



Students' preferences with university teaching practices: analysis of testimonials with artificial intelligence

Carmen Álvarez-Álvarez¹ · Samuel Falcon²

Accepted: 29 April 2023 / Published online: 16 May 2023
© The Author(s) 2023

Abstract

University teaching practices impact student interest, engagement, and academic performance. This paper presents a study that uses artificial intelligence (AI) to examine students' preferences for university teaching practices. We asked students in various fields open-ended questions about the best teaching practices they had experienced. Due to the large amount of data obtained, we used the AI-based language model Generative Pretrained Transformer-3 (GPT-3) to analyse the responses. With this model, we sorted students' testimonials into nine theory-based categories regarding teaching practices. After analysing the reliability of the classifications conducted by GPT-3, we found that the agreement between humans was similar to that observed between humans and the AI model, which supported its reliability. Regarding students' preferences for teaching practices, the results showed that students prefer practices that focus on (1) clarity and (2) interaction and relationships. These results enable the use of AI-based tools that facilitate the analysis of large amounts of information collected through open methods. At the didactic level, students' preferences and demand for clear teaching practices (in which ideas and activities are stated and shown without ambiguity) that are based on interaction and relationships (between teachers and students and among students themselves) are demonstrable.

Keywords Teaching practices · Teaching quality · Satisfaction · Higher education · Artificial intelligence

✉ Samuel Falcon
samuel.falcon@ulpgc.es

Carmen Álvarez-Álvarez
alvarezmc@unican.es

¹ Department of Education, Universidad de Cantabria, Avda. de los Castros S/N, 39005 Santander, Spain

² Department of Education, University of Las Palmas de Gran Canaria, Calle Juan de Quesada, n° 30, 35001 Las Palmas de Gran Canaria, Las Palmas, Spain

Introduction

University teaching practices are a major area of interest for educational researchers (Harbour et al., 2015; Slavin & Lake, 2008). Teaching practices play a role in stimulating students' interest, engagement, learning, and academic performance (Vercellotti, 2018). A paradigm shift in university teaching is currently taking place; expository teaching practices are being questioned and gradually replaced by active methodologies, professional simulation practices and interactive practices, among others (Carr et al., 2015; Roberts, 2019). However, it is necessary to understand students' preferences for teaching practices because they impact the emotional engagement and performance of students (Smith & Baik, 2021).

A suitable way to assess students' preferences for teaching practices is through open-ended questions (Hills et al., 2016). However, the data coding used in this method prevents working with a large number of samples and requires considerable processing time (Rahman, 2016). Currently, these problems can be overcome due to advances in artificial intelligence (AI), as they facilitate and optimise the performance of these tasks (Hirschberg & Manning, 2015). Employing a language processing model makes it possible to accurately and efficiently analyse large amounts of text. This, in turn, allows researchers to gain a deeper understanding of the topic of study and to thus draw more meaningful conclusions (Johnson & Onwuegbuzie, 2004).

Therefore, the aims of this paper are (1) to assess university students' preferences for teaching practices through open-ended questions and (2) to encode the information using an AI-based tool. In this way, we can assess whether the tool is sufficiently reliable to analyse data collected through open-ended questions. In addition, this method enables us to identify which teaching practices are preferred by university students, which could help researchers and teachers account for these aspects when designing teaching and learning programmes.

The introduction is divided into two parts. The first part elaborates on the importance of learner preferences, while the second provides a more detailed description of the AI-based tools capable of analysing large amounts of text.

Students' preferences for teaching practices

Previous studies have addressed the need for academic communication that motivates emotional engagement on the part of university students through the teaching practices employed by their teachers (Chalmers et al., 2018; Könings et al., 2011; Tronconi et al., 2021). Several studies advocate reducing the use of lectures for large groups and employing active methodologies with regular feedback for students (Carr et al., 2015; Chalmers et al., 2018; Hardman, 2016; Moliní Fernández & Sánchez-González, 2019; Roberts, 2019; Steen-Utheim & Wittek, 2017). To date, however, there has been little discussion about students' preferences within these methodologies. A study in which the way that students experience their university classes, how they value their active learning experiences and what preferences they have in this respect is assessed is needed to maximise their emotional engagement with the subject matter and ensure their success (Alegre & Villar, 2017; Könings et al., 2011; Slater & Davies, 2020).

Previous studies on students' preferences have mainly focused on specific areas of knowledge or specific teaching practices. For instance, Minhas et al. (2012) assessed the preferences of health students for teaching practices and found a preference for

seminar-based learning over lectures. Another study conducted by Opdecam et al. (2014) assessed the preferences of first-year university students for teamwork-based activities over lectures. Their results showed that the students' preferences clearly differed depending on gender (women preferred teamwork more than men did), level (students with a lower profile preferred teamwork) and motivation (teamwork had a high acceptance among the most intrinsically motivated students). Nevertheless, a global understanding of preferred teaching practices in general is lacking.

In regard to the study of students' preferences, we encounter a major problem: the existence of different names for similar constructs and similar names for different constructs. This problem is known as the jingle-jangle fallacy (Marsh, 1994; Marsh et al., 2003) and leads to confusion, information elusiveness and misinterpretation. To avoid these fallacies, in this study, we build on the categories identified in a recent systematic review of teaching practices in universities (Smith & Baik, 2021). The nine categories of teaching practices applied in this research are as follows:

1. **Clarity:** Teaching practices in which the structure and content of knowledge are clear to students. These practices require planning, organisation and structure in the content and delivery of lectures and practical classes.
2. **Research:** The use of methodological approaches that encourage problem solving, enquiry and testing, such as problem- or case-based learning.
3. **Application:** The conducting of exercises and activities that require the use of knowledge gained through active learning in different situations or contexts.
4. **Experiential:** A particular type of application in which practical and experiential, authentic or real learning is developed through the learner's own experience (e.g., work placements in other institutions).
5. **Challenges:** Practices meant to achieve interest and deep cognitive engagement through didactic proposals that challenge students' thinking, expression, or action.
6. **Relevance:** A set of teaching practices that highlight the value, purpose or impact of the interventions to be addressed in a learning or professional development process.
7. **Interaction and relationships:** A set of communicative and relational processes between teachers and students and among the students themselves (collaborative learning, peer tutoring, classroom dialogue, etc.).
8. **Consolidation:** A set of correction, recovery and revision practices meant to help identify errors and other comprehension problems to address them correctly and adequately.
9. **Self-regulation:** Self-assessment and self-monitoring practices conducted independently by the students themselves to plan, organise and correct their own comprehension errors, which leads to the achievement of cognitive training and a greater awareness of progress.

These practices capture Smith and Baik's (2021) findings at a general level of abstraction. By employing these categories, we can classify the information from student responses to open-ended questions at a level that can be easily understood by both teachers and students. In this way, it may be easier to integrate students' preferred practices into teachers' professional performance.

Text analysis with AI

We can take different approaches to answer the question of which teaching practices are preferred by university students. For instance, some studies attempt to answer this question

through the use of self-report questionnaires (Aridah et al., 2017), while others do so through open collection methods (Hills et al., 2016). Consequently, the method of data analysis differs. In the first case, the analysis is quantitative and usually performed by using statistical techniques, while in the second case, the analysis is qualitative and conducted through content analysis of emerging categories (Cohen et al., 2000; Lodico et al., 2010).

Between these two approaches, the use of open collection methods such as open-ended questions or interviews facilitates a better expression of ideas among students, allowing researchers to gain a deeper understanding of the relevant issue (Johnson & Onwuegbuzie, 2004). However, studies following this methodology are often limited in sample size or data processing time due to the large amounts of collected information that need to be coded and analysed (Rahman, 2016).

A review of the various studies related to the nine categories of good practice, as identified by Smith and Baik (2021), shows that, thus far, most studies designed to collect large amounts of information from a large number of samples used questionnaires or standardised scales as data collection techniques for quantitative analysis (Könings et al., 2011; Minhas et al., 2012; Opdecam et al., 2014; Vercellotti, 2018). In contrast, most other studies designed to collect information from a small number of samples or more easily manageable amounts of data used qualitative analysis (Steen-Utheim & Wittek, 2017) and descriptive statistics (Hardman, 2016). However, another form of data analysis is emerging thanks to AI.

Advances in AI over the last decade have made it easier to solve problems involving the processing of large amounts of data across all fields of knowledge. This is especially notable in biology, where the development of AlphaFold 2, an AI-based tool capable of predicting the three-dimensional structure of proteins, has been a milestone (Callaway, 2020; Jumper et al., 2021). Since researchers have gained access to this tool, the number of preprints and scientific publications in the field has increased significantly, as has our knowledge (Callaway, 2022). Similarly, great achievements are being made in fields related to textual analysis due to advances in natural language processing (Hirschberg & Manning, 2015).

The use of AI-based tools for textual analysis in education is a practice that has yielded positive results for a decade. For instance, consider the case of sentiment analysis, which involves a tool capable of extracting sentiment (positive, negative, or neutral) from large amounts of text (Rani & Kumar, 2017). Over the last few years, researchers have used this tool to study students' evaluations of massive open online courses (MOOCs) and teachers through the use of open-ended questions (Geng et al., 2020; Rybinski & Kopciuszewska, 2021; Zhou et al., 2020). Some of this research has concluded that this tool is useful for course satisfaction evaluations, (Cunningham-Nelson et al., 2019), teaching analysis (Leong et al., 2012) and course improvement (Pong-inwong & Songpan, 2019). This might be an effective way to assess students' preferences for teaching practices, but thanks to the transformer revolution (Vaswani et al., 2017), a door to the thorough analysis of responses to open-ended questions has been opened.

Transformers allowed language processing models to shift from being dependent on human training (in a supervised way) to being trained automatically (or self-supervised) through large corpora of textual data. This paradigm shift has resulted in large pretrained models with millions of parameters that are able to understand human language much better than their predecessors (Qiu et al., 2020). In this context, models that use deep learning to understand and generate high-quality text, such as the Generative Pre-trained Transformer-3 (GPT-3), have emerged (Floridi & Chiriatti, 2020). GPT-3 has two strengths: (1) its ability to understand written instructions in natural language (as one person would

speak to another); and (2) its flexibility, since, having been trained only to understand and generate human-like text, it can perform many tasks for which it has not been specifically trained. These tasks include classification, sentiment analysis, programming, and textual summarisation, among many others (OpenAI, 2022). Despite the novelty of these models, several authors have already expressed a need to use them in the execution of tasks such as the coding of large amounts of information to then study their reliability (Qiu et al., 2020). Moreover, by using these novel models, many of the problems associated with the coding and analysis of information collected through open methods can be overcome.

Objectives of the current study

In this study, using AI (GPT-3), we analysed students' answers to an open-ended question regarding their preferred teaching practices. Furthermore, we identified the university teaching practices that most satisfy students. The findings will determine whether the reliability results are good enough for the use of this tool in analysing open-ended student responses. This could open a door for the use of AI-based tools in the analysis of large amounts of qualitative data. Moreover, by using this method, we can discover students' preferences for university teaching practices and thus better guide the processes of methodological change.

Methods

Participants

Participants were undergraduate and postgraduate students from 90 classes (42 in the first term and 48 in the second term) at the University of Cantabria, Spain representing different disciplines. The total number of respondents was 1081 (601 women and 480 men).

Procedure

We informed both teachers and students of the study objectives, and then we visited each class so that students could complete the questionnaires. Students completed the surveys in the classroom under the supervision of the teacher and researchers. These surveys consisted of several scales (Álvarez-Álvarez et al., 2022), but only the open-ended question on best teaching practices was considered in this study. The data were treated ethically and in accordance with the guidelines of academic university research, which stipulate confidentiality and objectivity.

Instruments

Teaching practices

Following previous studies in which specific open-ended questions are asked and the answers are then analysed using AI (Hynninen et al., 2019), we assessed best teaching practices from a student's point of view by reviewing responses to the following open-ended question: "*Comment and explain in your own words the best practice you have seen*

in this class and explain why you think it is successful in as much detail as you can so that other teachers can imitate it”.

To code the information collected through the use of the open-ended question, we used a classification system for teaching practices developed by Smith and Baik (2021) in their systematic review. This system consists of 9 categories of teaching practices, to which we added the category “0”, referred to as “none”, to classify student responses stating that there is no good teaching practice. The resulting rubric is detailed in Table 1.

GPT-3

We used the GPT-3 model (Brown et al., 2020) to code the responses to our open-ended question. Specifically, we used text-davinci-002, with the temperature set to 0.1 and the Top P set to 1. The instructions for the model included the sentence “*Classify the comments in one of the following categories:*”, followed by the categories defined in the rubric (Table 1). Afterwards, we provided the answers to the open-ended question for classification.

Data analysis

To calculate the reliability of the coding conducted with GPT-3, two researchers independently coded a random selection equal to 10% of the total sample following the procedure conducted by other researchers to assess the reliability of coding (Russ, 2018). Both GPT-3 and our coders classified each response into a single category of teaching practices. Reliability was calculated as the percentage of agreement (Brownell et al., 2013; King & La Paro, 2015) using the ReCal3 tool (Freelon, 2010). After coding students’ responses with GPT-3, we conducted a descriptive analysis of the results using JASP 0.16.2 (JASP Team, 2022).

Results

Coding reliability

The overall percentage of agreement between the two researchers and GPT-3 was 89.07%, which is considered satisfactory (O’Connor & Joffe, 2020). Among the various categories, this overall agreement percentage varied individually from 64.81% for the category “*Interaction and relationships*” to 97.53% for the category “*Importance*” (Table 2).

Descriptive analysis

The frequency of each category used to classify responses according to the rubric is presented below (Table 3). In addition, one representative example from each category is presented.

We can see that there are two categories of teaching practices that stand out, “*Interaction and relationships*” and “*Clarity*”, in which 30.53% and 28.86% of the responses were classified, respectively. The next highest ranking categories were “*Application*”, with 13.14% of responses classified, “*Investigation*”, with 10.64%, “*Experience*”, with

Table 1 Rubric used to classify teaching practices. Adapted from Smith and Baik (2021)

Code	Name	Characteristics	Definition
1	Clarity	Structure of the content representations, Alignment, Experience, Relationship	Make the structure of knowledge and the progression of learning clear to students. There are three levels: (a) Curriculum design: clear objectives and alignment between objectives, activities and assessments; clear organisation of topics within a subject ("disciplinary content structure") (b) Lesson design: clear planning and organisation of content and activities (c) Delivery: clear explanations and structuring of content by experts
2	Investigation	PBL (problem-based learning), CBL (case-based learning), active learning, inquiry-based learning	Use of approaches/methods which aim to encourage questioning, problem solving, investigation and testing. Sometimes referred to as 'inquiry-based learning' or 'active learning'. Examples of common pedagogical approaches are problem-based or case-based learning
3	Application	Active learning, application of knowledge, flipped classroom, PBL (problem-based learning), CBL (case-based learning)	Engage learners in exercises/activities to apply knowledge and understanding
4	Experience	Active learning, PBL (problem-based learning), episodic richness	A particular type of application involving practical and experiential learning, sometimes referred to as "authentic learning" or "real world" practice
5	Challenges	Stimulating interest, inquiry-based learning	Stimulate interest and foster deep cognitive engagement. Sometimes mentioned in connection with "inquiry-based learning" and problem-based learning
6	Importance	Value (for learners), Episodic richness, Stimulate interest	Helping students to see the value/purpose of what they are learning. There are two levels: (1) Pedagogical approaches: experiential learning, problem- or case-based learning. (2) The way the teacher teaches: e.g., using authentic examples of disciplinary ideas or constructs for students
7	Interaction and relationships	Collaborative learning, Interaction/dialogue, Student-teacher relationship, Collaborative assessment, Peer tutoring	Enable and facilitate peer interaction and learning in a social context; Encourage positive interaction between pupils and teachers

Table 1 (continued)

Code	Name	Characteristics	Definition
8	Consolidation	Random practice, Examination practice, Remedial practice, Structure of content representations; relationships between ideas	Provide appropriate types of remedial and review practice, where the material to be learned is "made up" during the study sessions following the first session in which the material is learned. Consolidate understanding and correct misconceptions
9	Self-regulation	Metacognitive training, Modelling, Awareness of learning/progress, Independent learning	Facilitating learners' self-assessment, management of their own learning (e.g. planning, organisation, monitoring, corrective action, revision), learning how to learn and reflecting on how they come to learn
0	None (added to the original categories)	The student considers that there has been no good practice	

Table 2 Percentage of agreement in coding 10% of the total sample

	Teaching practice category code										M
	0	1	2	3	4	5	6	7	8	9	
Overall agreement percentage	96.91	84.57	90.12	86.42	88.27	96.30	97.53	64.81	91.98	93.83	89.07
GPT-3 and researcher 1	95.37	80.56	86.11	82.41	85.19	94.44	96.30	62.96	92.59	92.59	86.85
GPT-3 and researcher 2	99.07	83.33	89.81	85.19	90.74	100.00	98.15	65.74	92.59	93.52	89.81
Researcher 1 and researcher 2	96.30	89.81	94.44	91.67	88.89	94.44	98.15	65.74	90.74	95.37	90.56

M = Mean. The codes correspond to the following categories: 0 = None; 1 = Clarity; 2 = Investigation; 3 = Application; 4 = Experience; 5 = Challenges; 6 = Importance; 7 = Interaction and relationships; 8 = Consolidation; 9 = Self-regulation

Table 3 Frequency of each category of teaching practice

Category of teaching practice	Frequency	Percentage	Representative example
None	25	2.31	Nothing
Clarity	312	28.86	The slides and the teacher's good method of explanation
Investigation	115	10.64	The clinical cases help to better understand the subject matter
Application	142	13.14	Use of apps or websites where to put into practice concepts of this subject
Experience	77	7.12	The best thing is the lab practicals because that is where I really see what I have been taught in class
Challenges	10	0.93	He asks broad questions that are difficult to answer to make us think things through
Importance	7	0.65	The teacher has included a variety of examples so that we can better internalise the content of the subject. She gives real examples that have happened or could happen, so that we can think about how we would react to different experiences
Interaction and relationships	330	30.53	Classroom debate, in which we all participate
Consolidation	26	2.40	The teacher answers all our questions and gives us a daily review of the previous lessons
Self-regulation	37	3.42	Following our contributions to an activity, the teacher intervenes and helps us to work on and understand them more effectively
Total	1081	100.00	

7.12%, “*Self-regulation*”, with 3.42%, “*Consolidation*”, with 2.40% and “*None*”, with 2.31%. Finally, the categories “*Challenges*” and “*Importance*” were almost residual, with 0.93% and 0.65% of responses classified, respectively.

Discussion

The present study aimed (1) to analyse students' answers to an open-ended question on their preferred teaching practices using AI (GPT-3); and (2) to identify the university teaching practices that most satisfy students to better understand their preferences. The results showed that GPT-3 was able to classify responses to the open-ended question with a reliability remarkably similar to that of humans. They also showed that university students prefer practices that focus on clarity and those that focus on interaction and relationships. These findings are discussed below.

First, it is necessary to comment on the reliability of the coding performed by the AI-based model. As claimed by Qiu et al. (2020), one of the remaining challenges following the recent emergence of large pretrained language models is determining how to use them to code large amounts of information and then using that coding information to study their reliability. Surprisingly, the percentage of agreement between humans was remarkably similar to that between humans and AI, even in the category with the lowest percentage of agreement. The average percentage of agreement between researchers 1 and 2 was 90.56%, which was not far from that between researcher 2 and GPT-3 (89.81%) or between researcher 1 and GPT-3 (86.85%). These findings help overcome the challenge proposed by Qiu et al. (2020) by demonstrating GPT-3's usefulness and reliability in coding large amounts of information. This opens the door to the use of AI-based models for coding and data analysis in other types of qualitative research. In this way, it will be possible to have a larger number of samples and shorter analysis times without losing the richness of the information obtained through open-ended collection methods, which often allow a researcher to reach more elaborate conclusions (Johnson & Onwuegbuzie, 2007).

Regarding the second objective, the results show a preference for those practices that focus on (1) clarity and (2) interaction and relationships. Students demand clear teaching practices where ideas and activities are presented and displayed unambiguously and show order, design, and planning. They also advocate the use of practices that are based on interaction and relationships (between teachers and students and among the students themselves) to share their concerns and doubts and to engender support of their learning processes in university classrooms. It is encouraging to compare these findings with those of Hattie (2008), who, after analysing more than 800 meta-analytic studies, found that effective teachers communicate clear content and assessment criteria and apply feedback both among students and between teachers and students. The results of this study also showed students' preferences for teaching practices that focus on the investigation and application of knowledge. According to previous studies, these practices are useful when teaching students (Ambrose et al., 2010). These findings have implications for university teaching practices, demonstrating student interest in clear and active methodologies (Minhas et al., 2012; Opdecam et al., 2014). University teachers who wish to meet students' preferences and achieve greater engagement in their teaching experiences need to rethink both aspects.

Limitations and future perspectives

Despite the contributions of this study, it also has some limitations that need to be addressed. According to a report carried out by UNESCO's education sector (2019), the use of AI in educational research incurs several challenges that need to be considered by researchers working with AI. Among them is the creation of inclusive models that are not biased due to the training of models with inadequate databases. In this study, this dimension is not considered, but future studies need to test for possible biases in the use of AI-based models. However, another challenge set by UNESCO is to increase the use of AI in educational research, so future research should continue to use this type of model to further explore and enable the advantages of this methodology for researchers. In addition, the results of this study were obtained using a pretrained base model. This implies that there is still room for improvement if the model were to be refined with data that had been previously classified by the researchers.

Another limitation of the current study was that only the preferences of students from Spain were assessed. Previous research has shown cultural differences in students' preferences for teaching practices (Macfayden et al., 2003; Yang & Tsai, 2008). A cross-cultural study that includes teachers from other countries is needed to assess any differences in their use of engaging messages. Similarly, the global view followed in this study prevented us from analysing the results by student attributes such as knowledge or gender, even when previous studies have shown differences in students' preferences across these variables (Opdecam et al., 2014). In future studies, it will be worthwhile to identify cohort trends in students' preferences.

Finally, the findings of this research provide insights into the design of future interventions aimed at modifying university teaching practices. A methodological change in teaching practices could lead to an improvement in student interest, engagement, and performance (Smith & Baik, 2021; Vercellotti, 2018). One possible means of achieving this is through feedback-based interventions that leverage the use of technology (Falcon et al., 2023; Rodgers et al., 2019). Students could provide feedback on their preferred teaching practices, which can be analysed instantly with GPT-3 so that a teacher can adapt to the preferences of their students. Further research should be undertaken to explore this possibility.

Conclusions

In the present study, we performed an automatic classification using an AI-based tool of student responses to open-ended questions regarding their preferences for teaching practices. We then examined the results to determine which teaching practices are preferred by university students. First, we found that the reliability of the AI model regarding the classification task was similar to that of humans. Then, the results showed that students preferred practices that focus on clarity and those that focus on interaction and relationships. These findings open the door for the use of pretrained text generation models for large textual analysis and classification tasks. In addition, they provide university teachers with guidelines for developing their teaching practices. Learning how to better plan and develop lessons has been and will be a professional challenge for university

teachers, and this study contributes to the identification and categorization of students' preferences by highlighting the importance of clarity and interaction.

Acknowledgements The authors appreciate the collaboration given in the development of the field work by the people involved.

Author contributions CÁ-Á: Conceptualization, Methodology, Investigation, Writing—Original Draft, Writing—Review & Editing, Supervision, Project administration, Funding acquisition. SF: Formal analysis, Writing—Original Draft, Writing—Review & Editing.

Funding Open Access funding provided thanks to the CRUE-CSIC agreement with Springer Nature. This work was supported by the University of Cantabria (Spain) and the funding received within its IV Call for Teaching Innovation Projects. It has also been funded by the University of Las Palmas de Gran Canaria, Cabildo de Gran Canaria, and Banco Santander by the pre-doctoral training programme for research personnel.

Data availability The data that support the findings of this study are available on request from the corresponding author. These data are not publicly available due to privacy or ethical restrictions.

Declarations

Competing interests The authors have no relevant financial or non-financial interests to disclose.

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/>.

References

- Alegre, O. M., & Villar, L. M. (2017). Indicadores y control estadístico para el seguimiento y evaluación de preferencias de aprendizaje de estudiantes universitarios. *Revista De Educación a Distancia (RED)*. <https://doi.org/10.6018/red/55/2>
- Álvarez-Álvarez, C., Sánchez-Ruiz, L., Sarabia Cobo, C., & Montoya-del Corte, J. (2022). Validación de un cuestionario para la evaluación de la interacción en la enseñanza universitaria. *REDU. Revista De Docencia Universitaria*, 20(1), 145–160. <https://doi.org/10.4995/redu.2022.15918>
- Ambrose, S. A., Bridges, M. W., Dipietro, M., Lovett, M. C., Norman, M. K., & Mayer, R. E. (2010). *7 research-based principles for smart teaching* (1st ed.). John Wiley.
- Aridah, A., Atmowardoyo, H., & Salija, K. (2017). Teacher practices and students' preferences for written corrective feedback and their implications on writing instruction. *International Journal of English Linguistics*, 7(1), 112. <https://doi.org/10.5539/ijel.v7n1p112>
- Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., Neelakantan, A., Shyam, P., Stryer, G., Askell, A., Agarwal, S., Herbert-Voss, A., Krueger, G., Henighan, T., Child, R., Ramesh, A., Ziegler, D. M., Wu, J., Winter, C., ... Amodei, D. (2020). Language models are few-shot learners. *Advances in Neural Information Processing Systems, 2020-Decem*.
- Brownell, C. A., Svetlova, M., Anderson, R., Nichols, S. R., & Drummond, J. (2013). Socialization of early prosocial behavior: parents' talk about emotions is associated with sharing and helping in toddlers. *Infancy*, 18(1), 91–119. <https://doi.org/10.1111/j.1532-7078.2012.00125.x>
- Callaway, E. (2020). "It will change everything": Deepmind's Ai makes gigantic leap in solving protein structures. *Nature*, 588, 203–204.
- Callaway, E. (2022). What's next for AlphaFold and the AI protein-folding revolution. *Nature*, 604, 234–238.

- Carr, R., Palmer, S., & Hagel, P. (2015). Active learning: The importance of developing a comprehensive measure. *Active Learning in Higher Education*, 16(3), 173–186. <https://doi.org/10.1177/1469787415589529>
- Chalmers, C., Mowat, E., & Chapman, M. (2018). Marking and providing feedback face-to-face: Staff and student perspectives. *Active Learning in Higher Education*, 19(1), 35–45. <https://doi.org/10.1177/1469787417721363>
- Cohen, L., Manion, L., & Morrison, K. (2000). *Research methods in education* (5th ed.). Routledge.
- Cunningham-Nelson, S., Baktashmotlagh, M., & Boles, W. (2019). Visualizing student opinion through text analysis. *IEEE Transactions on Education*, 62(4), 305–311. <https://doi.org/10.1109/TE.2019.2924385>
- Falcon, S., Admiraal, W., & Leon, J. (2023). Teachers' engaging messages and the relationship with students' performance and teachers' enthusiasm. *Learning and Instruction*, 86, 101750. <https://doi.org/10.1016/j.learninstruc.2023.101750>
- Floridi, L., & Chiriatti, M. (2020). GPT-3: Its nature, scope, limits, and consequences. *Minds and Machines*, 30(4), 681–694. <https://doi.org/10.1007/s11023-020-09548-1>
- Freelon, D. G. (2010). ReCal: Intercoder reliability calculation as a web service. *International Journal of Internet Science*, 1, 20–33.
- Geng, S., Niu, B., Feng, Y., & Huang, M. (2020). Understanding the focal points and sentiment of learners in MOOC reviews: A machine learning and SC-LIWC-based approach. *British Journal of Educational Technology*, 51(5), 1785–1803. <https://doi.org/10.1111/bjjet.12999>
- Harbour, K. E., Evanovich, L. L., Sweigart, C. A., & Hughes, L. E. (2015). A brief review of effective teaching practices that maximize student engagement. *Preventing School Failure*, 59(1), 5–13. <https://doi.org/10.1080/1045988X.2014.919136>
- Hardman, J. (2016). Tutor–student interaction in seminar teaching: Implications for professional development. *Active Learning in Higher Education*, 17(1), 63–76. <https://doi.org/10.1177/1469787415616728>
- Hattie, J. (2008). Visible learning: A synthesis of over 800 meta-analyses relating to achievement. *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. Routledge. <https://doi.org/10.4324/9780203887332>
- Hills, C., Levett-Jones, T., Warren-Forward, H., & Lapkin, S. (2016). Teaching and learning preferences of 'Generation Y' occupational therapy students in practice education. *International Journal of Therapy and Rehabilitation*, 23(8), 371–379. <https://doi.org/10.12968/ijtr.2016.23.8.371>
- Hirschberg, J., & Manning, C. D. (2015). Advances in natural language processing. *Science*, 349(6245), 261–266.
- Hynninen, T., Knutas, A., Hujala, M., & Arminen, H. (2019). Distinguishing the themes emerging from masses of open student feedback. 2019 42nd International Convention on Information and Communication Technology, Electronics and Microelectronics, MIPRO 2019 - Proceedings, 557–561. <https://doi.org/10.23919/MIPRO.2019.8756781>
- JASP Team. (2022). *JASP (Version 0.16.2)*.
- Johnson, R. B., & Onwuegbuzie, A. J. (2004). Mixed methods research: A research paradigm whose time has come. *Educational Researcher*, 33(7), 14–26. <https://doi.org/10.3102/0013189X033007014>
- Johnson, R. B., & Onwuegbuzie, A. J. (2007). Toward a definition of mixed methods research. *Journal of Mixed Methods Research*, 1(2), 112–133. <https://doi.org/10.1177/1558689806298224>
- Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., Bridgland, A., Meyer, C., Kohl, S. A. A., Ballard, A. J., Cowie, A., Romera-Paredes, B., Nikolov, S., Jain, R., Adler, J., ... Hassabis, D. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature*, 596, 583. <https://doi.org/10.1038/s41586-021-03819-2>
- King, E., & La Paro, K. (2015). Teachers' language in interactions: An exploratory examination of mental state talk in early childhood education classrooms. *Early Education and Development*, 26(2), 245–263. <https://doi.org/10.1080/10409289.2015.989029>
- Könings, K. D., Brand-Gruwel, S., & van Merriënboer, J. J. G. (2011). The match between students' lesson perceptions and preferences: Relations with student characteristics and the importance of motivation. *Educational Research*, 53(4), 439–457. <https://doi.org/10.1080/00131881.2011.625155>
- Leong, C. K., Lee, Y. H., & Mak, W. K. (2012). Mining sentiments in SMS texts for teaching evaluation. *Expert System Application*, 39, 2584–2589. <https://doi.org/10.1016/j.eswa.2011.08.113>
- Lodico, M. G., Spaulding, D. T., & Voegtle, K. H. (2010). *Methods in educational research: From theory to practice* (Vol. 28). John Wiley.

- Macfayden, L. P., Chase, M. M., Reeder, K., & Roche, J. (2003). Matches and mismatches in intercultural learning: design and facilitation of an online intercultural course. *UNESCO Conference on Intercultural Education*, 15–18.
- Marsh, H. W. (1994). Sport motivation orientations: Beware of jingle-jangle fallacies. *Journal of Sport & Exercise Psychology*, 16(4), 365–380.
- Marsh, H. W., Craven, R. G., Hinkley, J. W., & Debus, R. L. (2003). Evaluation of the Big-Two-Factor theory of academic motivation orientations: An evaluation of jingle-jangle fallacies. *Multivariate Behavioral Research*, 38(2), 189–224. https://doi.org/10.1207/S15327906MBR3802_3
- Minhas, P. S., Ghosh, A., & Swanz, L. (2012). The effects of passive and active learning on student preference and performance in an undergraduate basic science course. *Anatomical Sciences Education*, 5(4), 200–207. <https://doi.org/10.1002/ase.1274>
- Moliní Fernández, F., & Sánchez-González, D. (2019). Fomentar la participación en clase de los estudiantes universitarios y evaluarla. *REDU. Revista De Docencia Universitaria*, 17(1), 211. <https://doi.org/10.4995/redu.2019.10702>
- O'Connor, C., & Joffe, H. (2020). Intercoder reliability in qualitative research: Debates and practical guidelines. *International Journal of Qualitative Methods*. <https://doi.org/10.1177/1609406919899220>
- Opdecam, E., Everaert, P., Van Keer, H., & Buyschaert, F. (2014). Preferences for team learning and lecture-based learning among first-year undergraduate accounting students. *Research in Higher Education*, 55(4), 400–432. <https://doi.org/10.1007/s11162-013-9315-6>
- OpenAI. (2022). Examples - OpenAI API. <https://Beta.Openai.Com/Examples>.
- Pong-inwong, C., & Songpan, W. (2019). Sentiment analysis in teaching evaluations using sentiment phrase pattern matching (SPPM) based on association mining. *International Journal of Machine Learning and Cybernetics*, 10(8), 2177–2186. <https://doi.org/10.1007/s13042-018-0800-2>
- Qiu, X. P., Sun, T. X., Xu, Y. G., Shao, Y. F., Dai, N., & Huang, X. J. (2020). Pre-trained models for natural language processing: A survey. *Science China Technological Sciences*, 63(10), 1872–1897. <https://doi.org/10.1007/s11431-020-1647-3>
- Rahman, M. S. (2016). The advantages and disadvantages of using qualitative and quantitative approaches and methods in language “testing and assessment” research: A literature review. *Journal of Education and Learning*, 6(1), 102. <https://doi.org/10.5539/jel.v6n1p102>
- Rani, S., & Kumar, P. (2017). A sentiment analysis system to improve teaching and learning. *Advances in Learning Technologies*. <https://doi.org/10.1109/MC.2017.133>
- Roberts, D. (2019). Higher education lectures: From passive to active learning via imagery? *Active Learning in Higher Education*, 20(1), 63–77. <https://doi.org/10.1177/1469787417731198>
- Rodgers, W. J., Kennedy, M. J., VanUitert, V. J., & Myers, A. M. (2019). Delivering performance feedback to teachers using technology-based observation and coaching tools. *Intervention in School and Clinic*, 55(2), 103–112. <https://doi.org/10.1177/1053451219837640>
- Russ, R. S. (2018). Characterizing teacher attention to student thinking: A role for epistemological messages. *Journal of Research in Science Teaching*, 55(1), 94–120. <https://doi.org/10.1002/tea.21414>
- Rybinski, K., & Kopciszewska, E. (2021). Will artificial intelligence revolutionise the student evaluation of teaching? A big data study of 1.6 million student reviews. *Assessment and Evaluation in Higher Education*, 46, 1127–1139. <https://doi.org/10.1080/02602938.2020.1844866>
- Slater, D. R., & Davies, R. (2020). Student preferences for learning resources on a land-based postgraduate online degree programme. *Online Learning Journal*, 24(1), 140–161.
- Slavin, R. E., & Lake, C. (2008). Effective programs in elementary mathematics: A best-evidence synthesis. *Review of Educational Research*, 78(3), 427–515. <https://doi.org/10.3102/0034654308317473>
- Smith, C. D., & Baik, C. (2021). High-impact teaching practices in higher education: A best evidence review. *Studies in Higher Education*, 46(8), 1696–1713. <https://doi.org/10.1080/03075079.2019.1698539>
- Steen-Utheim, A., & Wittek, A. L. (2017). Dialogic feedback and potentialities for student learning. *Learning, Culture and Social Interaction*, 15, 18–30. <https://doi.org/10.1016/j.lcsi.2017.06.002>
- Tronchoni, H., Izquierdo, C., & Anguera, M. T. (2021). Regulación de la interacción participativa en clases universitarias expositivas. Propuesta formativa co-constructiva basada en la metodología observacional como estrategia mixed methods. *Publicaciones*, 52(2), 89–110. <https://doi.org/10.30827/publicaciones.v52i2.20751>
- UNESCO Education Sector. (2019). *Artificial Intelligence in Education: Challenges and Opportunities for Sustainable Development*. <https://en.unesco.org/themes/education-policy>
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit Jakob and Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2017). Attention Is All You Need. In I. Guyon, U. V Luxburg, S. Bengio, H. Wallach,

- R. Fergus, S. Vishwanathan, & R. Garnett (Eds.), *ADVANCES IN NEURAL INFORMATION PROCESSING SYSTEMS 30 (NIPS 2017)* (Vol. 30).
- Vercellotti, M. L. (2018). Do interactive learning spaces increase student achievement? A comparison of classroom context. *Active Learning in Higher Education*, 19(3), 197–210. <https://doi.org/10.1177/1469787417735606>
- Yang, F. Y., & Tsai, C. C. (2008). Investigating university student preferences and beliefs about learning in the web-based context. *Computers and Education*, 50(4), 1284–1303. <https://doi.org/10.1016/j.compedu.2006.12.009>
- Zhou, J., & Ye, J. M. (2020). Sentiment analysis in education research: a review of journal publications. *Interactive Learning Environments*. <https://doi.org/10.1080/10494820.2020.1826985>

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

Carmen Álvarez-Álvarez Carmen is Doctor in Pedagogy and works as a lecturer in the Department of Education at the University of Cantabria. Her research interests focus on theory–practice relations, school organisation, reading promotion and rural schools.

Samuel Falcon Samuel is a PhD student at the University of Las Palmas de Gran Canaria. He is developing his line of research on the study of secondary school teachers' messages using AI-based tools.