Emotion extraction from Likert-Scale questionnaires
– an additional dimension to Psychology Instruments –

Renata Magalhães[1] [0009-0005-4725-2883], Francisco S. Marcondes[1] [0000-0002-2221-2261], Dalila Durães[1] [0000-0002-8313-7023], and Paulo Novais[1] [0000-0002-3549-0754]

ALGORITMI Centre / LASI, University of Minho, Braga, Portugal
pg34321@alunos.uminho.pt, francisco.marcondes@algoritmi.uminho.pt, {dad|pjon}@di.uminho.pt

Abstract. Sentiment analysis tasks are used in various domains, including education. Likert-scale questionnaires are often used to gain insights into the respondents' views in various contexts. However, these questionnaires can allow for more information than they are designed for. This research paper explores an emotion classification technique for extracting emotional information from likert-scale questionnaires. A case study is presented in which a tailored questionnaire was employed to gather students' opinions on school-related matters, such as learning importance, academic performance and family and peer involvement and support. The students ($n = 845$) answered the questionnaire using a scale from totally disagree to totally agree. Through this questionnaire-based approach, data on students' emotions was collected.

Keywords: sentiment analysis · emotion classification · natural language processing · likert scale questionnaires · education.

1 Introduction

Likert-scale questionnaires are used in a variety of fields, such as marketing, education, medicine, nutrition, and nursing, but mostly in psychology [1]. It aims to measure responses to a statement within a scale of agreement from totally disagree to totally agree.

This research hypothesis is that it is possible to extract additional features from already existing Likert-Scales questionnaires, which would provide additional analysis dimensions. For this paper, the objective is to explore that hypothesis within the scope of emotion classification. Thus, given a Likert-scale questionnaire, the emotion of the respondent is to be assessed.

This paper was developed within the scope of a project aiming to prevent school failure and dropout. A Likert-scale questionnaire was developed as a supporting tool and applied to 845 students in northern Portugal. It was upon that questionnaire that this proposal is relying on, yet generalizable for any Likert-scale questionnaire with similar features. As a remark, the questionnaire is restrained due to copyright. Overall, the understanding that comes from emotion classification allows for improved interventions and strategies as well as support systems to address emotional well-being of the students.

Ethical Statement. The questionnaire utilized was designed by a team of educational psychologists. Considering that the participants in this study were underaged, parents were asked to sign an informed consent form. Researchers were only in charge of data analysis.

2 Related Work

Sentiment analysis, also known as opinion mining, is an area of research in the field of text mining that extracts and classifies sentiment, emotions, and attitudes expressed in text data [2]. Techniques such as natural language processing (NLP) and machine learning are used to enable the identification and analysis of emotions expressed in written or spoken language [3]. This is widely applied in various fields such as healthcare [4], customer feedback analysis [5], social media [6], and in education [7].

Emotion classification is an area of natural language processing, which allows the detection of emotions through data retrieved from the target user [8]. Current literature finds a varity of models for emotion classification performed either by text, audio, image or video [9]. Emotion lexicons are available for use, either manually annotated (*e.g.* NRC Emotion Lexicon and LIWC) or generated automatically using machine learning algorithms (*e.g.* WordNet Affect and SentiWordNet). These were not adquate for the context of this project as it is necessary to have spontaneous and genuine text in order to obtain accurate results. It is not the case as the aim is to assess the emotions that might be articulated on the questionnaire respondent given a question written by another person.

A questionnaire-approach, the AEQ, was developed in 2011 [10] to measure emotions in students' learning and performance. The framework used for defining emotions was the control-value theory of achievement, and different emotions were studied (enjoyment, hope, pride, relief, anger, anxiety, hopelessness, shame, and boredom). Overall, the AEQ contains 24 scales that tap these nine emotions occurring in three different academic achievement settings: class-related emotions, learning-relating emotions and test emotions. Students had to respond to this questionnaire with a 5 point Likert scale (1 = completely disagree, 5= completely agree). The AEQ is not adequate for the context of this paper as it aims to specifically assess student's emotions, whereas the goal is not to extract and classify emotions directly, but to investigate the possibility of extracting emotions as additional feautres from a questionnaire that was not designed for that purpose.

3 The Emotion Dimension of Likert-Scale Questionnaires

Likert-scale questionnaires present statements to which the respondent has to reply to within a scale. This could be a scale of agreement (totally disagree to totally agree), a satisfaction scale (very dissatisfied to very satisfied), a likelihood scale (very unlikely to very likely), a good to bad scale (very bad to very good) or a frequency scale (never to always). This is usually presented in a five point scale, in which 1 represents the lowest score, 3 represents neutrality and 5 represents the highest score [11]. Scores 1 and 5 are usually variations of intensities of either the positive or negative poles.

Through Likert-Scale questionnaires it is possible to measure responses to a statement, that is, personal opinions and they are often used in correlation to emotional states. Psychologists have been struggling to find agreement in regards to a definition of emotion, yet the notion that it is complex is consensual. Plutchik's model of emotion is perhaps the most popular model for dimensional models [12] and is widely used in a varitety of research, including this study. Plutchik defines emotions as

> *a multifaceted and inferred response to a stimulus, including subjective feelings, cognitions, impulses to actions and behaviour* [13, p. 551]

Following Plutchik's definition of emotion as a response to a stimulus, considering a Likert-scale questionnaire inquiry a stimulus, it is possible to say that, in addition to responding to a question, the respondent is also articulating an emotion that is dimmed by the intensity of the response. Therefore, combining the emotion associated to the question text with the reply to that question, it is possible to suggest the emotion that the respondent may be articulating.

Consider, for a reference, the Theory of Brain Functions and Behaviour *cf.* [14]. In short, it is a 4-point Likert scale questionary for assessing Behavioral Inhibition System (BIS), which regulates anxiety responses to stimuli associated with punishment, absence of reward, and novelty, while the Behavioral Approach System (BAS) governs appetitive motivation and responds to cues indicating rewards, nonpunishment, and escape from punishment. As a self-report questionnaire it has been used in various contexts, including for the elicitation of students' emotional states [15] in university examinations.

Looking at BIS-BAS, the first statement reads "A person's family is the most important thing in life". If the respondent responds with "very true for me", it means that the respondent values and cherishes family. Emotionally speaking, this statement articulates feelings of anxiety, submission, and sentimentality. However, if the respondent responds with "very false for me", it does not necessarily mean that the person hates their family. Emotionally speaking, this statement articulates a combination of feelings of outrage, contempt, and morbidness or derisiveness.

According to Plutchik's circumplex model, outrage is a combination of surprise and anger, contempt is a combination of disgust and anger and morbidness or derisiveness are a combination of disgust and joy. This suggests a complex combination of contrasting feelings of frustration, energy, distrust, disinterest and contempt.

Plutchik's three-dimensional circumplex model illustrates the interconnections between emotions using a color wheel analogy. It incorporates a vertical dimension for intensity, a circular dimension for similarity, and eight sectors representing primary emotions, with the empty spaces representing dyads, which are mixtures of two primary emotions [13]. The model's eight primary emotions dimensions can be arranged in four pairs of opposites: joy and sadness, fear and anger, anticipation and surprise, disgust and trust, as shown in figure 1.

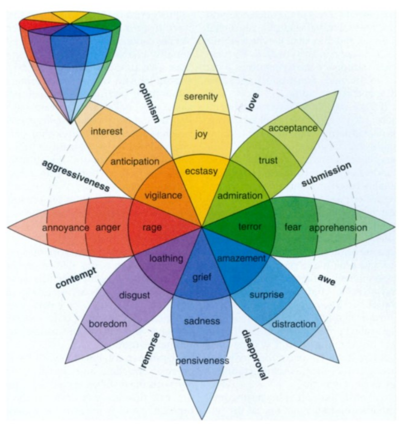

Fig. 1: Plutchik's circumplex model of emotions [13, p.349].

Plutchik states that emotions can combined in the same way that blue and red make purple, and, for instance, mixing joy and acceptance produces love; disgust and anger produces hatred or hostility.

It is of extreme importance to assess a whole questionnaire before proceeding to its emotional classification, to make sure there are no assumptions but evidence of emotion articulation. That is, by observing the presence of specific emotions in several statements of the questionnaire, that emotion can be accurately extracted, whereas if it is just present in one statement, it could not necessarily mean that that specific emotion is being articulated by the respondent as he is just marking an answer on a scale but there is no reason or explanation as to why that answer was chosen. If a pattern or frequent presence is found, however, it is plausible to extract and classify that emotion.

This means that emotion classification is performed by classifying each statement on the questionnaire, with a score being assigned to each emotion present on that statement. For this paper, a lexical approach *cf.* [16] was used, for an instance refer to table 1a. Thus, when the respondent reads the statement, it receives a stimulus, meaning that an emotion is being articulated. However, it is also necessary to consider the answer in order to verify its suitability. Each emotion articulated in table 1a is then multiplied by the response weight presented in table 1b and added to the respondent score. Since each emotion has its inverse, therefore, if the respondent chose "fully disagree" for the statement in table 1a does not mean, for an instance, `anxiety = -0.4` but `hate = 0.4`.

Table 1: Demonstration of the proposed approach.

statement	anxiety	sentiment.	submission
A person's family is the most important thing in life	0.4	0.6	0.7

(a) Example of dataset for emotion classification proposed in this study.

response	fully agree	partially agree	not sure	partially disagree	fully disagree
emotion weight	1.0	0.5	0.0	-0.5	-1.0

(b) Demonstration of emotion classification calculation.

4 Emotion Extraction on a Drop-out Risk Assessement Instrument

The proposed approach was applied to a Likert-scale questionnaire of about 50 statements which was developed by a team of educational psychologists and was administered to a total of 845 students, aged between 10 and 17 years old from northern Portugal. The purpose of the questionnaire was to evaluate multiple aspects including personal motivation, engagement in school, task management, perseverance, beliefs and attitudes towards the educational environment, active participation, family support, interpersonal relationships, and students' perceptions of teachers and the school environment. Participants were requested to indicate their level of agreement or disagreement to each statement, in which the responses available are "totally disagree", "disagree", "I am not sure", "agree", and "totally agree".

Examples of statements include "It is important to me to learn as much as I can", "I forget important deadlines" and "My family/parents are there when I need them" and, therefore, each section of the questionnaire articulates specific emotions within the respondent.

The first step was to identify emotions that were of interest in the questionnaire. Then, researchers and psychologists engaged in discussions to identify the emotions that would make sense in the context of the project, as according to current emotion detection practices [17].

Psychologists reinforced the importance of studying emotions that usually occur in an educational context and that are comprised in different spectrums of Plutchik's circumplex model.

Thus, the following emotions were analyzed within the study's specific context: happiness (also known as joy), trust, optimism, interest, boredom, anxiety, distraction and shame. Within Plutchnik's Wheel of Emotions, joy and trust are recognized as two of the foundational primary emotions. Additionally, both optimism and interest are accounted for within the wheel. Optimism is characterized as a combination of anticipation and joy, while interest represents a lower intensity form of anticipation and vigilance. Anxiety, on the other hand, emerges from a combination of anticipation and fear, shame is rooted in a fusion of fear and disgust, boredom is a lower intensity of loathing and disgust, and distraction is a lower intensity of amazement and surprise, as well as a direct opposite of interest.

Each section contained emotional inferences for all emotions, but different emotions were more prevailing in different sections. The predominant emotions in the importance of learning and academic performance section are anxiety, interest and shame, in task management, preserverance and self-regulation are distraction and boredom, in beliefs and attitudes related to the educational environment is optimism, in attitudes and behaviours related to active participation and engagement in school is interest, in family support and involvement in school-related matters are happiness and anxiety, in friendship and peers relationships within the school context are happiness, trust and interest and in perceptions of students regarding their teachers and the overall school environment is trust.

A dataset was created that contains all the statements in the questionnaire, as well as a column for each of the chosen emotions. For each statement in the questionnaire a score was assigned to each emotion bearing in mind which emotions are articulated to the respondent when agreeing with the statement. This annotation was based on the intensity, typical sensations, similar words and utility of that specific emotion present in Plutchnik's Wheel. Assuming that the student fully agreed with that statement, a specific emotion or combination of emotions would take place, that is, a certain emotional inference would be produced. Subsequently, the scores for each emotion based on the student's response was calculated as proposed.

As illustration, an excerpt of the questionnaire classification can be seen in table 2. Once the questionnaire was filled out, a normalized value between -1 to 1 for each emotion, per student, is calculated as illustrated in table 3.

Table 2: Emotion Classification for the Questionnaire.

question	happiness	trust	optimism	interest	boredom	anxiety	distraction	shame
Q1_1	0.3	0.2	0.5	0.8	0	0.05	0	0
Q1_2	0.2	0.2	0.5	0.8	0.1	0.05	0	0.05
Q1_3	0.2	0.1	0.3	0.4	0	0.3	0.05	0.1

Table 3: Emotion Classification results.

student	happiness	trust	optimism	interest	boredom	anxiety	distraction	shame
1	0,70	0,70	0,66	0,66	0,5	0,55	0,47	0,56
2	0,76	0,80	0,76	0,75	0,83	0,75	0,84	0,80
3	0,54	0,50	0,51	0,45	0,47	0,66	0,45	0,66

The emotion classification model developed in the context of this study allowed for the gathering of students emotional scores on happiness, trust, optimism, interest, boredom, distraction, anxiety and shame, which were then combined into a dataset containing other important features for analysis such as school year and gender, for example.

The dataset utilized for this study contained a total of 845 students, from which 415 are female and the remaining 430 male. In terms of the school year, distribution is as follows: 212 from grade 5, 210 from grade 6, 138 from grade 7, 155 from grade 8 and 130 from grade 9, therefore showing an equitable gender and school year distribution.

After careful analysis of the results obtained, it was found that the emotional state that was most present among the eight classified was anxiety, followed by shame, boredom, distraction, interest, happiness, trust and optimism, at last. The mean value for anxiety is 0.59 while for optimism is 0.48. All mean values can be seen distribution can be visualized in the table 4.

Table 4: Mean Values for all Emotions.

Anxiety	**Shame**	**Boredom**	**Distraction**	**Interest**	**Happiness**	**Trust**	**Optimism**
0.59	0.57	0.53	0.53	0.50	0.50	0.49	0.48

While anxiety was the emotion that contained the highest mean value, it is worth mentioning that the highest scores were found in grade 5 students. Keeping in mind that this questionnaire was administered in the beginning of the school year, this can be explained due to the fact that students move to a new school at grade 5. Anxiety is the combination of anticipation and fear, meaning that it is characterized by feelings of being alert, stressed and scared. This transition, which often includes students integrating a new class with new teachers and staff members at a different school can take a toll on students' anxiety scores. It is also worth mentioning that overall anxiety results were higher for male students. In fact, male students obtained a higher score of happiness, trust, boredom, anxiety, distraction and shame, whereas female students achieved higher levels of interest and optimism. It was also observed that male students report higher scores of anxiety in grades 5, 7 and 8, whereas female students report higher anxiety scores at grade 6 and 9. However, the difference of scores in grade 6 is almost non-existential.

Grade 5 students obtained higher values of all emotions in comparison to other school years. Overall, happiness, trust, optimism, and interest show a gradual decline as students progress from 5th grade to 9th grade. This could suggest a potential decrease in emotional well-being during the school years.

The transition from 6th grade to 7th grade appears to show the first signs of a notable impact on emotional experiences. Happiness, trust, and optimism experience a sharp decline, while boredom and distraction increase. This transition period may pose challenges for students' emotional well-being as it is a time of more responsibility and workload. In Portugal, students' transition to the 7th grade is known for being challenging as students start having more subjects than they did in the 5th and 6th grade.

However, lowest results occur at grade 8, especifically for happiness, interest, trust and optimism. There is a slight increase of these emotions from grade 8 to grade 9. Although reaching their lowest score also at grade 8, shame and anxiety scores overall remain more consistent throughout all school years in comparison to other emotions. Positive emotions register a significant decrease at grade 8, while anxiety, shame, boredom and distraction remain more consistent. This could indicate that grade 8 is an emotionally demanding year for students. While female students also report a small decrease in positive emotion scores at grade 8, this steady decrease is more remarkably noticed in male students.

Figure 2 shows the distribuiton of emotion scores per school year.

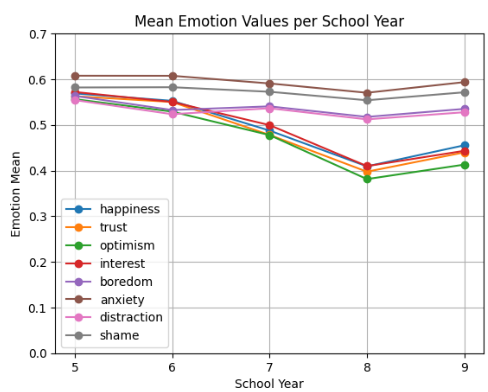

Fig. 2: Emotion Mean Distribution Per School Year.

Out of the eight emotions, shame is the one that remains more constant in different school years. With a slighty bigger mean value for male students, shame seems to remain consistent in the 5th and 6th grade, decreasing slightly in the 8th grade and experiencing a small increase by the 9th grade. Shame is a combined emotion, meaning that it occurs with a combination of fear and disgust. Typical sensations of shame include feeling bitter, and stressed. Male students report higher scores of shame at every school year, except for the 9th grade.

5 Discussion and Future Work

The proposed metholody shows that additional features can be extracted from likert-scale questionnaires. The case study presented shows that this questionaire-based emotion classification approach enabled an exploratory analysis of students' emotional states, taking place at the beginning of the school year, which contributes to the body of knowledge in sentiment analysis and emotion classification.

It was found that anxiety scores were the highest. This result was not expected. In contrast, optimism scores were found to be the lowest. Distraction and boredom are higher at grade 7 than 6, whereas interest, happiness, trust and optimism are found to be the opposite, perhaps due to the fact that students' workload is much higher in grade 7 and students usually have new teachers at this time. The lowest scores for all positive emotions were found at grade 8, which was expected due to the reputation that grade 8 holds for being a demanding year in Portugal.

An unexpected finding is that average anxiety results were slightly higher for male students. Previous studies report that female students usually achieve higher anxiety scores than male students [10] [18]. Though only a small difference was found, this result can be paired with shame, for which male students also reported a slightly higher result than female students. However, answers to the questionnaire indicate that overall male students are more worried about improving their scores, showing others that they are good with school work and overall maintaing a good image among their peers, therefore explaining higher anxiety scores and validating the results obtained from the proposed emotion classification approach.

This research highlights the importance of developing and refining emotion detection techniques specifically tailored to specific domains and languages. As sentiment analysis continues to expand beyond English [2], it is crucial to invest in the development of accurate and extended emotion lexicons in other languages, as the majority of current lexicons are mostly available for the English language and are found to be incomplete, incorrectly translated, or simply containing transliterations of the original English terms, as is the case of the NRC Emotion Lexicon (also known as EmoLex). This presents a challenge for sentiment analysis in languages other than English. This research does not encounter that limitation as individual classification should be made for each statement in the questionnaire, performed in the questionnaire's language.

Multiple studies indicate that support and guidance from both teachers and parents are crucial factors in fostering student motivation and interest in school [19]. Taking this into account, the questionnaire used in this study pinpoints all the important factors in terms of academic success, ranging from the students' own views on their educational goals to their peers and family support, therefore creating an non-intrusive, undemanding, and reliable form of gaining insights into students' emotional states. Sentiment analysis and more specifically emotion detection can be used as a monitoring as well as predictive tool, creating advantages for the educational system at a low cost [20]. A future work hypotheses would be to have students repeat the whole questionnaire at the beginning and end of the school year to compare results and therefore making it possible to monitor students' emotional state throughout the school year.

This is just an example of the utlization of the proposed approach. When applied to questionnaires of different fields of study, this metholody can have the same impact, allowing for specific interventions in different domains.

6 Conclusion

The methodology proposed by this research shows that by classifying a likert-scale questionnaire for the purpose of emotion classification will result in a fine-tuned model that can be applied to any questionnaire of any field of study.

A case study was presented in which this approach was applied to a tailored likert-scale questionnaire administered to a sample of 845 K-12 students at the beginning of the school year. This questionnaire, designed by a team of educational psychologists, presents statements in relation to the importance of learning and academic performance within the school context and students had to answer within a scale of agreement and disagreement. It was possible to extract insights of aspects of students' experiences within the educational environment, and emotional levels for happiness, trust, optimism, interest, anxiety, distraction, boredom and shame.

In conclusion, this study showcases a new emotion classification model, implemented through likert-scale questionnaires. In an educational context it can aid in detecting students' emotional levels and allowed for a timely intervention or prevention of school failure. Though, this method can be applied to questionnaires of different areas of study, allowing for tailored interventions wihtin different contexts.

Acknowledgements This work is financed by National Funds through the Portuguese funding agency, FCT - Fundação para a Ciência e a Tecnologia within project 2022.06822.PTDC.

References

1. Jonald L Pimentel. A note on the usage of likert scaling for research data analysis. *USM R&D Journal*, 18(2):109–112, 2010.
2. Walaa Medhat, Ahmed Hassan, and Hoda Korashy. Sentiment analysis algorithms and applications: A survey. *Ain Shams engineering journal*, 5(4):1093–1113, 2014.
3. Jingfeng Cui, Zhaoxia Wang, Seng-Beng Ho, and Erik Cambria. Survey on sentiment analysis: evolution of research methods and topics. *Artificial Intelligence Review*, pages 1–42, 2023.

4. Madison Milne-Ives, Caroline de Cock, Ernest Lim, Melissa Harper Shehadeh, Nick de Pennington, Guy Mole, Eduardo Normando, and Edward Meinert. The effectiveness of artificial intelligence conversational agents in health care: systematic review. *Journal of medical Internet research*, 22(10):e20346, 2020.
5. Amjad Iqbal, Rashid Amin, Javed Iqbal, Roobaea Alroobaea, Ahmed Binmahfoudh, and Mudassar Hussain. Sentiment analysis of consumer reviews using deep learning. *Sustainability*, 14(17):10844, 2022.
6. Maryum Bibi, Wajid Arshad Abbasi, Wajid Aziz, Sundus Khalil, Mueen Uddin, Celestine Iwendi, and Thippa Reddy Gadekallu. A novel unsupervised ensemble framework using concept-based linguistic methods and machine learning for twitter sentiment analysis. *Pattern Recognition Letters*, 158:80–86, 2022.
7. Delali Kwasi Dake and Esther Gyimah. Using sentiment analysis to evaluate qualitative students' responses. *Education and Information Technologies*, 28(4):4629–4647, 2023.
8. Weiyuan Li and Hua Xu. Text-based emotion classification using emotion cause extraction. *Expert Systems with Applications*, 41(4):1742–1749, 2014.
9. J Jayalekshmi and Tessy Mathew. Facial expression recognition and emotion classification system for sentiment analysis. In *2017 International Conference on Networks & Advances in Computational Technologies (NetACT)*, pages 1–8. IEEE, 2017.
10. Reinhard Pekrun, Thomas Goetz, Anne C Frenzel, Petra Barchfeld, and Raymond P Perry. Measuring emotions in students' learning and performance: The achievement emotions questionnaire (aeq). *Contemporary educational psychology*, 36(1):36–48, 2011.
11. Sarang Shaikh and Sher Muhammad Doudpotta. Aspects based opinion mining for teacher and course evaluation. *Sukkur IBA Journal of Computing and Mathematical Sciences*, 3(1):34–43, 2019.
12. Lai Po Hung and Suraya Alias. Beyond sentiment analysis: A review of recent trends in text based sentiment analysis and emotion detection. *Journal of Advanced Computational Intelligence and Intelligent Informatics*, 27(1):84–95, 2023.
13. Robert Plutchik. The nature of emotions: Human emotions have deep evolutionary roots, a fact that may explain their complexity and provide tools for clinical practice. *American scientist*, 89(4):344–350, 2001.
14. Charles Carver and Teri White. Behavioral inhibition, behavioral activation, and affective responses to impending reward and punishment: The bis/bas scales. *Journal of Personality and Social Psychology*, 67:319–333, 08 1994.
15. Dino Krupić and Philip J Corr. Individual differences in emotion elicitation in university examinations: A quasi-experimental study. *Personality and Individual Differences*, 71:176–180, 2014.
16. Dan Jurafsky and James H. Martin. *Speech and Language Processing.* draft (https://web.stanford.edu/ jurafsky/slp3/), third edition, 2023.
17. Saif M Mohammad. Best practices in the creation and use of emotion lexicons. *arXiv preprint arXiv:2210.07206*, 2022.
18. Leili Hosseini and Homayoun Khazali. Comparing the level of anxiety in male & female school students. *Procedia-social and behavioral sciences*, 84:41–46, 2013.
19. Gaetana Affuso, Anna Zannone, Concetta Esposito, Maddalena Pannone, Maria Concetta Miranda, Grazia De Angelis, Serena Aquilar, Mirella Dragone, and Dario Bacchini. The effects of teacher support, parental monitoring, motivation and self-efficacy on academic performance over time. *European Journal of Psychology of Education*, 38(1):1–23, 2023.

20. Leo CE Huberts, Marit Schoonhoven, and Ronald JMM Does. Multilevel process monitoring: A case study to predict student success or failure. *Journal of Quality Technology*, 54(2):127–143, 2022.