, the study intends to propose, a preliminary analysis for quantifying the divergence between the subjective interpretation of the single lecturer and the grade assigned by the jury. By analyzing such discrepancy, the aim is to outline the common discrepancy that governs the studio course Juries. The procedure has been developed and applied during the summer semester of 2018 at Alanya Hamdullah Emin Pasa University, Department of Architecture, program in the English language. The paper illustrates at first the procedure and proposes a study case, whereas in the 2nd part the results are analyzed and further development is considered.

Figure 1. Contextualization of the research.

2. Materials and Methods

As stated above, the present research has been conducted to quantify the divergences between the assigned grade and the grades expressed by jurors. The analysis has been conducted during a 15 lectures semester, based on 2 intermediate juries and a final one. As a method of analysis, the percentage error has been considered as a tool to evaluate the discrepancy between the assigned final grade (decided by the jury) and the individual grades (expressed by each juror).

As well known, the percentage error represents the difference between a measured or experimental value and a true or exact value (Swamidass, 2000). For the present analysis, the grade assigned by the jury has been assumed as exact value, whereas the individual grades expressed by jurors have been considered as experimental values. The grades have been assigned on a scale from 0 to 100, with steps of 10 points corresponding to each variation, e.g., 80-89 = BA, 90-100 = AA. Therefore, besides the percentage error analysis a second analysis, the tolerance interval (Meeker et al, 2017), has been performed to quantify the percentage of grades that expresses the matching within the tolerance of 10%, corresponding to a single grade variation. The analysis has been conducted considering the grades of 29 undergraduate students who attended the course. The students have been divided into 3 groups accordingly to their level, i.e., Group 1 - Freshman, Group 2 - Sophomore, and Group 3 - Junior. For every jury, each of the jurors expressed an own grade, whereas for each submitted design a grade has been successively decided by the jury. As an example, below is proposed the analysis conducted on Group 1.

2.1. Freshmen Group

The first course to be analyzed has been conducted by 2 lecturers with different backgrounds and unknown to each other. As regards the 1st Midterm Jury (Figure 2a), Lecturer 1 has been recorded an average grading discrepancy of 29.5%, whereas for Lecturer 2 the gap reached an average order of 42.4%. Successively, the 2nd Midterm Jury (Figure 2b) registered a lowering trend in terms of discrepancy. In particular, for Lecturer 1 it decreased to an average value of 12.8%, whereas for Lecturer 2 the gap was in the order of 9.4%. As regards the final jury (Figure 2c) here the analysis has shown a much closer gap expressed by the jurors. In particular, for Lecturer 1 the average error has been recorded in order of 9.62%, whereas for the Lecturer 2, of 11.8%. From this analysis, it emerged that the first Midterm Jury has been characterized by the highest level of discrepancy, whereas the 2nd Midterm Jury and the Final Jury confirmed the same trend of discrepancy. Lastly, the tolerance analysis has been conducted (Figure 2d) and from it emerged that throughout the course only 20% of the grades have been assigned within the same grade criteria.

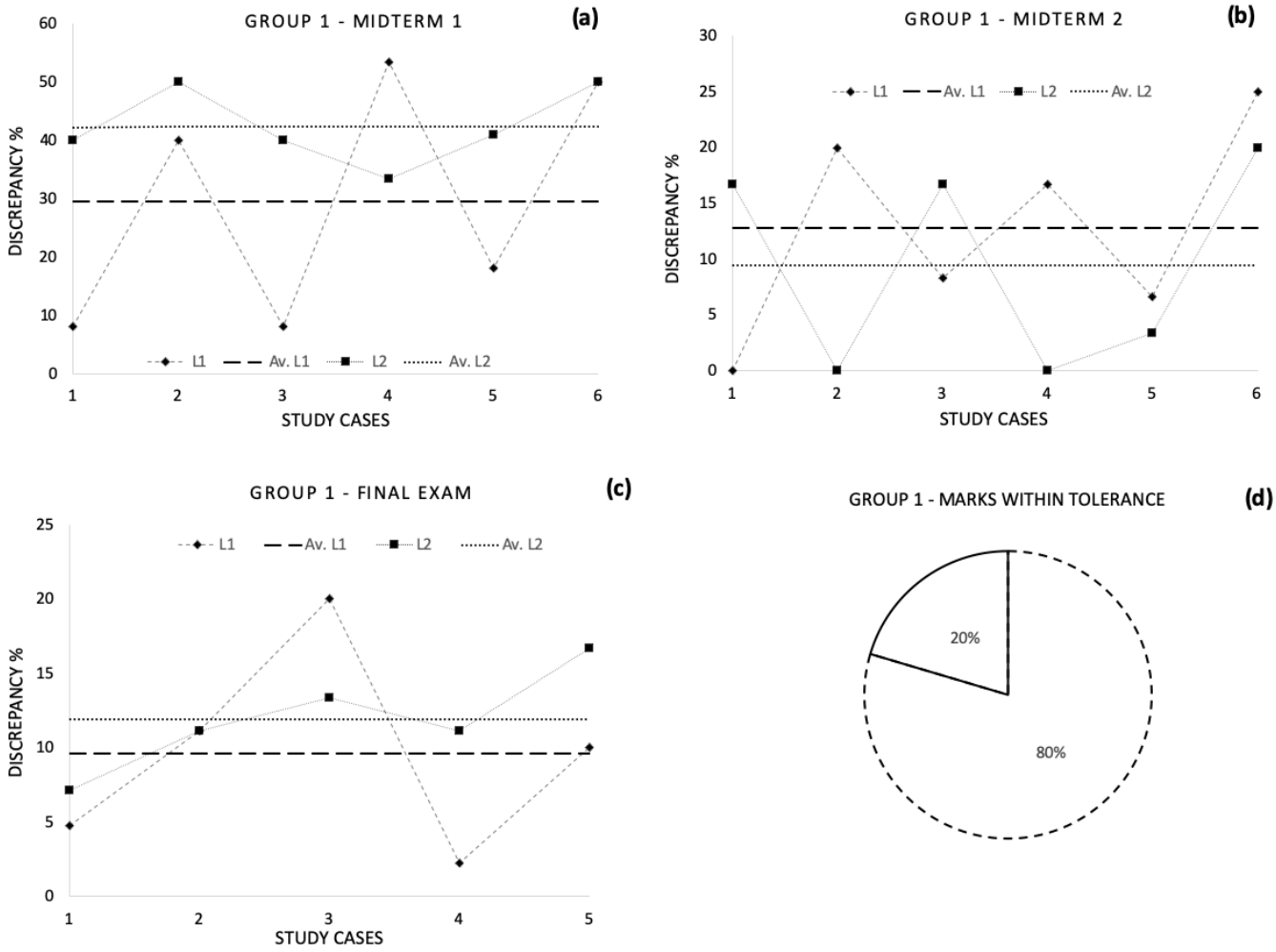

Figure 2. Group 1 - Grading discrepancy and tolerance analyses.

3. Analysis and Results

Successive analyses have been applied on the marks collected from the lecturers who joined the juries of the Sophomore and Junior groups. In both cases, the midterm evaluations have been performed by 2 lecturers, whereas the final juries have been conducted by 4 lecturers. At first, are presented the analyses conducted on the grades assigned to the single groups whereas successively a comparison has been made considering all the participants in each jury.

3.1. Sophomore Group

The second set of grades has been analyzed for both midterm juries and the final one. For the 1st Midterm, the analysis recorded an average grading discrepancy of 29.9% as regards Lecturer 1, whereas for Lecturer 2 the gap was of the order of 15.9% (Figure 3a). Successively, the 2nd Midterm Jury (Figure 3b) registered a lowering trend in terms of discrepancy. For Lecturer 1 it decreased to an average value of 18.8%, whereas for Lecturer 2 the gap was in the order of 12.8%. As regards the analysis of the grades proposed in the final jury (Figure 3c) here a jury of 4 lecturers expressed their grades at first and successively discussed over the assigned grade. The results of this analysis showed that for Lecturer 1 the average grading discrepancy is of the order of 5.5%, 16% for both Lecturer 2 and Lecturer 3, whereas for Lecturer 4 the average grading discrepancy recorded was 20%. The successive tolerance analyses, for the accepted mismatch, showed that 25% of the grades fall into the same grade range (Figure 3d).

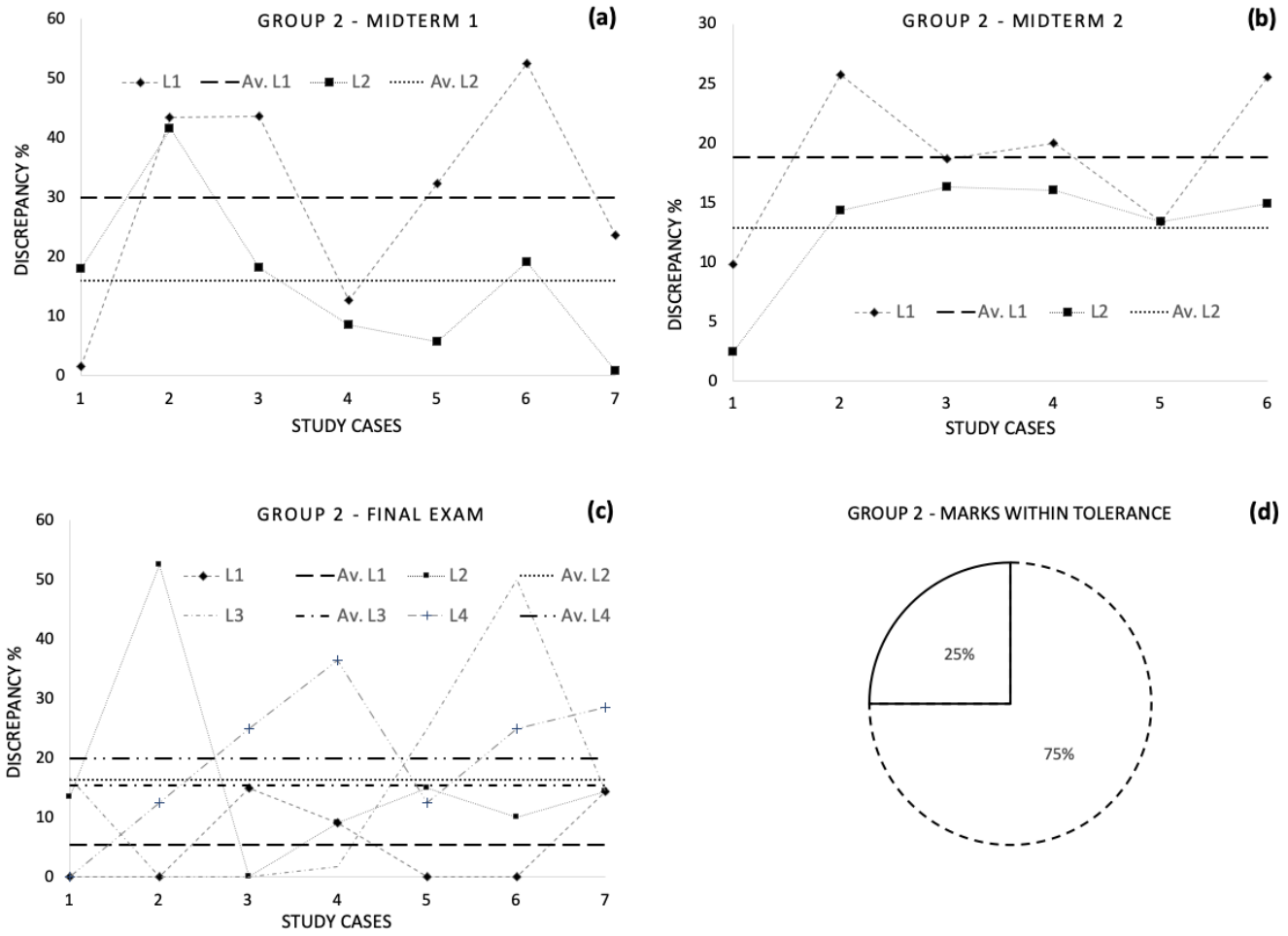

Figure 3. Group 2 - Grading discrepancy and tolerance analyses.

3.2. Junior Group

Lastly and similarly to the previous data, the 3rd set of grades has been analyzed for both midterm juries and the final one. In particular, it emerged that for the 1st Midterm Jury the analysis recorded an average grading discrepancy of 17.8% as regards Lecturer 1, whereas for Lecturer 2 the gap was of the order of 16.3% (Figure 4a). Successively, the 2nd Midterm Jury (Figure 4b) registered an uprising trend in terms of discrepancy. For Lecturer 1 it increased to an average value of 24%, whereas for Lecturer 2 the gap was in the order of 19.6%. As regards the analysis of the grades proposed in the final jury (Figure 4c) here a jury of 4 lecturers expressed their grades at first and successively discussed over the assigned grade. The results of this analysis showed that for Lecturer 1 the average grading discrepancy is of the order of 7.4%, 10% for Lecturer 2, whereas values of 20.8% and 12.2% have been recorded respectively for Lecturer 3 and Lecturer 4. The successive tolerance analysis (Figure 4d) showed that 26% of the grades respect the defined tolerance mismatch.

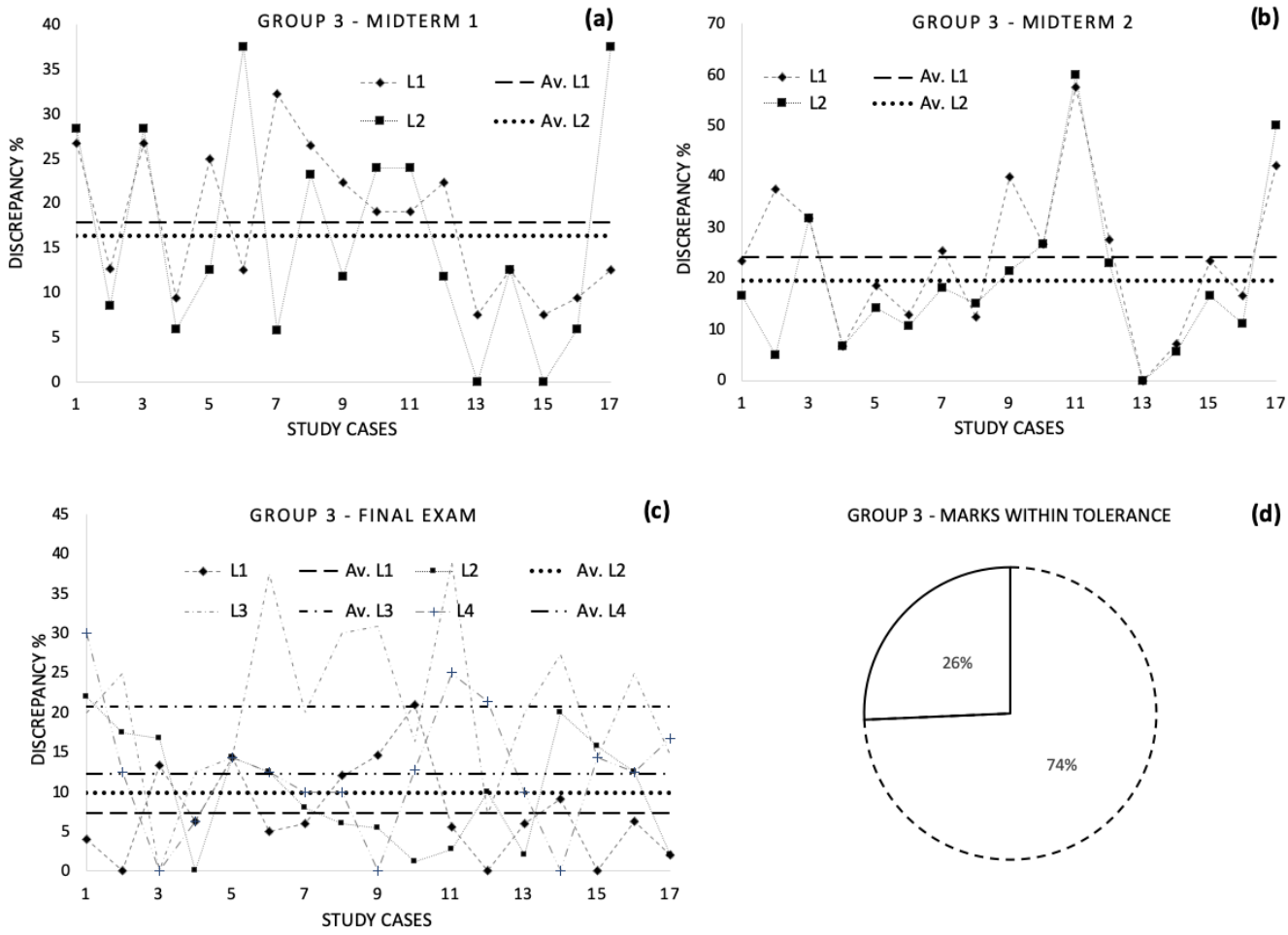

Figure 4. Group 3 - Grading discrepancy and tolerance analyses.

3.3. Overall Analyses

After the first group of analyses in which the grades of each course and each exam have been compared individually, the 2nd set of analyses has been performed considering all the grades assigned during each jury. Here the purpose was to analyze if and how the trends in grading discrepancy and tolerance change through the semester. Hence, an initial grading discrepancy analysis has been performed considering the results of the 1st Midterm Jury for all the Groups (Figure 5a). The analyses showed average errors of 23% and 21.4% respectively for Lecturer 1 and Lecturer 2. Highlighting an average error discrepancy of 22.2%. Similarly, the analysis conducted on the marks expressed in the 2nd Midterm jury outlined an average grading discrepancy of 20.6% for Lecturer 1 and of 16.1% for Lecturer 2 (Figure 5c). These results highlighted an average error discrepancy of 18.4%. However, in both the case of Midterm 1 and Midterm 2, the tolerance analyses remarked values value of 26% (Figure 5 b, d). Lastly, the analysis conducted on the sets of marks assigned in the final juries (Figure 6a) highlighted values of grading discrepancy in the order of 7.3% (Lecturer 1), 11.8% (Lecturer 2), 15.4% (Lecturer 3) and, 19.2% (Lecturer 4) respectively, with an average error order of 13.4%. Moreover, in this case, the tolerance analysis outlined a value of 28%, slightly higher than the Midterm Juries (Figure 6b).

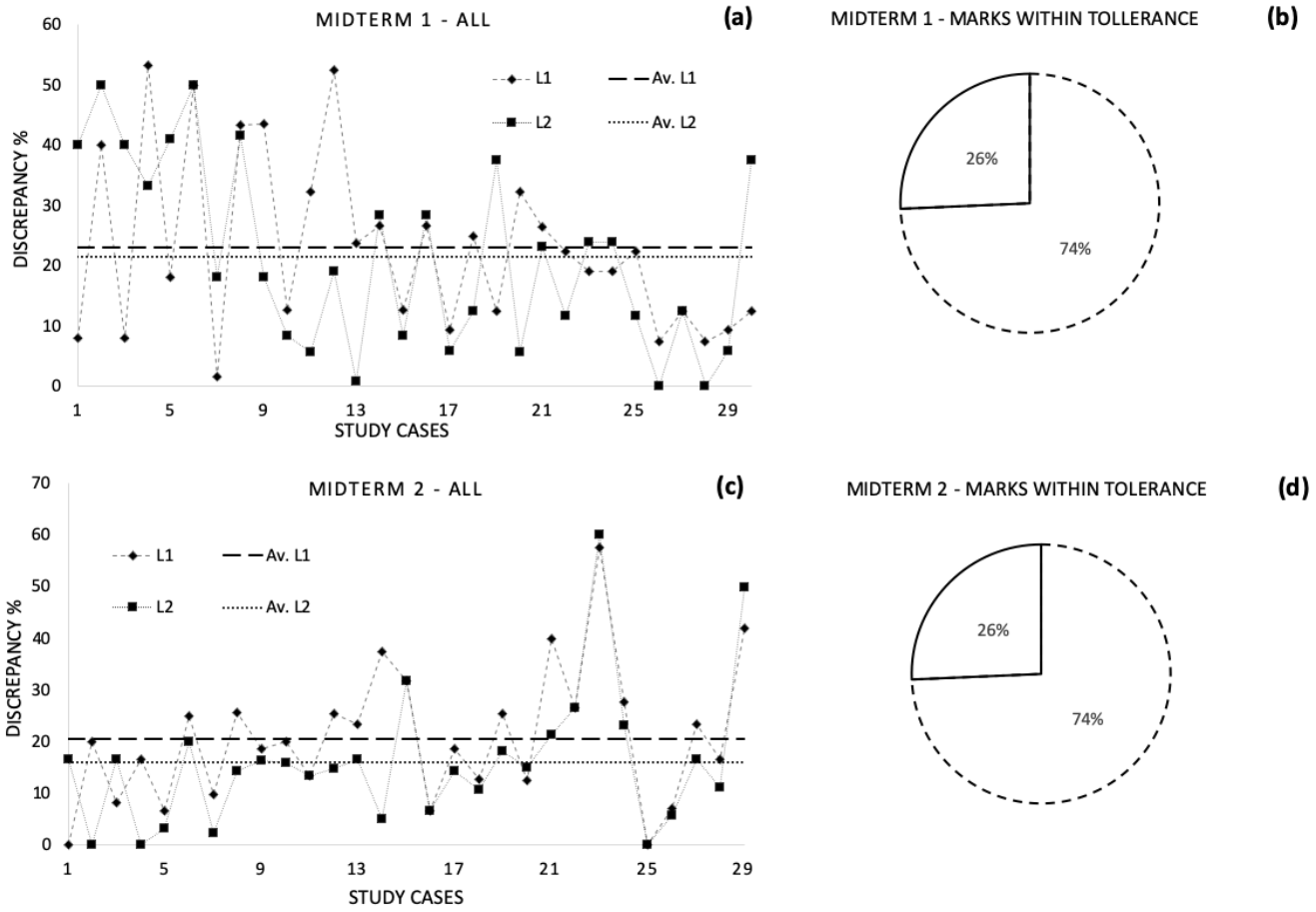

Figure 5. Midterm Juries - Grading discrepancy and tolerance analyses.

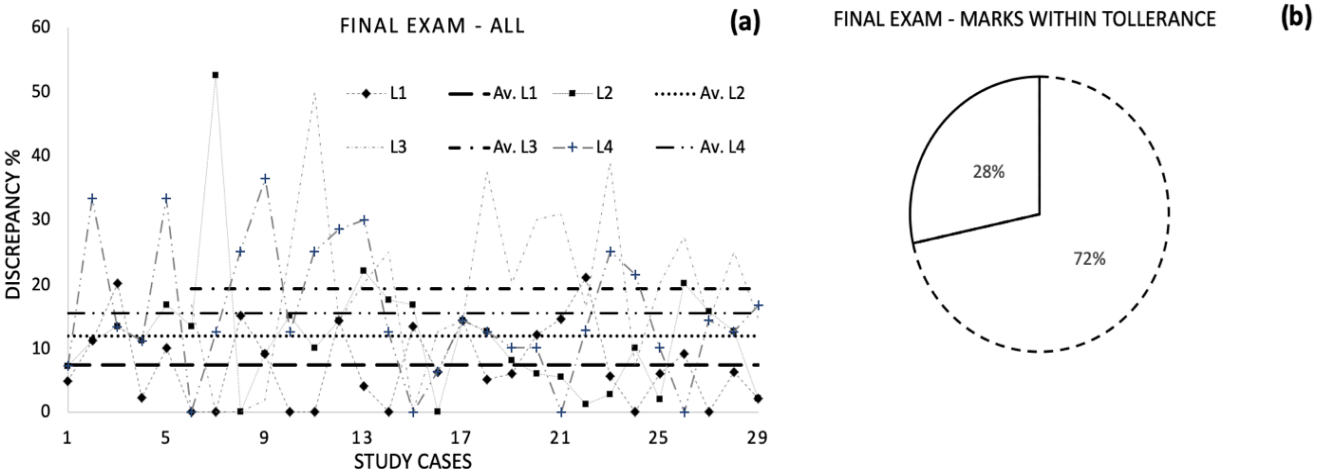

Figure 6. Final Juries - Grading discrepancy and tolerance analyses.

4. Discussion

4.1. Results of the analyses

The analyses conducted in this experiment focused the attention on the grading discrepancy and tolerance analyses in an architectural design studio course. In particular, the findings have been oriented toward the intermediate and final juries. Observing the results summarized in Table 1, it can be highlighted that the highest discrepancy values have been recorded in the 1st Midterm Jury, whereas the lowest in the Final Jury (25.3% Vs 12.9%). Similarly, the tolerance analysis showed an

increment proportioned to the seniority of the students (20% for the Freshmen Vs 26% for the Juniors). Even though the average discrepancy for all the Midterm Jury analyses can be assumed to be of the order of 20%, it is significant to highlight the gap of almost 30% between the maximum and the minimum values recorded in terms of discrepancy (42.4% Vs 9.4%). A mismatch, that in the case of Final Juries is significantly reduced to the order of 15% (in the range of 20.8% Vs 5.5%).

Table 1. Analyses' summary

	Grading Discrepancy (%)								*Tolerance Analysis*
	Midterm 1		*Midterm 2*		*Final Jury*				*(%)*
	L1	*L2*	*L1*	*L2*	*L1*	*L2*	*L3*	*L4*	
Group 1	29.5	42.4	12.8	9,4	11.8	9.6	/	/	20
Group 2	29.9	15.9	18.8	12.8	5.5	16.0	16.0	20.0	25
Group 3	17.8	16.3	24.0	19.6	7.4	10.0	20.8	12.2	26
Average	25.3		16.2		12.9				23.7

4.2. Author's opinions

Considering the results of the analyses, it can be highlighted a significant discrepancy characterized the grading process. However, the division of the course into 3 parts, allowed to reduce this gap significantly and progressively. This fact might be attributed to 2 different conditions, one related to students' performance and one related to lecturers' teamwork. As regards the students' performance and, as well known, the project definition rarely follows a linear trend line and most of the time it is highlighted by an exponential improvement in the final part. Hence, a higher definition of the design parameters and process' allowed a more detailed and objective evaluation. On the other hand, considering the different expertise and backgrounds of the jurors, it can be assumed that the numerical values corresponding to the subjective evaluation and interpretation vary significantly up to the progress of the design process. Besides, as long as the course continues and the teamwork of the lecturers/jurors increases in terms of cooperation, the standards and parameters can be more commonly redefined. These aspects might be assumed at the base of decreasing the grading discrepancy analyzed.

5. Conclusion

In this research, an analysis of the grading discrepancy in architectural juries has been proposed. The study was developed to quantify the divergence between the final grades and the marks proposed by each juror during an exam session. The percentage error and tolerance interval were considered tools for analyses. The whole procedure has been applied in the vertical architectural design studio at Alanya Hamdullah Emin Pasa University, Department of Architecture, during the summer semester of 2018.

The analyses outlined variable levels of grading discrepancy and tolerance interval, which however showed a decreasing trend toward the end of the semester. If in the 1st Midterm Jury, the average discrepancy error was of the order of 25.3%, this has been reduced to 16.2% in the 2nd Midterm Jury and 12.9% in the Final Jury. Similarly, the tolerance analysis showed the overall percentage of grades which mismatch was lower than 10%, progressively increased throughout the course.

As preliminary research, the procedure showed encouraging results, however, since the methodology has been applied only once, to understand its reliability it would be beneficial to apply the model to a higher number of jurors, preferably with different backgrounds. If the discrepancy trend remains confirmed, for further studies a standardized grading method might be worth definition and recommendation.

Acknowledgment

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Conflict of interests

The authors declare no conflict of interest.

References

Alagbe, O.A., Oluwatayo, A.A., Aderonmu, P.A., Alalade, G.M. (2015). Difference in grading parameters in architectural schools and its impact on the competency rating of future professionals. *Frontiers of Architectural Research, 4*(3), 230-236. https://doi.org/10.1016/j.foar.2015.06.001

Charalambous, N., Christou, N. (2016). Re-adjusting the Objectives of Architectural Education. *Procedia - Social and Behavioral Sciences, 228*(1), 375-382. https://doi.org/10.1016/j.sbspro.2016.07.056

Datta, A. (2007). Gender and Learning in the Design Studio. *Journal for Education in the Built Environment, 2*(2), 21-35. https://doi.org/10.11120/jebe.2007.02020021

De La Harpe, B., Peterson, J.F., Frankham, N., Zehner, R., Neale, D., Musgrave, E., McDermott, R. (2009). Assessment Focus in Studio: What is Most Prominent in Architecture, Art and Design? *Journal of Art & Design Education, 28*(1), 37-51. https://doi.org/10.1111/j.1476-8070.2009.01591.x

Dutton, T.A. (1987). Design and Studio Pedagogy. *Journal of Architectural Education, 41*(1), 16-25. https://doi.org/10.2307/1424904

Goldschmidt, G., Hochman, H., & Dafni, I. (2010). The design studio "crit": Teacher–student communication. *Artificial Intelligence for Engineering Design, Analysis and Manufacturing, 24*(3), 285-302. https://doi.org/10.1017/S089006041000020X

Kvan, T. (2001). The pedagogy of virtual design studios. *Automation in Construction, 1*(3), 345-353. https://doi.org/10.1016/S0926-5805(00)00051-0

Kvan, T., Jia, Y. (2005). Students' learning styles and their correlation with performance in architectural design studio. *Design Studies, 26*(1), 19-34. https://doi.org/10.1016/j.destud.2004.06.004

Meeker, W.Q., Hahn, G.J., Escobar, L.A. (2017). In *Statistical Intervals: A Guide for Practitioners and Researchers, Second Ed.* John Wiley & Sons. ISBN 9781118594841. https://doi.org/10.1002/9781118594841

Oh, Y., Ishizaki, S., Gross, M.D., Yi-Luen Do, E. (2013). A theoretical framework of design critiquing in architecture studios. *Design Studies, 34*(3), 302-325. https://doi.org/10.1016/j.destud.2012.08.004

Pavlovic, M. (2020). An Experimental Approach to the Sophomore Architectural Design Studio. *Civil Engineering and Architecture, 8*(5), 942-949. https://doi.org/10.13189/cea.2020.080521

Ragheb, G. (2016). Assessment in Architectural Design Education Case Study: First-Year Students at Pharos University. Journal of Civil Engineering and Architecture*, 10*(9), 1015-1026. https://doi.org/10.17265/1934-7359/2016.09.005

Sagun, A., Demirkan, H., Goktepe, M. (2008). A Framework for the Design Studio in Web-Based Education. *Journal of Art & Design Education, 20*(3), 332-342. https://doi.org/10.1111/1468-5949.00282

Swamidass, P.M. (2000). *MEAN ABSOLUTE PERCENTAGE ERROR (MAPE). In: Swamidass P.M. (eds) Encyclopedia of Production and Manufacturing Management.* Boston, MA: Springer. ISBN 978-1-4020-0612-8. https://doi.org/10.1007/1-4020-0612-8_580

Tzonis, A. (2014). Architectural education at the crossroads. *Frontiers of Architectural Research, 3*(1), 76-78. https://doi.org/10.1016/j.foar.2014.01.001

Utabeta, N., Hassanpour, B., Zaharim, A., Spalie, N. (2012). Re-assesing Criteria-Based Assessment in Architecture Design Studio. In: *Thaung K. (eds) Advanced Information Technology in Education. Advances in Intelligent and Soft Computing, vol 126. Springer, Berlin, Heidelberg.* https://doi.org/10.1007/978-3-642-25908-1_30

Wayne, A., Mugerauer, R. (1991). Excellent studio teaching in architecture. *Studies in Higher Education*, *16*(1), 41-50. https://doi.org/10.1080/03075079112331383081