Marques *et al.*
Journal of New Approaches in Educational Research (2025) 14:9
https://doi.org/10.1007/s44322-025-00030-3

Journal of New Approaches in Educational Research

ORIGINAL RESEARCH Open Access

Beyond bias in student satisfaction surveys: exploring the role of grades and satisfaction with the learning design

Francielle Marques[1*], Davinia Hernández-Leo[1] and Carlos Castillo[1,2]

*Correspondence:
francielle.marques@upf.edu

[1] Departament of Engineering, Universitat Pompeu Fabra, Barcelona, Catalonia, Spain
[2] ICREA, Barcelona, Catalonia, Spain

Abstract

Course satisfaction surveys play a relevant role in Higher Education, aiding in the quality assessment of courses and informing academic promotions. Nonetheless, understanding potential biases and influential factors within these surveys is crucial to their equitable utilization within universities. This study delves into a deconstruction of satisfaction ratings considering three learning design factors (content, methodology, and workload) and their interplay with student grades. Especially emphasizing the need for institutional analytics to engage in Generative Uncertainty, aiding productive inquiries using data. Institutional analytics of the 2021–2022 and 2022–2023 survey results from a Spanish university revealed that learning design aspects strongly correlate with students' holistic perception of a course. The correlation between student grades and student satisfaction related to learning design is either weak or moderate. These analytical findings imply that there may be bias in students' responses to course satisfaction surveys (e.g., lower grades leading to lower satisfaction). However, this bias doesn't consistently manifest.

Keywords: Learning Design, Learning Environment, Learner satisfaction, Institutional Analytics, Biases

1 Introduction

Student satisfaction surveys are widely used to assess the students' view of higher education. Although the measure helps higher education institutions make more informed decisions about professor's promotions, quality of courses, and teaching, student satisfaction can have systematic variations in different dimensions. Several authors have indicated the multifaceted aspect of student satisfaction.

The effects can be related to undesirable factors, such as class size, course, gender of the professor, type of class, course levels, and disciplines, among others (Bedggood & Donovan, 2012; Radchenko, 2020; Radmacher & Martin, 2001). Although there are studies reflecting on the impact of learning design and performance on student satisfaction (Li et al., 2016; Rienties & Toetenel, 2016; Rienties et al., 2015), more evidence is needed in diverse contexts and with large datasets, as there is still a widely spread belief that the notion that professors who award higher grades receive more favorable

Springer

© The Author(s) 2025. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

student evaluations (Griffin, 2004; Remedios & Lieberman, 2008; Stumpf & Freedman, 1979). The study by Rosen (2018) drew attention to the importance of understanding underlying correlations and trends in student satisfaction. Centra (2003) showed, using regression analysis to estimate the expected grade, that student satisfaction was not impacted by teachers giving higher grades, and lower course work had minimal effects on course evaluations. The authors concluded that instructional practices and design affected student feedback and performance. However, to the best of our knowledge, there is still a missing deconstruction about the effects of satisfaction with different facets of the learning design and its implications in students' activities - such as the workload.

Indeed, frameworks that connect learning analytics with learning design high- light the relevance of considering several types of data collected from students (performance, satisfaction, context) to inform understanding of how to improve learning environments (Hernández-Leoet al., 2019). The learning design comprises the choices made when planning for learning in a course, such as content, methodology (type of learning tasks, structure of teaching sessions, and materials), and imply (expected and actual) students' workload (Laurillard, 2013). Measurements about students' satisfaction with the learning design offer a perspective of insights and support teaching and academics decision-making (Burns & Corwin, 2017; Hernández-Leo et al., 2019).

We can categorize the effects that may influence student satisfaction into three main groups:

1. Factors outside the control of the instructor
    - Instructor demographics
    - Class size
    - Class level
2. Factors within the control of the instructor
    - Methodology
    - Workload
    - Compliance of obligations
    - Teaching plan
3. Student performance
    - Average performance
    - Distribution of performance

By categorizing the effects, we can have a more comprehensive understanding of these factors. Factors present in group 1 may influence student satisfaction and should be considered to mitigate possible unwanted biases. Factors in group 2 are the ones that can possibly be improved. Factors in group 3 are more difficult to interpret, and some authors seem to assume they also introduce unwanted biases, with the expectation that teaching satisfaction should be independent of student performance. In fact, this requires a more nuanced interpretation. We believe that within some range, it is ok to expect that teaching satisfaction should be independent of student

performance, but courses in which students perform exceptionally poorly or exceptionally well should be looked at with care.

In this frame, we state that integrated analytics of measurements considering students' performance (such as their grades) and an in-depth analysis of students' answers to course satisfaction surveys that focus on elements related to learning design can further shed light on interrelated potential biases.

In this study, we will use the perspective of institutional analytics using a dataset provided by Pompeu Fabra University, a university in Barcelona, Spain. The study is also original in that the analytics emphasizes a deconstruction of satisfaction ratings considering three learning design factors (content, methodology, workload) and their interplay with learning performance measurements.

2 Background

For higher education institutions, effective institutional analytics has become essential (Wise et al., 2023), using data to inform and design solutions for better decision-making. This approach encompasses a range of applications, such as student enrollment, student satisfaction, admissions, curriculum development, and faculty performance. In fact, institutional analytics can help the University community create a more responsive and effective educational environment.

While student satisfaction surveys are widely used to inform institutions about promotions and awards for university staff, several studies have found biases related to gender (Marques et. al (2024); Heffernan, 2023; Mengel et al., 2018), level of the course, and the age of the professor (Flegl & Andrade Rosas, 2019), for example. Research in this area shows conflicting results (Clayson, 2009), indicating that bias may be tied to the context and time, revealing the need for continuous analyses with the aim of finding potential implications of biases. In the context of institutional analytics, we state that it is important to consider and study the tools (i.e., surveys, dashboards), observing and analyzing deeper for a better understanding of different factors (Ortiz-Beltrán et al., 2022). Campaña et al. (2016) analyzed differences in student satisfaction by field of study, also using socio-demographic characteristics, factors related to expectations, and motivations. Torrado and Blanca (2022) assessed student satisfaction with online courses, creating a Spanish version of the Learner Satisfaction Survey. Unlike previous research in the context of the Spanish higher education system, our study goes further in the analysis, integrating data related to performance and analyzing extreme cases of satisfaction with workload.

Similarly to learning analytics, institutional analytics should enable institutions to engage in "Generative Uncertainty," (Campos et al., 2024), an interpretive stance in which productive inquiries into data are manifested. Using assumptions on the presence of uncertainty to validate data analyses while generating patterns of asking new inquiries under uncertainty. It allows the continuous development of hypotheses that lead to further reflection and action during the process of decision-making and improvement in higher education. For institutional analytics, generative uncertainty stances can create a space for discussions to uncover better solutions to existing problems, but also to produce new institutional knowledge due to the debate generated.

Further, for a comprehensive analysis, it's essential to integrate diverse factors, allowing a nuanced identification of trends and patterns in student satisfaction and enhancing factors that outline cases that go beyond the trend. In this paper, we will illustrate different co-founders, such as student satisfaction with learning design and grades, and we will delve into a study of workload to understand the courses that do not follow a pattern in student satisfaction.

3 Research questions

This paper investigates how factors of the educational setting and student performance interplay with student satisfaction with the learning design. Our analysis considers an in-depth analysis of the notion of "satisfaction with the learning design" as a subset of the satisfaction items in course surveys that focuses separately on the content, the teaching methodology, and the perceived workload. Our main research question is: "Do grades measuring student performance bias student satisfaction ratings with the learning design?". To further understand, we will investigate the following research questions:

1. RQ1—How does student satisfaction with learning design affect overall satisfaction with the course measures?
2. RQ2—Do grades bias the student satisfaction with the learning design?
3. RQ3—Does the context influence student satisfaction with the learning design?

4 Methods

4.1 Context and data collection

In this work, we will refer to a "subject" (e.g., introductory physics), and a "group" (e.g., group A in introductory physics). Because some subjects have many students, especially in the first and second year, subjects may be divided into multiple groups. Our unit of analysis is a group.

4.1.1 Student satisfaction survey

We use a dataset recording answers to the official student satisfaction survey during the 2021-2022 and 2022-2023 academic years at Pompeu Fabra University with a total of 3066 bachelor-level groups distributed in 25 degrees. We use a dataset that considers only groups with a sufficient number of responses.

We adopt a method introduced in "eXplorance Blue", which is a software package for course evaluation surveys that has been used by the university[1]. This method, which they name Reliability Assessment Score (RAS) (1), is based on the bound of the error of estimation in "Elementary Survey Sampling" (section 4.3) (Scheaffer et al., 2011), setting the standard deviation (s) to 1.1.

We keep only groups with RAS ≤ 0.5.

$$RAS = 2\sqrt{\frac{1.1^2}{n} \cdot \frac{N-n}{N}} \tag{1}$$

If:

- $0 \leq RAS < 0.25$ is considered to have a good number of responses.

- $0.25 \leq RAS \leq 0.5$ is considered to have a sufficient number of responses.
- $RAS > 0.5$ is considered to have an insufficient number of responses.

Our final dataset contains 586 bachelor-level groups distributed in 9 study centers or schools (communication, economics and business, engineering, health and life sciences, humanities, law, law and economics, political and social sciences, translation, and language sciences) offering 25 degrees.

In this study, we focus on the survey questions related to methodology (MTD: "The teaching methodologies used (design of class sessions, activities to be carried out by students, and teaching materials used) have helped me in the learning process."), workload (WLD: "The volume of work required is adapted to the credits of the subject."), satisfaction with teaching (TEA: "I am satisfied with the teaching received."), as well as the overall satisfaction with the subject (SBJ: "In general terms, I am satisfied with this subject.").

The remaining survey questions are related to teaching compliance (OBL: "The teacher has adequately fulfilled their teaching obligations (teaching plan, punctuality, attention to students, etc.)") and compliance with the teaching plan (PLN: "The contents taught within the subject and the competencies worked are adjusted to what the teaching plan says".). Students answer these questions on a numerical scale from zero (complete dissatisfaction) to ten (complete satisfaction). The survey contains six questions and is applied entirely through an online application that students access using their university credentials.

Towards the end of each quarter or semester, but before the final exams, each student receives an email with a link to the assessment surveys. There is one survey for each subject-group in which they have been enrolled during that quarter or semester. Student surveys can be submitted until the period for final exams begin; most courses have one or more assessments during the quarter or semester, hence student performance is not completely determined by final exams. Professors receive a summary of the student evaluations weeks after the quarter or semester ends, so these evaluations can not influence their grading of the final exams.

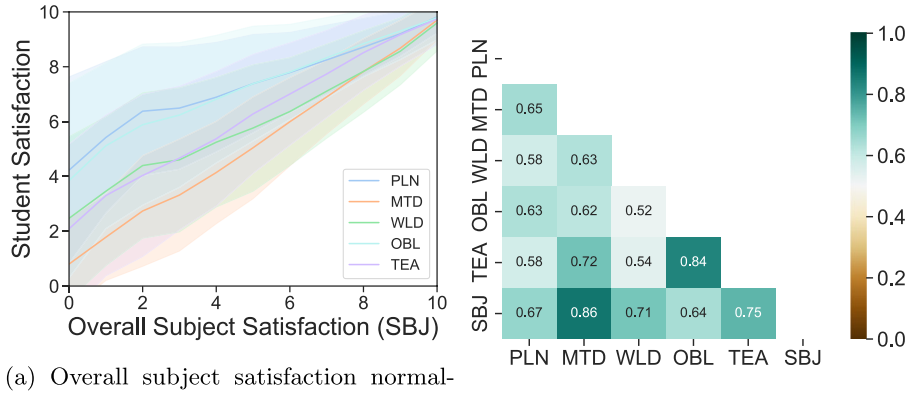
4.1.2 Student performance

In addition to student satisfaction scores, we also have access to aggregate statistics of students' performance by group.

These aggregate include the total number of total students and the number of students graded as failed (grade < 5.0), approved ($5.0 \leq$ grade < 7.0), notable ($7.0 \leq$ grade < 9.0) and excellent (grade ≥ 9.0).

We remark that we do not have access to individual grades, and hence, we cannot compute an average grade for each group.

However, we can roughly estimate this average by using the middle point of each range of grades. This way, the average performance of students in a group is estimated as $(\text{P(failed)} \times 2.45 + \text{P(approved)} \times 5.95 + \text{P(notable)} \times 8 + \text{P(excellent)} \times 9.5)/4$.



(a) Overall subject satisfaction normalized as a function of the ratings of the survey questions.

(b) Pearson correlation heatmap for student satisfaction survey

Fig. 1 Overall subject satisfaction as a function of the student satisfaction scale (**a**) and Correlations among survey questions (**b**)

5 Results

5.1 The interplay between student performance and satisfaction measurements with learning design

5.1.1 Correlation between answers to questions on student satisfaction

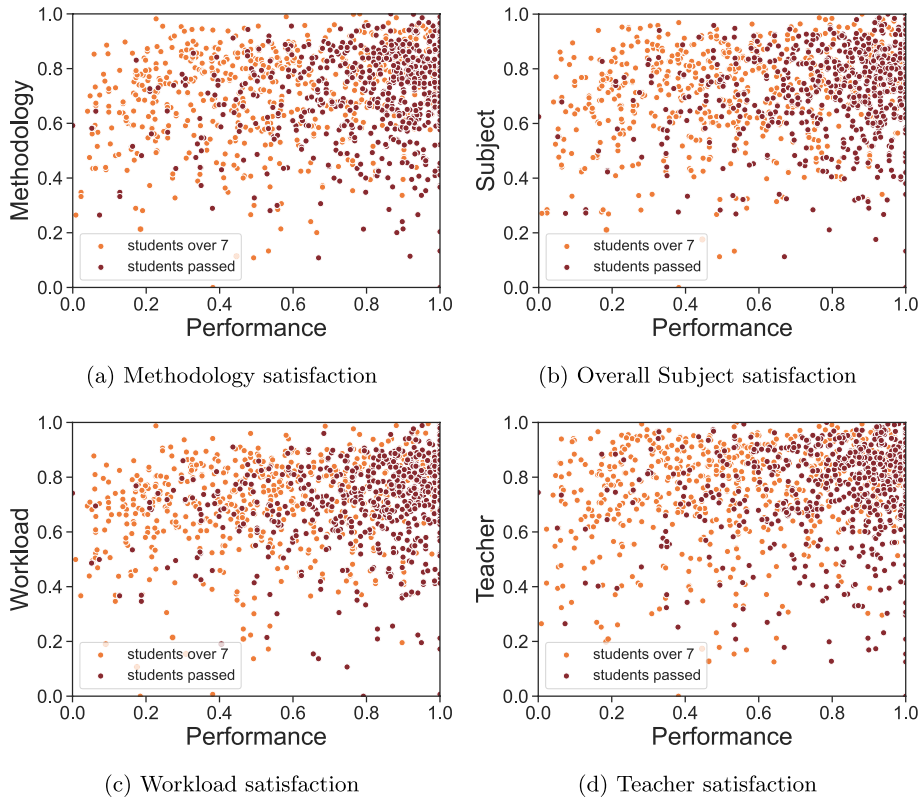
Overall satisfaction with the subject (SBJ) generally correlates with all the other questions from the satisfaction survey. This can be seen in Fig. 1a. We also consider the Pearson correlation between answers to different parts of the satisfaction survey to answer the RQ1 (Fig. 1b).

We observe that teaching methodologies (MTD) and overall satisfaction with the teacher (TEA) are highly correlated with overall satisfaction with the subject (SBJ). From that, we can see that the satisfaction with the learning design facets related to teaching methodologies has a higher correlation with the overall satisfaction perceived. While the satisfaction with the compliance with the teaching plan (PLN) and the perceived workload (WLD) resulting from the learning design proposed seems to have a relatively lower correlation and influence on the overall course satisfaction, we found a relatively lower correlation. The correlation between satisfaction with workload (WLD) and satisfaction with teaching received (TEA), as well as compliance with teaching obligations (OBL), is weak. Similarly, the correlation between satisfaction with teaching plans (PLN) and satisfaction with teaching received (TEA) is also low. Compliance with teaching obligations (OBL) and overall satisfaction with the teacher (TEA) are highly correlated. Additionally, we performed the t-test of the significance of the correlation coefficient and found $p < 0.001$ for all correlations, meaning that they are statistically significant.

5.1.2 Correlation between student satisfaction and student performance

Each group corresponds to two dots in Fig. 2: the proportion of students graded 5 or more, which is the proportion of students passing a course, and the proportion of students graded 7-10, which are those graded “notable” and “excellent”.

In both cases, student satisfaction scores are distributed over a broad range.



(a) Methodology satisfaction (b) Overall Subject satisfaction

(c) Workload satisfaction (d) Teacher satisfaction

Fig. 2 Correlation between student performance and student satisfaction measures. Each subject is represented by two dots: one indicating students who get a "notable" or "excellent" grade (lighter), and one indicating the students who pass the subject

We can see that the top-left quadrant is sparsely populated, while the top-right quadrant is denser. This means that the highest student satisfaction is not obtained by courses in which student performance is abnormally low, i.e., courses in which less than 20% of students pass or, conversely, in which 80% or more of the students fail. Yet, most of the courses with high student satisfaction are in courses that a large majority of students pass or in which a large majority of students pass with a high grade.

Statistical analyses indicated weak correlations (ranging from 0.16 to 0.25) were statistically significant (p-values < 0.0001). Additionally, several courses deviated from the overall trend.

5.1.3 The interplay between context, student performance, and satisfaction measurements

To answer the research question RQ3 - Does the context influence student satisfaction with the learning design? In Fig. 3, we have the correlations based on the factors being analyzed considering the dataset.

We observe that class size negatively correlates with student satisfaction with methodology, workload, and overall satisfaction with the subject. This also can be

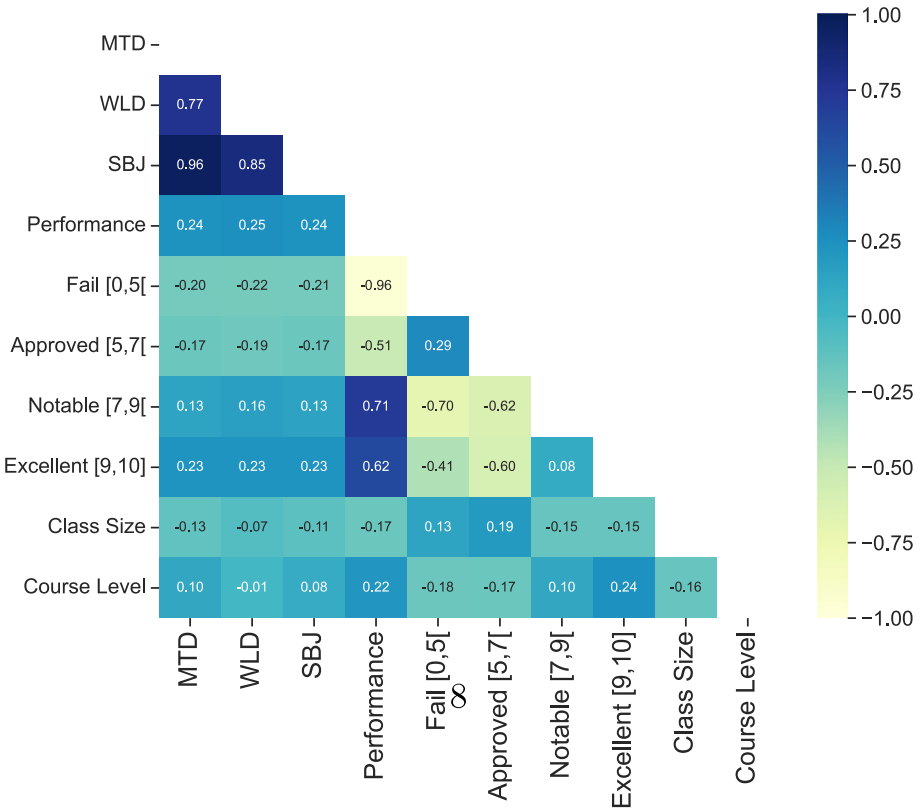


Fig. 3 Correlations with learning design satisfaction, students performances, and educational setting

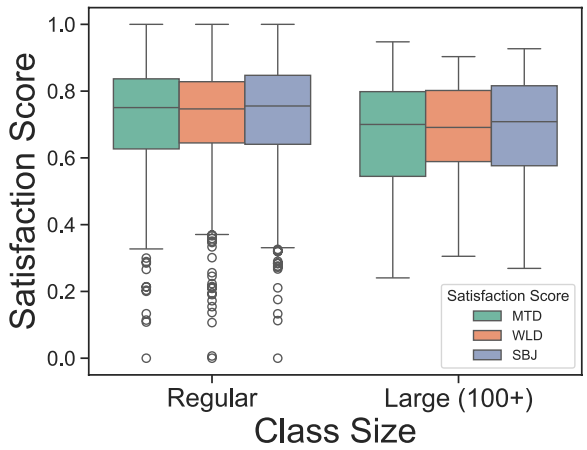


Fig. 4 Student satisfaction score by class size

noted in Fig. 4, where larger classes, having 100 students or more, tend to be penalized regarding satisfaction scores.

On the other hand, course level shows a weak correlation with overall satisfaction with the subject.

The correlation of student satisfaction with performance is weak: 0.24 with methodology, 0.25 with the workload, and 0.24 with overall satisfaction with a subject.

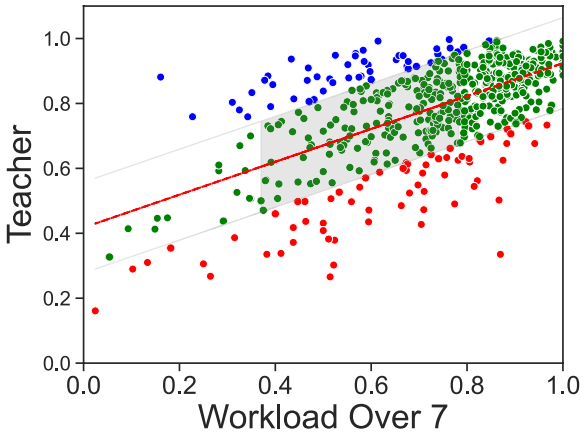
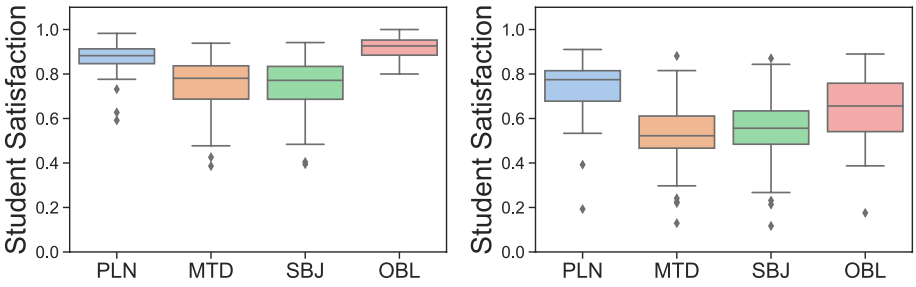


Fig. 5 Correlation between teacher satisfaction and the proportion of students who considered the workload deserved an evaluation of 7 or higher



(a) Student satisfaction above range (b) Student satisfaction below range

Fig. 6 Evaluations by course level, considering the data points above range and below the range that goes beyond the trend, left and right, respectively

5.1.4 The interplay between workload and satisfaction measurements

In Fig. 5, we computed workload evaluations above 7 by course. We use workload over 7 because it is considered a "good" rating.

We observe that there is a correlation between satisfaction with the teaching received and workload. The higher the satisfaction with the workload, the better the satisfaction with the teaching received.

We segmented the Fig. 5 in ranges to help a better understanding of the correlations within the dataset studying the workload. Compared to methodology and satisfaction with the subject, the workload has more data points in the cases that go beyond the trend when evaluating student's evaluations over 7. To examine cases that deviate from the trend, we investigated in Fig. 6 20% outliers data points. To contextualize, above the range, there are 49 data points, and below the range, there are 66 data points.

In Fig. 6a, we observe that when the satisfaction with teaching received (TEA) is higher than what would be expected considering the fraction of students that are not satisfied with the workload (WLD), the satisfaction with compliance with teaching obligations (OBL) and compliance with teaching plan (PLN) is fairly high and has low variance. This suggests that students tend to be more "forgiving" of inappropriate workloads (mostly higher than expected workloads, we suppose) if the professor honors the plan and fulfills

well his/her obligations. On the other hand, Fig. 6b shows that when the satisfaction with the teaching received is lower than expected considering workload (WLD), satisfaction with multiple aspects is lower on average. Interestingly, the biggest difference in the ratings when comparing both Fig. 6 is observed in the satisfaction with the methodology (MTD), which is highly correlated with the satisfaction with the overall subject (SBJ), followed by compliance with teaching obligations.

6 Discussion

Both university faculty perceptions and research studies present evidence regarding bias in course evaluations based on student satisfaction that is not entirely consistent with each other. In this study, we analyze the perspective of institutional analytics with measures related to performance and its effects on specific facets of student satisfaction related to three facets of the learning design of the courses.

We investigate a dataset (with reliable data, RAS index) containing student satisfaction, educational setting dimensions, and student performance. The results of this study align with previous researcher findings (Centra, 2003; Li et al., 2016; Rajabalee & Santally, 2021; Rienties & Toetenel, 2016) when we observe that satisfaction with the teaching methodology and workload are highly correlated with the overall subject satisfaction (RQ1). This consistency is more pronounced when we evaluate the interactions of MTD, WLD, and SBJ, where we highlight the impact of the methodology on the overall perception of satisfaction with the subject.

We found a weak correlation regarding the relationship between student satisfaction measures and student performances (RQ2) as the results are sparse and distributed among the ratings. As in Centra (2003), teachers giving higher grades to their students do not impact the student's perceptions of satisfaction. Although in Badri et al. (2006); Garrouste and Le Saout (2020), the authors found effects of performance measures in the student's ratings, our results evidence that in our context, the hypothesis about professors giving good grades leads to greater student satisfaction seems to have a lower or minimal effect in the results. Further, we investigate other factors that could influence student satisfaction, relating also to the student's performance (RQ3). We note that class size and course level present a lower correlation with the learning design measures. The performance also is not highly correlated with the MTD, WLD, and SBJ, showing consistency in the dimensions evaluated with our previous analysis.

To investigate the cases that go beyond the trends, we illustrated the cases related to workload and teacher satisfaction. We found differences in the two groups analyzed, indicating the impact on satisfaction measures depending on the workload and its relation to satisfaction with the teaching received. Especially when we observe the satisfaction with the methodology, which demonstrates the biggest difference between the two groups analyzed (above and below the range).

This study supports the importance of institutional analytics and generative uncertainty in considering different perspectives impacting student satisfaction when addressing decision-making in higher education. In contrast to a belief spread among university instructors and contradictory results in previous research regarding our main research question, "Do grades measuring student performance bias student satisfaction ratings with the learning design?" we can see that satisfaction with learning

design facets is mostly distributed when controlling student performance. The distribution includes higher and lower scores, suggesting that grades are not always biasing students' satisfaction ratings with both learning design facets and course satisfaction. There is also the common belief about workload. We see that courses with good methodological design can have good student satisfaction ratings beyond bias or dependencies with grades and satisfaction with the workload.

7 Conclusions

Student satisfaction results are insightful and widely considered in program quality assessment and academic promotion processes. Although the measure promotes and improves decision-making in higher education, it might contain biases, where learning about the constructs and possible biases becomes necessary to provide fair use of the student satisfaction results. Institutional analytics, such as the approach contributed by this study, can help to understand these constructs.

Our research shows that student satisfaction results need to be interpreted holistically, from a generative uncertainty stance, considering potential bias but also beyond bias. We found a strong correlation between student satisfaction with teaching methodology and workload and their overall perception of the subject. This underscores the significant influence of students' views on learning design decisions in shaping their overall satisfaction. Conversely, grades displayed a weak correlation with the learning design measures in the survey. There are cases of high satisfaction with the course and low satisfaction with the workload when other factors (and especially methodology) are highly rated. This suggests that, in this context, professors who give higher grades aren't necessarily guaranteed better student evaluations. We noted then that performance had little influence on overall subject satisfaction, as evidenced by their somehow low correlations.

Overall, the measures of students in higher education and student satisfaction measures hold significant importance. Analyzing survey results with an awareness of context is essential, especially recognizing the impact of learning design facets on satisfaction and performance. Such a comprehensive approach mitigates potential biases in higher education decisions. It also promotes awareness among professors and academic managers about the importance of decisions in the learning design, transcending other (sometimes limited) biases or commonly believed dependencies to improve the quality of the courses. Institutional analytics dashboards should care about the "bias and beyond bias" uncertainties in their institutions to help them understand the effects of factors in their contexts and identify elements for further exploration and teaching improvement. While this study provides valuable insights, to draw further conclusions, additional features (not present in our dataset) are needed, as well as data from more academic years, to allow more robust analyses. We plan further research considering the different teaching methodologies of the courses, and the interactions of students with the Learning Management System (LMS). Additionally, one could envision qualitative studies considering the comments that some students leave in the evaluations, which is data that we do not have at the moment.

7.1 Limitations

The study analyzed only two academic years and one university. This way, the results are tied to the specific university context and do not allow a broader view to generalize the findings. Yet, it illustrates the need of supporting generative uncertainty in institutional analytics. Also, we cannot access individual scores, which prevents us from having an average grade for each course. The following studies can address these limitations and promote insights into other contexts.

Acknowledgements
The work was done in collaboration with UPF services (Strategy, Data and Quality Service and the Data Protection Officer) through the UPF Learning Lab initiative.

Authors' contributions
Francielle Marques: Conceptualization, Investigation, Data curation, Formal analysis, writing—original draft, Writing—review & editing; Davinia Hernández-Leo: Funding Acquisition, Coordination, Conceptualization, Investigation, Supervision, Writing – review & editing; Carlos Castillo: Funding Acquisition, Coordination, Conceptualization, Investigation, Supervision, Writing – review & editing;

Funding
This work is co-funded by MICIU/AEI/https://doi.org/10.13039/501100011033(PID2020-112584RB-C33, PID2023-146692OB-C33, CEX2021-001195-M) and the Department of Research and Universities of the Government of Catalonia (SGR 00930).
Hernández-Leo (Serra Húnter) also acknowledges the support by ICREA under the ICREA Academia programme.
Castillo also acknowledges support from the EU-funded projects "SoBigData++" (grant agreement 871,042), "FINDHR" (grant agreement 101,070,212).

Data availability
Not applicable.

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests. Not applicable.

Received: 18 June 2024 Accepted: 19 January 2025
Published online: 11 March 2025

References

Badri, M. A., Abdulla, M., Kamali, M. A., & Dodeen, H. (2006). Identifying potential biasing variables in student evaluation of teaching in a newly accredited business program in the uae. *International Journal of Educational Management, 20*(1), 43–59.
Bedggood, R. E., & Donovan, J. D. (2012). University performance evaluations: What are we really measuring? *Studies in Higher Education, 37*(7), 825–842.
Burns, S., & Corwin, K. (2017). Automating student survey reports in online education for faculty and instructional designers. *Proceedings of the seventh international learning analytics & knowledge conference* (pp. 590–591).
Campaña, J. C., Gimenez-Nadal, J. I., & Molina, J. A. (2016). The satisfaction of university students in Spain: differences by field of study. *Applied Economics Letters, 23*(7), 506–509.
Campos, F., Nguyen, H., Ahn, J., & Jackson, K. (2024). Leveraging cultural forms in human-centred learning analytics design. *British Journal of Educational Technology, 55*(3), 769–784.
Centra, J. A. (2003). Will teachers receive higher student evaluations by giving higher grades and less course work? *Research in Higher Education, 44*, 495–518.
Clayson, D. E. (2009). Student evaluations of teaching: Are they related to what students learn?: A meta-analysis and review of the literature. *Journal of Marketing Education, 31*(1), 16–30.
Flegl, M., & Andrade Rosas, L. A. (2019). Do professor's age and gender matter or do students give higher value to professors' experience? *Quality Assurance in Education, 27*(4), 511–532.
Garrouste, M., & Le Saout, R. (2020). Good teaching and good grades. Can you buy pedagogy? *Annals of Economics and Statistics* (139), 29–60. https://www.jstor.org/stable/10.15609/annaeconstat2009.139.0029.
Griffin, B. W. (2004). Grading leniency, grade discrepancy, and student ratings of instruction. *Contemporary Educational Psychology, 29*(4), 410–425.

Heffernan, T. (2023). Abusive comments in student evaluations of courses and teaching: The attacks women and marginalised academics endure. *Higher Education, 85*(1), 225–239.
Hernández-Leo, D., Martinez-Maldonado, R., Pardo, A., Munñoz-Cristóbal, J. A., & Rodríguez-Triana, M. J. (2019). Analytics for learning design: A layered framework and tools. *British Journal of Educational Technology, 50*(1), 139–152.
Laurillard, D. (2013). *Teaching as a design science: Building pedagogical patterns for learning and technology*. Routledge.
Li, N., Marsh, V., & Rienties, B. (2016). Modelling and managing learner satisfaction: Use of learner feedback to enhance blended and online learning experience. *Decision Sciences Journal of Innovative Education, 14*(2), 216–242.
Marques, F., Hernández-Leo, D., & Castillo, C. (2024). Measuring gender bias in student satisfaction in higher education: a cross-department study. *Cogent Education*, *11*(1), 2375183.
Mengel, F., Sauermann, J., & Zölitz, U. (2018). Gender bias in teaching evaluations. *Journal of the European Economic Association, 17*(2), 535–566.
Ortiz-Beltrán, A., Hernández-Leo, D., Amarasinghe, I. (2022). Surviving and thriving: How changes in teaching modalities influenced student satisfaction before, during and after covid-19. *Australasian Journal of Educational Technology*, 72–88. https://ajet.org.au/index.php/AJET/article/view/8958.
Radchenko, N. (2020). Biases in student evaluations of teaching: An american case study.
Radmacher, S. A., & Martin, D. J. (2001). Identifying significant predictors of student evaluations of faculty through hierarchical regression analysis. *The Journal of Psychology, 135*(3), 259–268.
Rajabalee, Y. B., & Santally, M. I. (2021). Learner satisfaction, engagement and performances in an online module: Implications for institutional e-learning policy. *Education and Information Technologies, 26*(3), 2623–2656.
Remedios, R., & Lieberman, D. A. (2008). I liked your course because you taught me well: The influence of grades, workload, expectations and goals on students' evaluations of teaching. *British Educational Research Journal, 34*(1), 91–115.
Rienties, B., Toetenel, L., Bryan, A. (2015). " scaling up" learning design: impact of learning design activities on lms behavior and performance. *Proceedings of the fifth international conference on learning analytics and knowledge* (pp. 315–319).
Rienties, B., & Toetenel, L. (2016). The impact of learning design on student behaviour, satisfaction and performance: A cross-institutional comparison across 151 modules. *Computers in Human Behavior, 60*, 333–341.
Rosen, A. S. (2018). Correlations, trends and potential biases among publicly accessible web-based student evaluations of teaching: a large-scale study of ratemyprofessors.com data. *Assessment & Evaluation in Higher Education, 43*(1), 31–44.
Scheaffer, R.L., Mendenhall, W., Ott, R.L., Gerow, K.G. (2011). *Elementary survey sampling (6th ed)*. Thomsom Brooks/Cole.
Stumpf, S. A., & Freedman, R. D. (1979). Expected grade covariation with student ratings of instruction: Individual versus class effects. *Journal of Educational Psychology, 71*(3), 293.
Torrado, M., & Blanca, M. J. (2022). Assessing satisfaction with online courses: Spanish version of the learner satisfaction survey. *Frontiers in Psychology, 13*, 875929.
Wise, A.F., Brackett, A., Maddox, B. (2023). Flexible coupling of learning analytics research and practice in the university: A collective strengths approach. *Lak23 conference proceedings* (p. 25).

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.