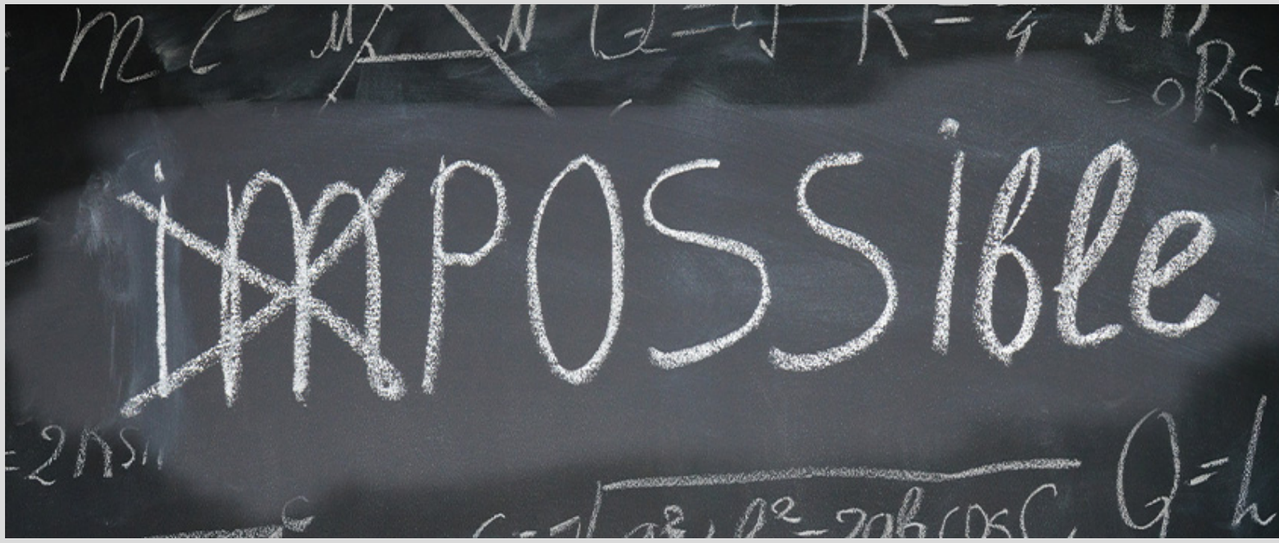
JOURNAL ON EFFICIENCY AND RESPONSIBILITY IN EDUCATION AND SCIENCE

The ERIES Journal is being managed by an international editorial board as a regular scientific journal. A rigorous process of papers' reviews (double-blind peer review) is fully supported by a web-based submission system. The journal is published electronically four times a year, on March 31, June 30, September 30 and December 31 of the current year.

The journal is indexed in:

- BASE - Bielefeld Academic Search Engine
- Directory of Open Access Journals
- EBSCO database
- Emerging Sources Citation Index - Web of Science™ Core Collection
- ERIC - Education Resources Information Center
- ERIH PLUS
- Google Scholar
- Open Academic Journals Index
- ResearchBib
- SCOPUS
- The list of reviewed periodicals in the Czech Republic



www.eriesjournal.com

Scientific journal of the Czech University of Life Sciences Prague JOURNAL ON EFFICIENCY AND RESPONSIBILITY IN EDUCATION AND SCIENCE, distributed by the Faculty of Economics and Management. Published quarterly. Executive editors: Ing. Martin Flégl, Ph.D., Ing. Jiří Fejfar, Ph.D., PhDr. Michaela Cocca and Armando Cocca, PhD. Editorial Office**:** ERIES Journal, Czech University of Life Sciences Prague, CZ 165 21 Prague 6 - Suchdol, Czech Republic, email: editor@eriesjournal.com, tel: +420 224 382 355.

volume 17
issue 3

2024

An international peer-reviewed journal published by
Faculty of Economics and Management
Czech University of Life Sciences Prague

editor@eriesjournal.com
www.eriesjournal.com
Online ISSN: 1803-1617
Printed ISSN: 2336-2375

EDITORIAL BOARD

EDITORIAL OFFICE

ERIES Journal, Czech University of Life Sciences Prague (CZU), CZ 165 00 Prague 6 - Suchdol, Czech Republic
email: editor@eriesjournal.com * tel: +420 224 382 355
Registration number: MK ČR E 21414 © Czech University of Life Sciences Prague, Czech Republic, 2024

INSTRUCTIONS FOR AUTHORS

The Journal on Efficiency and Responsibility in Education and Science publishes papers of the following categories: full research papers, short communications, review studies and book reviews (on invitation only).

- FULL RESEARCH PAPERS
- SHORT COMMUNICATION
- REVIEW STUDY

Papers are published in English. A paper may comprise an empirical study using an acceptable research strategy, such as survey, case study, experiment, archival analysis, etc. It may contain a theoretical study aimed at advancing current theory or adapting theory to local conditions or it may arise from theoretical studies aimed at reviewing and/or synthesizing existing theory. Concepts and underlying principles should be emphasized, with enough background information to orient any reader who is not a specialist in the particular subject area.

Submission checklist

The paper. The paper is carefully formatted according to the template of the journal (see bellow). Special attention is paid to the exact application of the Harvard referencing convention to both continuous citations and list of references. If an electronic source has the DOI number assigned, also it will be provided in the list of references. Manuscripts are submitted via the editorial system in the DOC.

Research highlights. The core results, findings or conclusions of the paper are emphasized in 2-4 bullet points (max. 150 characters per bullet point including spaces). The highlights are submitted as a text into the submission form in the editorial system.

Copyright form. The submission of a paper will imply that, if accepted for publication, it will not be published elsewhere in the same form, in any language, without the consent of the Publisher. The manuscript submitted is accompanied by the copyright form signed by the corresponding author who declares the agreement of all authors with the conditions in the Form. The Form is submitted into the editorial system in the PDF format.

Suggested reviewers. It is required to suggest two experts appropriate to evaluation of the paper. The experts should be out of the affiliation of the author(s), Czech University of Life Sciences Prague, and also both experts should be from different affiliations. The reviewers are submitted into the text fields in the submission form of the editorial system.

Preparation of the manuscript (technical notes)

Authors are responsible for applying all requirements that are specified in the journal's paper template in individual sections. Especially, the paper must provide a short review of current state in the area of the paper's aim in Introduction. The paper should refer significant sources, particularly scientific journals or monographs.
Papers must be closely scrutinized for typographical and grammatical errors. If English is not author's first language then the paper should be proof-read by a native English-speaking person, preferably one with experience of writing for academic use. Spelling should follow the Oxford English Dictionary.
Tables, graphs and illustrations should be drawn using a suitable drawing package. Colour may be used. Place all diagrams and tables where you wish them to appear in the paper. Ensure your diagrams fit within the margins and are resizable without distortion.

Review procedure

Following Editorial recommendation, papers are submitted to a double-blind peer review process before publication. Commentary by reviewers will be summarized and sent by email to authors, who can choose to revise their papers in line with these remarks. Re-submitted papers should be accompanied by the description of the changes and other responses to reviewers' comments (see above), so that the desk-editor can easily see where changes have been made.

Copyright

Authors are fully responsible for the paper's originality and for correctness of its subject-matter, language and formal attributes. Author's statement should be enclosed declaring that the paper has not been published anywhere else.

The submission of a paper will imply that, if accepted for publication, it will not be published elsewhere in the same form, in any language, without the consent of the Publisher. Before publication, authors will be asked to complete a copyright release, giving the publisher permission to publish the paper in a specific issue of this Journal. Overall copyright ownership of the paper, however, remains with the author/s. It is the authors' responsibility to obtain written permission to quote material that has appeared in another publication.

EDITORIAL

We are pleased to present you with the third issue of 2024 (vol. 17, no. 3). This issue includes eight articles by authors from the Czech Republic, Indonesia, Israel, Hungary, Nigeria, and the Philippines, on topics related to students' learning motivation and self-efficacy, analyses of the impact of the COVID-19 pandemic on learning and examination, and challenges in mathematics education.

In the first article, "Results of Mathematics Examinations Before, During, and After the Covid-19 Related Restrictions", Eva Ulrychová, Renata Majovská, and Petr Tesař investigated the impact of the Covid-19 pandemic-related restrictions on results of mathematics examinations at the University of Finance and Administration in Prague. The authors compared the standard testing results from the period immediately before the pandemic (academic year 2018/2019) and alternative forms of testing during the pandemic (academic years 2019/2020 and 2020/2021). Further, the authors compared examination results after the return to the standard teaching and examination regime (academic year 2021/2022) with previous periods (academic years 2018/2019, 2019/2020, 2020/2021). The results revealed that the use of non-standard forms, although more challenging for teachers to control, did not lead to better results, as the results in the correspondence form were similar to the standard form and even worse in the online form.

In the second article, "Factor Analysis on the Motivation for Extensive Reading Questionnaire," Helta Anggia and Anita Habók examined factors adapted from the Motivation for Reading Questionnaire together with some items based on the extensive reading, principles, and Technology Acceptance Model to develop and validate a new measure of university students' motivation for reading following an extensive reading intervention. The study recruited 558 undergraduate students of English as a foreign language in Indonesia. The structure of the questionnaire was validated using exploratory and confirmatory factor analyses. Moreover, to determine the dependability of the instrument, the authors calculated the internal consistency reliabilities of the instrument as a whole and per factor. The study demonstrated that the subscales of the Motivation for Reading Questionnaire, in conjunction with the Technology Acceptance Model and extensive reading theories, can be used to assess students' motivation to read following extensive online reading programs.

In the third article, "Self-Reported Zoom Exhaustion and Fatigue Levels among Physical Education Teacher Education Students in a State University in the Philippines," Julius Ceazar G. Tolentino and John Paul P. Miranda assessed the self-reported Zoom exhaustion and fatigue (ZEF) levels of physical education teacher education (PETE) students who were attending flexible learning set up as a solution to continue classes during the COVID-19 pandemic. The study included responses from 303 students aged between 19 and 21. The analysis revealed that the PETE students were very tired and exhausted both in general and visually after a series of video conferences within a semester. Furthermore, the students were socially-, motivationally-, and emotionally- moderately tired and exhausted. As the authors conclude, the analysis findings may have direct implications for the physical education teacher education program but cannot be generalized to other teacher education students.

In the fourth article, "The Impact of the Learning Environment on Self-efficacy and Achievement Goals of Israeli Pre-service Teachers," Marcel Amasha examined a hypothetical, theoretical model to explain how pre-service teachers' perceptions of their own learning environment affect belief in their self-efficacy in teaching, and how this shapes their achievement goals in teaching as future educators. The study included 278 pre-service teachers at all five colleges in Israel offering teacher training programs. The hypothesized structural model was deemed a good fit for the data and was able to explain 35% of the variance in the mastery goals of pre-service teachers, 24% of the variance in performance-approach goals, and 65% of the variance in performance-avoidance goals. The analysis also showed that perception of the learning environment has a strong and significant impact on teaching ability and the achievement goals of pre-service teachers.

In the fifth article, "Gender Differences in School Achievement and Attitudes towards Motivation in Secondary Economic Education," Jaromír Novák, Kateřina Berková, Andrea Kubišová, and Dana Kolářová investigated school achievement and attitudes of Czech and Slovak secondary school students towards motivation in economic subjects. The aim of the study was to analyze the factors of students' motivation, their relations with selected teacher's competencies, and students' school achievement with regard to their gender. The analyzed sample involved 572 secondary school students between 17 and 19 years old. The data

were analyzed using the Mann-Whitney U-test, Pearson, and Spearman correlation coefficients. The results identified that girls performed better than boys in all the subjects analyzed. At the same time, for all factors of motivation with significant gender differences, girls' motivation was stronger compared to boys.

In the sixth article, "Motivation and Behavioral Engagement: The Mediating Role of Mathematics Self-Efficacy in Primary Education", Achmad Hidayatullah, Ratno Abidin, and Abdul Muqit explored the interrelation between motivation, self-efficacy, and behavioral engagement in primary school mathematics learning. This study also examined the mediating role of self-efficacy in the relationship between motivation and behavioral engagement. Using structural equation modeling, the authors evaluated responses from 660 fifth and sixth-grade students in four schools in Surabaya, Indonesia. The analysis identified motivation as key to enhancing students' self-efficacy and behavioral engagement during mathematics learning. Additionally, self-efficacy was found to be linked with students' behavioral engagement and a mediator in the relationship between motivation toward mathematics and behavioral engagement during mathematics learning.

In the seventh article, "Navigating the College Students' Adversities: The Role of Academic Buoyancy and Motivation on Learning Achievement," Mohamad Arief Rafsanjani, Handri Dian Wahyudi, Retno Mustika Dewi, and Putri Ulfa Kamalia analyzed the nexus between academic buoyancy and students' learning achievement by introducing the mediating variable, motivational constructs, which include self-efficacy, persistence, and anxiety. For this purpose, the authors utilized Structural equation modeling (SEM) on responses from 493 college students in Indonesia. The results revealed that academic buoyancy directly affects learning achievement. Furthermore, motivational constructs (self-efficacy, persistence, and anxiety) significantly mediate the relationship between academic buoyancy and learning achievement. Therefore, students with high academic buoyancy often possess strong self-efficacy, which empowers them to engage actively in learning and seek help when needed.

Finally, in the eight article "Effect of 2-PL and 3-PL Models on the Ability Estimate in Mathematics Binary Items", Rukayat Oyebola Iwintolu, Oluwaseyi Opesemowo and Phebean Oluwaseyi Adetutu focused on examining the influence of 2-parameter logistic (PL) and 3-parameter logistic models on the students' estimates ability in mathematical binary items. The authors used Item Response Theory (IRT) in the research survey design, with a sample comprising 1015 senior secondary (SS) students. The Mathematics Achievement Test instrument was adapted from the General Mathematics Paper 1 of the Senior School Certificate Examination administered by the West Africa Examinations Council (WAEC). The analysis indicates that the 3-PL model bestowed more precise estimates of examinees' abilities than the 2-PL model. This suggests that using the 3-PL model for mathematics assessments in Nigerian secondary schools may create more accurate and reliable results.

We would like to thank all authors who have submitted their articles to the ERIES Journal, and special thanks go to all reviewers for their endless effort in revising the articles. We hope that all our readers will find this third issue of the year appealing. You can also follow the latest updates related to the ERIES Journal on its LinkedIn page, where we post information about the published articles, highest cited articles, and related upcoming events.

Sincerely

Martin Flégl
Executive Editor
ERIES Journal
www.eriesjournal.com
www.linkedin.com/company/eriesjournal/
www.erie.pef.czu.cz

CONTENT

Full research paper

RESULTS OF MATHEMATICS EXAMINATIONS BEFORE, DURING, AND AFTER THE COVID-19 RELATED RESTRICTIONS

Eva Ulrychová✉
Renata Majovská
Petr Tesař

University of Finance and Administration, Czech Republic

✉ ulrychova@mail.vsfs.cz

ABSTRACT

The article deals with the results of mathematics examinations at the University of Finance and Administration in Prague before, during, and immediately after the Covid-19 pandemic-related restrictions. The first objective is to evaluate whether the non-standard forms of testing (correspondence and online), used on an emergency basis during the pandemic, were adequate compared to the standard form (face-to-face) applied before the pandemic. The second objective is to assess whether and to what extent the results of the examinations have changed after the return of teaching and testing methods to normal. It turns out that the use of non-standard forms, although more challenging for teachers to control, did not lead to better results – the results in the correspondence form were similar to the standard form and even worse in the online form. The results of examinations administered in the standard form after the return to normal teaching were significantly better than in any of the periods studied, including the standard form of examination before the pandemic. Possible reasons for the results are analysed in the paper.

KEYWORDS

Covid-19 pandemic, form of examination, mathematics, online exam, statistical evaluation, written test

HOW TO CITE

Ulrychová E., Majovská R., Tesař P. (2024) 'Results of Mathematics Examinations Before, During, and After the Covid-19 Related Restrictions', *Journal on Efficiency and Responsibility in Education and Science*, vol. 17, no. 3, pp. 187–194. http://dx.doi.org/10.7160/eriesj.2024.170301

Article history
Received
May 2, 2023
Received in revised form
August 20, 2023
Accepted
July 11, 2024
Available on-line
September 30, 2024

Highlights

- *Alternative forms of written examinations in mathematics during the Covid-19 pandemic.*
- *Results of written examinations in mathematics after the return to normal teaching mode.*

INTRODUCTION

The Covid-19 pandemic and associated measures have necessitated a change in the approach to teaching and testing in all types of schools. However, as mentioned, for example, by Ho et al. (2021) or Hvorecký et al. (2021), many schools were not prepared for the situation, especially in the first wave of the pandemic, and teaching and examinations were thus "emergency" mode. Many teachers around the world have begun to think about how to replace standard teaching so that students can pass exams (Makamure and Tsakeni, 2020). They believed that using modern technology would help teachers and students overcome problems (Pokorny, 2021). This was also the case at Prague University of Finance and Administration. This article compares the results of mathematics examinations at this school during the pandemic with the results before the pandemic and after the restrictions related to the pandemic. Exam results may depend on many factors. For example, Fajčíková et al. (2020) point out the influence of the field of study, Joyce et al. (2015) study the impact of class time on academic performance, Ulrychová and Bílková (2018) investigate the impact of students' gender on their mathematics exam results, Majovska (2015) deals with the influence of online mathematics programmes on students' results, Darolia (2014) analyses the effect of working on the academic performance.

During the pandemic, additional factors have been added that can affect the outcome of the exam, which is conducted online. Haus et al. (2020) compare the different scenarios of written exams during the pandemic depending on the number of students to be monitored in parallel to avoid cheating. Hartnett et al. (2023) state that most students were positive with online exams, but digital inequalities were found between students' perceptions of digital competence.

Exam results could also be affected by the form of face-to-face or online teaching. Sun et al. (2008) list the main factors influencing students' satisfaction with online learning, both on the part of instructors and students. Issues related

to student satisfaction with online learning have come to the forefront of many scientific publications in recent years due to the restrictions accompanying the Covid-19 pandemic. Aristovnik et al. (2020) found that at the time of the pandemic, students from Europe were more likely to be satisfied with their schooling than students from other continents.

However, many studies report that the form of teaching does not strongly influence learning outcomes. For example, Pasáčková (2021) found that student success rates in mathematics did not change with the conversion from face-to-face to online teaching. Trends examining alternative online education outcomes were already emerging before the COVID-19 pandemic. Thompson and McDowell (2019) conducted a research study at an undergraduate college, comparing student successes in a mathematics course offered fully online, blended, and face-to-face, and concluded that the level of student performance is independent of the form of teaching. In research conducted by Ilgaz and Adanir (2020), there was no statistically significant difference in students' performance in online and traditional exams. Cahapay (2020) states that online or blended learning will become a common part of the curriculum.

The above studies often deal more with the form of teaching, not so much with the form of testing. Our article focuses primarily on the form of testing. Its aim is to compare the results of tests conducted in the standard way with two alternative forms. At the University of Finance and Administration, mathematics is taught as a two-semester course in the first year of study, Mathematics 1 in the winter semester and Mathematics 2 in the summer semester, both of which culminate in an exam. At the time of the Covid-19 pandemic, examinations were conducted by alternative means. The first aim of this paper is to assess whether these alternative forms were an adequate substitute for standard examination and whether the reduced ability to check the authorship of the test did not lead to better results to a greater extent. The summer semesters of three consecutive academic years, when the tests were administered in a different form each time, were chosen to compare the results.

Mathematics examinations are conducted standardly face-to-face at the University of Finance and Administration. In the summer semester of 2019/2020, examinations could not be held in a standard way due to the situation caused by the Covid-19 pandemic, specifically the closure of schools. However, sufficient conditions have not been created to allow the online exams to be conducted in a form that best matches the standard format. Therefore, an "emergency" non-traditional form of testing was chosen: students independently worked out tasks corresponding to the tasks from the regular exam test and sent the finished tasks for evaluation. In the academic year 2020/2021, examinations in both semesters were conducted exclusively online due to the ongoing pandemic.

This article compares the results of mathematics exams in the summer semester 2018/2019, when the standard form was used, with the results in the summer semester 2019/2020, when the correspondence form was used. Further, the results of the standard form in the summer semester 2018/2019 are compared with those of the online form in the summer semester 2020/2021. The results of all three forms, as well as of both alternative forms, are not compared. The main objective is to determine whether the correspondence and online forms are adequate substitutes for the standard form of the examination since the alternative forms do not allow sufficient checking to verify that students are working independently.

In 2021/2022, the academic process returned almost to normal mode. Teaching and examinations were conducted in the standard way. The authors of this article were interested in whether the period of the Covid-19 pandemic had an effect on the exam results. Therefore, after the end of the examination period of the summer semester of the academic year 2021/2022, the mathematics examination results of this semester were compared with the examination results of each of the summer semesters of 2019-2021, i.e., with the period of the last academic year before the Covid-19, when the examination was conducted in the standard form, and the period with Covid-19, when the examinations were conducted in two different alternative forms.

The article extends the paper Ulrychová, Majovská and Tesař (2022) presented at the 19th International Conference on Efficiency and Responsibility in Education (ERIE 2022).

The article is structured as follows. The Materials and Methods section describes the various forms of testing in more detail and specifies the data and statistical methods used. The section Results provides the summary of our research outcomes. The Discussion section compares the results with those of other studies and considers possible reasons for the results. The Conclusion section provides an overall summary of the findings.

MATERIALS AND METHODS

Alternative Teaching and Three Different Forms of Testing

At the University of Finance and Administration, mathematics lessons (lectures and seminars) proceeded in the standard way until the beginning of the summer semester of 2019/2020, when the schools were closed due to the Covid-19 pandemic. Students were then referred to several weeks of guided self-study, for which special study support materials were promptly created. Later, online streaming of standard lectures from empty classrooms without students was enabled. Streaming had the advantage over standard teaching in that recordings were made of the lectures, which students could replay as needed. The disadvantage, however, was that there were no seminars in which students could better understand the material and practice the necessary computational procedures.

In the academic year 2020/2021, both lectures and seminars were conducted online only. The lectures were recorded, as in 2019/2020; the seminars were not recorded. While the lectures were fully comparable to the standard ones, the seminars were more problematic, mainly due to the difficulty of checking the students' work.

In the academic year 2021/2022, teaching and examination returned to their standard form before the pandemic. In addition, learning was supported by providing students with lectures recorded, as during the pandemic.

The curriculum was the same in all considered academic years, regardless of the form of teaching. The exam tests had the same structure in all three forms of examinations (standard, correspondent, and online), which are described in more detail below. They consisted of tasks to calculate exercises from the field of mathematical analysis (behaviour of a function of one variable, Taylor polynomial, an indefinite and definite integral, derivative of a function of two variables). In the standard form of examination, students write the test in a classroom under the direct supervision of examiners. Before the final submission of the tests, the examiners quickly check each student's test, approve the correctly solved tasks, and allow the student to correct the remaining tasks (without the examiner specifying the errors in any way). The test consists of 10 items; to pass the exam, the student must solve at least 50% of them.

In the correspondence form, students prepared written work consisting of exercises corresponding to the standard exam test. However, each student first had to create his or her individual assignment according to precise instructions, which significantly limited the possibility of transferring calculations and results between students. Students could submit their finished work to the university's information system at any time during the exam period; no exam dates were announced. Therefore, this form of the exam gave students a long time to prepare, and students could continuously consult their work with anyone (including teachers). Therefore, the assessment was more rigorous than the standard form of testing. To pass the exam, students had to correctly solve all ten tasks (which corresponded to the regular test). In order to get a better grade, it was necessary to do extra exercises on topics that were not part of the standard test (but were taught in the lessons).

In the online form of the exam, students took the test remotely on the given exam date. Students had to keep the camera and microphone on throughout the exam, and the examiner supervised the exam's correct procedure. The structure of the test and the assessment criteria were the same as for the standard exam. However, compared to the standard exam, students did not have the advantage of a preliminary check of the test by the examiner before the final submission. After completing the test, students photographed or scanned the test and immediately uploaded it to the school information system in the prescribed format. The time to complete and submit the test was extended by ten minutes compared to the standard testing format to make a copy of the test and upload it to the information system. In the case of technical problems during the exam or when submitting the paper, students could promptly solve them with the examiner.

In our study, we purposely worked exclusively with students in the mathematics subject who had the same teachers, examiners, and examination and assessment system. We did not want to compare the results of mathematics exams with results in other subjects, where even the standard way of examination could be quite different. In addition, the demands on students and the rigour of assessment may vary from examiner to examiner, which may bias comparisons of examination results. In this sense, the mathematics subject examiners considered in this study are consistent, so the examination results are well comparable.

Methodology and Research Organization

We conducted our research in two parts. In the first part, we compared the results of the standard testing from the period immediately before the pandemic (academic year 2018/2019) and alternative forms of testing during the pandemic (academic years 2019/2020 and 2020/2021). We wanted to determine whether the alternative testing methods were adequate. In the second part of the research, we compared examination results after the return to the standard teaching and examination regime (academic year 2021/2022) with previous periods (academic years 2018/2019, 2019/2020, 2020/2021). Moreover, we compared the exam results separately for full-time students and for part-time students in the academic year 2021/2022 with the previous years.

We only processed data from the summer semesters of the mentioned academic years, so in the following tables, we denoted the summer semester of the academic year 2018/2019 briefly as S_2019, in other cases analogously. We had 101 results from the standard exam in S_2019, 96 from the correspondence exam in S_2020, 111 from the online exam in S_2021, and 72 from the standard exam in S_2022. The test results were converted to the numerical value: Excellent (A) = 1, Excellent minus (B) = 2, Very good (C) = 3, Very good minus (D) = 4, Good (E) = 5, Failed (F) = 6. For each of the examined groups S_2019 to S_2022, we calculated the sample average value (M), the unbiased estimate for variance (V), skewness (S), and kurtosis (K). The number of tests in a group was denoted by N. We calculated the same parameters separately for full-time students and part-time students. In other cases, we denoted these groups as F_S_2019 (full-time study) and P_S_2019 (part-time study) analogously.

We used the Bowman-Shenton skewness and kurtosis test for normality at a 0.05 significance level (Bowman and Shenton, 1975). In all cases, we compared two groups (hereafter generally referred to as Group_1 and Group_2 in the tables). Specifically, we compared S 2019 (Group_1) sequentially with S_2020 and S_2021 (always Group_2) and then S_2022 (Group_1) sequentially with S_2019, S_2020, S_2021 (always Group_2). Furthermore, F_S_2022 (Group_1) was compared successively with F_S_2019, F_S 2020, F_S 2021 (always Group_2) and P_S_2022 (Group_1) successively with P_S_2019, P_S_2020, P_S_2021 (always Group_2).

Based on our experience and assumptions, we formulated hypotheses:

- H_0: Distribution from which the test results of Group_1 and Group_2 come is the same.
- A_0: Distribution from which the test results of Group_1 and Group_2 come is not the same.

All the hypotheses were tested at the 0.05 significance level using the non-parametric Mann-Whitney U test (Mann and Whitney, 1947).

RESULTS

Table 1 shows the relevant data and calculated values for each group S_2019 to S_2022.

Group	*N*-number	*M*-average	*V*-variance	*S*-skewness	*K*-kurtosis	Normality
S_2019	101	3.545	2.512	-0.074	1.804	No
S_2020	96	3.688	2.765	-0.283	1.837	No
S_2021	111	4.135	2.936	-0.516	2.029	No
S_2022	72	2.944	2.900	0.346	1.797	No

Table 1: Examined data for 2019 - 2022 (source: own)

Group	*N*-number	*M*-average	*V*-variance	*S*-skewness	*K*-kurtosis	Normality
F_S_2019	57	3.509	2.5506	-0.139	1.691	No
F_S_2020	58	3.672	2.6803	-0.237	1.806	No
F_S_2021	85	4.271	2.9141	-0.644	2.197	No
F_S_2022	51	2.922	3.0740	0.368	1.812	No
P_S_2019	44	3.591	2.5739	0.005	1.923	Yes
P_S_2020	38	3.711	2.9681	-0.347	1.874	Yes
P_S_2021	26	3.692	2.8624	-0.162	1.816	Yes
P_S_2022	21	3.000	2.6013	0.293	1.709	Yes

Table 2: Examined data for 2019 - 2022 full-time and part-time students (source: own)

Group_1 / Group_2	Z	*p*-value	Cohen's *d*
S_2019 / S_2020	-0.693	0.488	0.088
F_S_2022 / F_S_2019	-1.803	0.071	0.350
S_2019 / S_2021	-2.686	0.007	0.357
S_2022 / S_2019	-2.347	0.019	0.365
P_S_2022 / P_S_2019	-1.354	0.176	0.367
P_S_2022 / P_S_2021	-1.284	0.199	0.419
P_S_2022 / P_S_2020	-1.449	0.147	0.426
S_2022 / S_2020	-2.759	0.006	0.442
F_S_2022 / F_S_2020	-2.361	0.018	0.442
S_2022 / S_2021	-4.312	0.000	0.697
F_S_2022 / F_S_2021	-4.070	0.000	0.780

Table 3: Results of Mann-Whitney test and Cohen's *d* (source: own)

Table 2 shows the relevant data and calculated values for each group F_S 2019 to F_S_2022 and P_S 2019 to P_S_2022.

Table 3 compares the pairs of groups as described above. It contains results of Mann-Whitney U tests, with the *p*-value estimated by normal approximation with a continuity correction (*Z*). Cohen's *d* was also calculated (Cohen, 1988). The rows of the table are ordered by the size of Cohen's *d*. Cohen's *d* indicates the size of the difference between the groups in terms of standard deviation units.

- A small effect size is usually around 0.2.
- A medium effect size is around 0.5.
- A large effect size is around 0.8 or higher.

In our case, the pairs in the first to fifth rows of Table 3 can be classified as small differences according to Cohen's *d*. The pairs in the sixth to ninth rows have a medium difference according to Cohen's *d*. The last two pairs, representing the results in 2021 (online exam) and the results in 2022 (standard exam), are significantly different. Surprisingly, the results in 2019 (S_2019, standard exam) and 2020 (S_2020, correspondence exam) are almost identical. Possible procedural reasons for these results are discussed in the following section.

DISCUSSION

As mentioned in the introduction, several factors can influence the results of examinations. It is difficult to determine to what extent the teaching method at the University of Finance and Administration contributes to the results. We do not have the opportunity to assess this sufficiently, as a change in the form of testing accompanied the change in the form of teaching. However, in accordance with the articles cited in the introduction, it can be concluded that the online form of teaching mathematics at the University of Finance and Administration may not have had a major impact on the examination results, as online lectures, in particular, correspond well to standard teaching. It is not a big problem to transmit information online (via camera or graphic tablet) from the teacher to the students and answer any questions. In this respect, no major change has occurred, especially in the part-time form of study, with no seminars.

The seminars present a problem because, in the standard form of teaching, students, under the teacher's guidance, solve examples on the blackboard and independently at their desks. In seminars conducted online, checking student work (calculations on paper) was complicated. In the standard form of teaching, students soon realize that the seminars are about student-teacher collaboration, not teacher rehearsal, and are not afraid to show their ignorance. In the academic year 2020/2021, when teaching was entirely online and students did not come into personal contact with the teacher at all, it was

often difficult to establish such cooperation. Students who were interested in mastering the curriculum had the opportunity to have the problematic parts explained to them, just like in the standard form of study, but many tried to hide their ignorance and avoid contact with the teacher. However, a great advantage compared to the past was the recordings of the lectures, which students could replay as needed.

Overall, it appears that students may have been as well prepared for the exam through the online form of instruction as through the standard form. In this paper, however, we consider how exam results may have been influenced by the form of testing, not the form of teaching. In the first phase, we compared the results of standard testing from the period immediately before the pandemic (academic year 2018/2019) with alternative forms of testing during the pandemic (academic years 2019/2020 and 2020/2021) to determine whether the use of alternative testing methods was adequate under the circumstances. Statistical evaluation shows that online testing results are worse than the standard form. The consistency of results in the correspondence and standard forms of testing is rather surprising.

In the correspondence form of the examination, students had the advantage of being allowed to work on the tests for a few months and had the opportunity to consult the tasks with anyone or have them checked before submission. The examiners were concerned that many students would thus obtain an excellent mark in the exam, regardless of their actual knowledge. On the other hand, the students had a more difficult situation with creating their own tasks, especially with the tougher test assessment. It turned out that many students were satisfied with completing the compulsory part of the test (and not always getting it right) – they did not attempt the extra part for a better grade. This may have been because the extra part consisted of tasks that were the content of the unpresented material, where students were referred to self-study, while the compulsory part was largely taught before the schools closed. However, students may have considered it sufficient to pass the exam regardless of the grade. The fact that the students did not have the same assignment may also have played a role. The concern that, in many cases, someone else authored the test instead of the student was not confirmed. In case of doubts about the authorship of the test, students were asked to take an individual online examination, but these cases were quite rare.

In the online form of testing, students were disadvantaged by the fact that they had not experienced face-to-face teaching at all since the beginning of their studies at the university; teaching was only online. Online lectures, transmitted via a camera or a graphic tablet, were an adequate substitute for face-to-face lectures, and students could replay the recording. However, online learning and computer-based testing may not suit all students. Kemp and Grieve (2014) found that students preferred face-to-face rather than online activities, but there was no significant difference in their test performance on the two alternatives. Boevé et al. (2015) concluded that computer-based exam total scores were similar to paper-based exam scores, but only about a quarter of students preferred a computer-based exam. Mendoza et al. (2021) argue that students' increasing anxiety during the pandemic has significantly affected their performance and that the transition to distance learning led to significant differences in students' understanding of mathematical concepts. Fejfar, Jadrná, and Fejfarová (2021), and Dvořáková et al. (2021) assess the advantages and disadvantages of distance education from the student's perspective.

The online form of the examination may have been more stressful for many students than the standard form of testing. In addition to the fact that they may not have been comfortable being watched by a camera, students may also have been nervous about the potential failure of technology, either during the exam or in copying, correctly formatting, and saving the finished test. Nervousness may have negatively affected the outcome of the exam. Elsalem et al. (2020) report that a third of students find online exams and the associated technical problems more stressful than standard exams. Furthermore, a not insignificant number of cheating attempts were detected. In this case, the result of the test was directly assessed as insufficient. When in doubt, students had to take an individual online examination. A significant factor may be that, unlike in the standard exam, students did not have the opportunity to correct some tasks after the examiner had previously checked them.

Among other studies comparing the results of standard and online exams, we mention in particular those conducted at schools whose focus and mathematics curriculum correspond to the University of Finance and Administration. For example, Klůfa (2021) compares the results of the oral part of the mathematics examination at the University of Economics in Prague, Otavová and Sýkorová (2021) from the same university compare the results of midterm tests, final tests, and final grades depending on the form of teaching and examination. Unlike the results presented in our paper, the online form of testing led to better results than the standard form. In the online form of testing, however, it depends very much on the conditions set; these are not specified in the above articles. For example, if the online exam is not monitored by cameras and students only upload a completed test, the results may be highly distorted due to the possibility of cheating.

The problems associated with cheating in online exams have received considerable attention, regarding proctoring during the exam (Atoum et al., 2017) and subsequent detection (D'Souza et al., 2017). Detecting cheating, however, often involves finding matches with other texts. In the mathematics exam, students work out problems using paper and pencil; checking the independence of their work is more difficult.

The results of research by Moravec, Ječmínek and Kukalová (2022) from the University of Life Sciences in Prague also show that the chances of passing the exam are higher with the online testing compared to the standard face-to-face form. The authors consider the higher success rate in online testing to result from online courses being more effective than traditional face-to-face courses, for example, Elfaki, Abdulraheem and Abdulrahim (2019). However, the authors acknowledge that the results of online testing may be biased due to the impaired ability to ensure cheating-free conditions; in designing the tests, they focused on preventing cheating rather than detecting it. Examination results from different classes were studied, but

multiple-choice tests were mostly used, which are not suitable for mathematics examinations. This may also be the reason for the different conclusions presented in our article compared to the articles mentioned above.

In the second phase, we compared examination results after the return to normal teaching and examination mode (academic year 2021/2022) with previous periods. The average grade in the summer semester 2021/2022 was found to be better than in any of the previous terms examined. Particularly in comparison to the year 2020/2021, when the exams were conducted online, the disadvantages mentioned above of this form may have played a role, especially the impossibility of pre-checking the test. During the summer semester 2021/2022 exam period, the examiners were surprised by the small number of students who failed the exam (grade "F"). In this period, however, a high proportion of students were registered to study the subject Mathematics 2 but did not appear for the exam at all (notation "-"). These were probably mostly students who were aware that they had not mastered the subject matter and would likely fail the exam. It is possible that the fact that probably underprepared students did not take the exam at all contributed to a better average grade.

It turned out that the percentages of students who did not appear for the exam were similar in 2018/2019 (47%) and 2021/2022 (49%), when the exams were conducted in the standard way in both years, as well as in 2019/2020 (35%) and 2020/2021 (35%) when the exams were conducted in alternative forms. In 2019/2020 and 2020/2021, i.e., using alternative forms of testing, the proportion of these students was lower than in 2018/2019 and 2021/2022. This confirms what examiners have already observed during the examination process, namely when alternative forms of testing are employed, students are more likely to attempt to pass the exam even if they are not adequately prepared. This may be due to students feeling embarrassed about displaying their lack of knowledge in a face-to-face setting with the examiner. Additionally, students may be hopeful that cheating will enable them to pass the exam through alternative forms, as demonstrated by Harmon and Lambrinos (2008). In fact, detected attempts at cheating, particularly in the online form of testing, were more frequent than in the standard form of testing. While the greater difficulty of detecting such attempts is a disadvantage of alternative forms of testing, the pandemic has fortunately not resulted in students being able to pass exams on a large scale by cheating.

Although the proportions of students who did not appear for the exam were similar in 2018/2019 and 2021/2022 (the standard examination form in both academic years), insufficient mastery of secondary school content may have also contributed to this in 2021/2022. This year followed the school year during which teaching was affected by the Covid-19 pandemic. At the University of Finance and Administration, mathematics is taught in the first year of study, building directly on secondary school mathematics. This is particularly evident in the summer semester when there is a greater need to use secondary school mathematical tools compared to the winter semester. During the academic year 2021/2022, many students complained about inadequate preparation in secondary school mathematics during the pandemic, when classes were conducted online without prior experience. Consequently, some students gave up studying mathematics at the university and did not even attend classes. This likely also contributed to the number of students who did not attempt the exam at the end of the summer semester and, thus probably, to the better average.

In this respect, the situation might be different for full-time versus part-time forms of study: students of the part-time form usually have a longer time gap since graduating from secondary school, and the problems with teaching during the pandemic mostly did not affect them. Therefore, the exam results from 2021/2022 were compared with previous years for full-time and for part-time students separately. For students of the part-time form, the results from 2021/2022 were consistent with all previous years; thus, the pandemic period did not affect the results for this form of study. In contrast, full-time students performed better in 2021/2022 compared to the pandemic period (with alternative forms of testing), probably because unprepared students did not show up for the exam at all.

Compared to 2018/2019, there was also agreement for the full-time form, as opposed to comparing the overall results without differentiating the form of study. If we were working at the 0.01 significance level, then even for the pair 2021/2022 and 2018/2019, the hypothesis H_0 would not be rejected. What is remarkable about the 2021/2022 exam results, in addition to the best average and low proportion of students failing the exam (grade "F": 8% in 2021/2022, 11% in 2018/2019, 14% in 2019/2020, 31% in 2020/2021), is the high proportion of top grades (grade "A": 29% in 2021/2022, 12% in 2018/2019, 16% in 2019/2020, 12% in 2020/2021). In the academic year 2021/2022, the number of "A" grades is the highest of all grades, unlike all previous periods, including the period with the standard form of examination.

The cause must, therefore, be sought elsewhere than in the form of testing. Such a good result could be due to the fact that students had – in contrast to the standard form of examination in 2018/2019 – recordings of lectures, including repetitions for the exam. According to the students' opinion, this helped them a lot in preparing for the exam. However, recordings of all lectures were also available in 2020/2021, but students did not have the opportunity to pre-check the test online as in the standard form. It is also possible that students approached their studies and exam preparation with more vigour after the tiredness and frustration during the pandemic.

Further research could focus more significantly on issues of potential cheating related to the use of ICT in testing. Here, it is necessary to start with the latest cybersecurity research, for example, according to Rahmani et al. (2021). However, there is a need to focus specifically on the conditions that are suitable for testing in mathematics. Multiple-choice tests are not appropriate, and due to the use of specific mathematical symbols, elaboration on a computer is not appropriate either. It is not possible to use automatic correction and scoring as it is possible in some other subjects (Böhmer et al., 2018). It is neither about creating tests suitable for such use (Ardid et al., 2015) nor about automated online exam proctoring (Atoum et al., 2017). The issue is how to achieve the best possible control in a situation where students work out the tasks classically using pencil and paper and immediately upload a copy of

the final product to the university information system, all without the personal supervision of the examiner.

There are many challenges and opportunities ahead for teachers. Malakeh et al. (2022) review and summarise research examining the impact of the pandemic on online examination globally. They highlight challenges and opportunities for policymakers, educators, researchers, and higher education decision-makers regarding online examinations. Teachers should be prepared for online teaching, but especially online testing, which brings more problems (Kyungmee and Fanguy, 2022). Khan et al. (2021) point to the fact that it is important in further research to pay attention to students' opinions about online testing, which can negatively affect the result of the test, for example, due to anxiety and stress. If online testing will continue to be used, teachers should have better technical facilities at their disposal (Abdelwahed, 2023). For example, online proctoring should help teachers in the future. It has become a necessity in online teaching (Waheeb, 2022).

CONCLUSIONS

The first objective of this paper was to determine whether the alternative forms (correspondence and online) of written examinations in mathematics, applied at the University of Finance and Administration during the Covid-19 pandemic, were a suitable substitute for the standard mode of examination. It turned out that the results in the correspondence form of testing, where students wrote the test almost without time limits and stress, were surprisingly very similar to the results in the standard form of testing. The results in the online form were worse than in the standard form. Possible reasons for these results are analysed in the Discussion section.

While alternative forms of testing are less likely to detect potential cheating, the results presented in this paper suggest that alternative methods did not lead to a higher proportion of completely unprepared students passing the exam. Both alternative forms of testing, as they were conducted at the University of Finance and Administration, can, therefore, be considered acceptable substitutes for standard testing in this regard. Although the correspondence form matched the results of the standard examination better than the online form, the correspondence form can only be considered an emergency solution because of the very low possibility of checking for cheating. However, if it is not possible to test in the standard way, the online form of testing is acceptable.

The second objective was to assess the results of the examinations after the return to the normal mode of teaching and the standard method of examination. Our research revealed that during this period, more students did not attempt to pass the exam at all. On the other hand, among the students who did pass the exam, there was a significant preponderance of those who were excellently prepared. The number of students who failed the exam was minimal. It appears that students adopted a more responsible approach to taking exams after returning to normal – those who were unprepared did not try to take the exam, while those who were prepared did so excellently. Although adopting hybrid teaching is likely inevitable in the future and beneficial for many subjects, the standard face-to-face form seems to be the most suitable for the written examination in mathematics.

REFERENCES

Abdelwahed, N. and Soomro, B. (2023) 'Attitudes and intentions towards the adoption of mobile learning during COVID-19: building an exciting career through vocational education', *Education and Training*, Vol. 65, No. 2, pp. 210–231. https://doi.org/10.1108/ET-02-2022-0048

Ardid, M., Gómez-Tejedor, J. A., Meseguer-Dueñas, J. M., Riera, J. and Vidaurre, A. (2015) 'Online exams for blended assessment. Study of different application methodologies', *Computers & Education,* Vol. 81, pp. 296–303. https://doi.org/10.1016/j.compedu.2014.10.010

Aristovnik, A., Keržič, D., Ravšelj, D., Tomaževič, N. and Umek, L. (2020) 'Impacts of the COVID-19 Pandemic on Life of Higher Education Students: A Global Perspective', *Sustainability*, Vol. 12, No. 20, p. 8438. https://doi.org/10.3390/su12208438

Atoum, Y., Chen, L., Liu, A. X., Hsu, S. D. and Liu, X. (2017) 'Automated online exam proctoring', *IEEE Transactions on Multimedia,* Vol. 19, No. 7, pp. 1609– 1624.

Boevé, A. J., Meijer, R., Albers, C., Beetsma, Y. and Bosker, R. (2015) 'Introducing computer-based testing in high-stakes exams in Higher Education: Results of a field experiment', *PLoS One,* Vol. 10, No. 12, p. e0143616. https://doi.org/10.1371/journal.pone.0143616

Bowman, K. O. and Shenton, L. R. (1975) 'Omnibus test contours for departures from normality based on √b1 and b2', *Biometrik*a, Vol. 62, No. 2, pp. 243–250. https://doi.org/10.1093/biomet/62.2.243

Böhmer, C., Feldmann, N. and Ibsen, M. (2018) 'E-exams in engineering education — online testing of engineering competencies: Experiences and lessons learned', *Proceedings of 2018th IEEE Global Engineering Education Conference (EDUCON)*, Santa Cruz de Tenerife, pp. 571–576. https://doi.org/10.1109/EDUCON.2018.8363281

Cahapay, M. B. (2020) 'Rethinking Education in the New Normal Post-COVID-19 Era: A Curriculum Studies Perspective', *Aquademia*, Vol. 4, No. 2. https://doi.org/10.29333/aquademia/8315

Cohen, J. (1988) *Statistical Power Analysis for the Behavioral Sciences*, 2nd Edition, Mahwah: Lawrence Erlbaum Associates.

Darolia, R. (2014) 'Working (and studying) day and night: Heterogeneous effects of working on the academic performance of full-time and part-time students', *Economics of Education Review*, Vol. 38, No. 1, pp. 38–50. https://doi.org/10.1016/j.econedurev.2013.10.004

D'Souza, K. A. and Siegfeldt, D. V. (2017) 'A Conceptual Framework for Detecting Cheating in Online and Take-Home Exams', *Decision Sciences Journal of Innovative Education,* Vol. 15, No. 4, pp. 370–391. https://doi.org/10.1111/dsji.12140

Dvořáková, K., Emmer, J., Janktová, R. and Klementová K. (2021) 'From F2F to ERT: University Students' Perception of Remote Learning During the First COVID-19 Lockdown', *Journal on Efficiency and Responsibility in Education and Science*, Vol. 14, No. 2, pp. 89–100. http://dx.doi.org/10.7160/eriesj.2021.140203

Elfaki, N. K., Abdulraheem, I. and Abdulrahim, R. (2019) 'Impact of E-Learning vs Traditional Learning on Student's Performance and Attitude', *International Journal of Medical Research & Health Sciences*, Vol. 8, No. 10, pp. 76–82.

Elsalem, L., Al-Azzam, N., Jum'ah, A. A., Obeidat, N., Sindiani, A. M. and Kheirallah, K. A. (2020) 'Stress and behavioral changes with remote E-exams during the Covid-19 pandemic: A cross-sectional study among undergraduates of medical sciences', *Annals of Medicine and Surgery*, Vol. 60, pp. 271–279. https://doi.org/10.1016/j.amsu.2020.10.058

Fajčíková, A., Fejfarová, M., Hlavsa, T. and Fejfar, J. (2019) 'Success Rate in Examinations: Does the Degree Programme Matter?', *Proceedings of the 16th International Conference on Efficiency and Responsibility in Education (ERIE 2019)*, Prague, pp. 43–49.

Fejfar, J., Jadrná, M. and Fejfarová, M. (2021) 'Covid-19: Distance Learning at FEM CZU Prague from the Perspective of Students (Spring 2020)', *Proceedings of the 18th International Conference on Efficiency and Responsibility in Education (ERIE 2021)*, Prague, pp. 22–28.

Harmon, O. R. and Lambrinos, J. (2008) 'Are online exams an invitation to cheat?', *Journal of Economic Education*, Vol. 39, No. 2, pp. 116–125. https://doi.org/10.3200/ JECE.39.2.116-125

Hartnett, M., Butler, P. and Rawlins, P. (2023) 'Online proctored exams and digital inequalities during the pandemic', *Journal of Computer Assisted Learning*, Vol. 53, No. 3, pp. 475–490. http://dx.doi.org/10.1111/jcal.12813

Haus, G., Pasquinelli, Y., Scaccia, D. and Scarabottolo, N. (2020) 'Online written exams during Covid-19 crisis', *Proceedings of International Conference e-Learning 2020, Japan*, pp. 79–86.

Ho, N. T. T., Sivapalan, S., Pham, H. H., Nguyen, L. T. M., Pham, A. T. V. and Dinh, H. V. (2021) 'Students' adoption of e-learning in emergency situation: the case of a Vietnamese university during COVID-19', *Interactive Technology and Smart Education*, Vol. 18, No. 3, pp. 246–269. https://doi.org/10.1108/ITSE-08-2020- 0164

Hvorecký, J., Beňo, M., Ferenčíková, S., Janoščová, R. and Šimúth, J. (2021) 'Czech and Slovak Educators' Online Teaching Experience: A Covid-19 Case Study', *Acta Informatica Pragensia*, Vol. 10, No. 3, pp. 236–256. http://dx.doi.org/10.18267/j.aip.162

Ilgaz, H. and Adanir, G. A. (2020) 'Providing online exams for online learners: Does it really matter for them?', *Education and Information Technologies*, Vol. 25, No. 2, pp. 1255–1269. https://doi.org/10.1007/s10639-019-10020-6

Itani, Malakeh, Itani, Maher, Kaddoura, S. and Al Husseiny, F. (2022) 'The impact of the Covid-19 pandemic on online examination: challenges and opportunities', *Global Journal of Engineering Education*, Vol. 24, No. 2, pp. 105–120.

Joyce, T., Crockett, S., Jaeger, D. A. (2015) 'Does classroom time matter?', *Economics of Education Review*, Vol. 46, pp. 64–77. https://doi.org/10.1016/j.econedurev.2015.02.007

Kemp, N. and Grieve, R. (2014) 'Face-to-face or face-to-screen? Undergraduates' opinions and test performance in classroom vs. online learning', *Frontiers in psychology*, Vol. 5. https://doi.org/10.3389/fpsyg.2014.01278

Khan, M. A., Vivek, V., Khojah, M., Nabi, M. K., Paul, M. and Minhaj, S. (2021) 'Learners' Perspective towards E-Exams during COVID-19 Outbreak: Evidence from Higher Educational Institutions of India and Saudi Arabia', *International Journal of Environmental Research and Public Health*, Vol. 18, No. 12, 6534. https://doi.org/10.3390/ijerph18126534

Klůfa, J. (2021) 'Evaluation of Oral Exams in Mathematics in Covid Time', *Proceedings of the 18th International Conference on Efficiency and Responsibility in Education (ERIE 2021)*, Prague, pp. 50–56.

Kyungmee, L. and Fanguy, M. (2022) 'Online exam proctoring technologies: Educational innovation or deterioration?', *British Journal of Educational Technology*, Vol. 53, No. 3, pp. 475–490. https://doi.org/10.1111/bjet.13182

Majovska, R. (2015) 'Open Source Programmes as the Support in Mathematical Education', *Proceedings of the 12th International Conference on Efficiency and Responsibility in Education (ERIE 2015)*, Prague, pp. 341–347.

Makamure, C. and Tsakeni, M. (2020) 'Covid-19 as an Agent of Change in Teaching and Learning STEM Subjects', *Journal of Baltic Science Education*, Vol. 19, No. 6A, pp. 1078–1091. https://doi.org/10.33225/jbse/20.19.1078

Mann, H. B. and Whitney, D. R. (1947) 'On a Test of Whether one of Two Random Variables is Stochastically Larger than the Other', *The Annals of Mathematical Statistics,* Vol. 18, No. 1, pp. 50–60.

Mendoza, D., Cejas, M., Rivas, G. and Varguillas, C. (2021) 'Anxiety as a prevailing factor of performance of university mathematics students during the COVID-19 pandemic', *The Education and Science Journal*, Vol. 23, No. 2, pp. 94–113. https://doi.org/10.17853/1994-5639-2021-2-94-113

Moravec L., Ječmínek J. and Kukalová G. (2022) 'Evaluation of final examination performance at Czech University of Life Sciences during the COVID-19 outbreak', *Journal on Efficiency and Responsibility in Education and Science*, Vol. 15, No. 1, pp. 47–52. http://dx.doi.org/10.7160/eriesj.2022.150105

Otavová, M. and Sýkorová, I. (2021) 'Comparison of Student Results before and during Covid-19 Pandemic', *Proceedings of the 18th International Conference on Efficiency and Responsibility in Education (ERIE 2021)*, Prague, pp. 98–104.

Pasáčková, J. (2021) 'Evaluation of Test Results during online and Face-to-Face Learning', *Proceedings of the 18th International Conference on Efficiency and Responsibility in Education (ERIE 2021)*, Prague, pp. 105–111.

Pokorny, M. (2021) 'Video Lessons and E-learning Can Overcome Ban of Face-to-face Lessons in Teaching Mathematics', *Proceedings of International Symposium on Educational Technology (ISET 2021)*, Japan, pp. 44–47. https://doi.org/10.1109/ISET52350.2021.00019

Rahmani, A. M., Mohammadi, M., Rashidi, S., Lánský, J., Mildeová, S., Safkhani, M., Kumari, S. and Hosseinzadeh, M. (2021) 'Questioning the Security of Three Recent Authentication and Key Agreement Protocols', *IEEE Access*, Vol. 9, pp. 98204–98217. http://dx.doi.org/10.1109/ACCESS.2021.3094875

Sun, P.-C., Tsai, R. J., Finger, G., Chen, Y.-Y. and Yeh, D. (2008) 'What drives a successful e-Learning? An empirical investigation of the critical factors influencing learner satisfaction', *Computers & Education*, Vol. 50, No. 4, pp. 1183–1202. https://doi.org/10.1016/j.compedu.2006.11.007

Thompson, V. L. and McDowell, Y. L. (2019) 'A Case Study Comparing Student Experiences and Success in an Undergraduate Mathematics Course offered through Online, Blended, and Face-to-Face Instruction', *International Journal of Education in Mathematics, Science and Technology (IJEMST)*, Vol. 7, No. 2, pp. 116–136. http://dx.doi.org/10.18404/ijemst.552411

Ulrychová, E. and Bílková, D. (2018) 'Gender Analysis of Math Exam Results at University of Finance and Administration', *Proceedings of the 12th International Days of Statistics and Economics (MSED)*, Prague, pp. 1815–1822.

Ulrychová, E., Majovská, R. and Tesař, P. (2022) 'Standard Form vs. Two Alternative Forms of Written Examinations in Mathematics', *Proceedings of the 19th International Conference on Efficiency and Responsibility in Education (ERIE 2022)*, Prague, pp. 165–176.

Waheeb, Y., Manoranjan, M. and Basem, S. (2022) 'Privacy-Preserving Online Proctoring using Image-Hashing Anomaly Detection', *Proceedings of EEE International Wireless Communications and Mobile Computing (IWCMC 2022)*, Dubrovnik, pp. 1113–1118, https://doi.org/10.1109/IWCMC55113.2022.9825119

FACTOR ANALYSIS ON THE MOTIVATION FOR EXTENSIVE READING QUESTIONNAIRE

Helta Anggia[1, 2]✉
Anita Habók[3, 4, 5]

[1]Doctoral School of Education, University of Szeged, Hungary

[2]Universitas Bandar Lampung, Bandar Lampung, Indonesia

[3]Institute of Education, University of Szeged, Hungary

[4]MTA-SZTE Digital Learning Technologies Research Group, University of Szeged, Hungary

[5]Digital Learning Technologies Incubation Research Group, University of Szeged, Hungary

✉ helta.anggia@edu.u-szeged.hu

Article history
Received
February 9, 2023
Received in revised form
June 14, 2023
Accepted
May 20, 2024
Available on-line
September 30, 2024

ABSTRACT

This study examined the factors adapted from the Motivation for Reading Questionnaire. We considered eight dimensions (Self-Efficacy, Reading Challenge, Reading Curiosity, Reading Involvement, Importance of Reading, Recognition for Reading, Reading for Grades, and Social Reasons for Reading). In addition, we included some items based on the extensive reading, principles, and technology acceptance model. The study recruited 558 undergraduate students of English as a foreign language in Indonesia via Google Forms. The structure of the questionnaire was validated using exploratory and confirmatory factor analyses. To determine the dependability of the instrument, internal consistency reliabilities of the instrument as a whole and per factor were calculated. We computed the average variance extracted and the Heterotrait-Monotrait Ratio of Correlation to determine convergent and discriminant validities. The results led to the omission of six items with loading values < 0.50. The omissions included one item for Reading Involvement (0.42) and five items for Social Reasons for Reading (0.47; 0.43; n/s.; n/s.; and n/s.). Lastly, the study presented the significance of the results and directions for future studies.

KEYWORDS

Confirmatory factor analysis, exploratory factor analysis, extensive reading, motivation, reading

HOW TO CITE

Anggia H., Habók A. (2024) 'Factor Analysis on the Motivation for Extensive Reading Questionnaire', *Journal on Efficiency and Responsibility in Education and Science*, vol. 17, no. 3, pp. 195–204. http://dx.doi.org/10.7160/eriesj.2024.170302

Highlights

- *The exploratory factor analysis helps the authors to identify the factors of the adapted questionnaire and the relevant items to the context.*
- *The confirmatory factor analysis ensures the validity of the 44 items of the adapted questionnaire.*
- *The average variance extracted (AVE) analysis indicated an acceptable convergent validity since items of the same factors loaded significantly.*
- *The heterotrait-monotrait (HTMT) analysis indicated a significant discriminant validity since the factors did not overlap.*

INTRODUCTION

In recent decades, Extensive reading has been identified as one of the most effective strategies for motivating university students to read (Renandya, 2007). Although the primary concern of teaching extensive reading has been motivation for reading (Chanthap and Wasanasomsithi, 2019; Hagley, 2017; Hendriwanto and Kurniati, 2019; Rezaee and Farahian, 2020; Shurentsetseg, Nandintsesteg and Nyamsuren, 2015), no instrument that assesses motivation for reading following an extensive reading intervention has been explicitly constructed under the principles of extensive reading. One of the most prominent instruments for reading motivation is the motivation for reading questionnaire (MRQ) developed by Wigfield and Guthrie (1997). Despite being validated in primary schools and employing multidimensional factors to measure reading motivation, the construct of the instrument requires revision. It lacks evidence of the large reading program's effects. This is understandable because Wigfield and Guthrie's background on reading motivation is more broad than specialized, such as EFL reading motivation.

Additionally, as digital learning has grown in popularity, extensive reading has shifted to an online format that uses technology as a medium of instruction (Cote and Milliner, 2015; Matsuda, 2020). As a result, the use of technology has become inevitable to extensive reading programs. Therefore, one may infer that the current constructs of MRQ must be adjusted following the technology acceptance model (TAM) proposed by Davis et al. (1989), who established the potential

relationship between perceived ease of use and perceived usefulness of technology and one's motivation to use the technology in a learning process. Meanwhile, Day (2015) established the famous ten principles of extensive reading based on the motivational aspects of students in extensive reading. The principles contain the intrinsic and extrinsic motivation aspects that Deci and Ryan (2000) put forward. In Indonesia, Higher education institutions have begun to pay attention to extensive reading, particularly programs that focus on English language learning. Moreover, the Indonesian government has provided support for extensive reading through Gerakan Literasi Nasional (National Literacy Movement) (Anandari and Iswandari, 2019). Consequently, the current study proposed a recontextualization of the MRQ using TAM and extensive reading principles to establish a robust assessment of students' reading motivation after an extensive reading intervention in the Indonesian EFL context.

Specifically, the study aims to analyse the MRQ components in light of Day (2015) and Davis (1989) theories to develop and validate a new measure of university students' motivation for reading following an extensive reading intervention. The first step was that we described the context that prompted the design of a reconstructed measure of reading motivation, explained the questionnaire item modification and development phase, and subsequently validated the data and made justification.

THEORETICAL BACKGROUND

Motivation for EFL Reading

Since Deci and Ryan (1985) introduced the initial concept of self-determination theory (SDT), the theory of motivation has substantially advanced. Apart from the SDT scales and intrinsic and extrinsic motivation, which evolved into a meta-theory of motivation, certain areas of expansion in motivation theory included expectancy value (Wigfield and Eccles, 2000), task avoidance and procrastination theory (Ferrari et al., 1995), TAM (Davis, 1989), and self-related beliefs (Habók et al., 2020). These concepts are pervasive and are adopted in many areas of learning motivation today, such as motivation for reading, particularly EFL reading motivation (Mori, 2002; Takase, 2007; Kim, 2011; Protacio, 2012; Park, 2015; Dakhi, 2018; Pirih, 2019). Similarly, this study integrated most of the theories mentioned above to rethink the construction of the MRQ (Wigfield and Guthrie, 1997) to create a more precise motivation for an extensive reading questionnaire following an extensive online reading intervention. We discussed the related theories as follows.

Initially, expectancy value theory may be considered crucial concerning how one's motivation is melded in relation to EFL reading (Wigfield and Eccles, 2000). Like many other subjects that involve one's ability-related beliefs, examining the influence of the competence beliefs of students, which have evolved over their school years, on whether or not to engage in positive reading behaviour during university is interesting. The findings that the competence belief of children decreases over school years may be logical now (Wigfield, 1994; Wolgast, 2018). Similarly, Tuominen et al. (2020) discovered that children who are transitioning from elementary to lower secondary schools experienced a stable positive achievement motivation. The authors also found that some students avoided tasks due to low competency. This finding reveals students' school-year competence beliefs. Long-term analysis of university students' EFL reading motivation will be fascinating because they vary in competence belief and task value representation. Subjective task values, such as the desire to perform well (attainment value), belief in future benefits (utility value), intrinsic motivation to do something (intrinsic value), and self-assessment of energy required to perform an activity (cost), play a significant role in determining one's future action (Wigfield and Eccles, 2000). The effect of the diverse expectancy values of university students on learning motivation, particularly in EFL reading, receives little attention and requires further studies.

Procrastination and task avoidance are also other facets of learning motivation that have received less scholarly attention. Ferrari et al. (1995) discussed the association among procrastination, task avoidance, and various factors such as perfectionism, low self-esteem, anxiety, achievement motivation, and intelligence. Although procrastination is intuitively associated with negative attitudes toward a particular task or assignment, Ferrari et al. (1995) found no positive association between procrastination and the majority of previously identified factors such as anxiety, negative achievement motivation, and low intelligence. In EFL reading, students' procrastination may be due to perfectionism in comprehending reading materials.

Another aspect of one's desire to engage in extensive online reading is the involvement of technology. Davis (1989) proposed that perceived ease of use and perceived usefulness of a particular technology influence the future actions of individuals. Furthermore, in his comprehensive elaboration, Davis proposed that beliefs and attitudes are co-determinants of action execution. Thus, the lack of ability-related beliefs in using a particular technology may decrease one's desire to read extensively via online technology.

The last aspect discussed is self-related beliefs. Habók et al. (2020) suggested that academic motivation mediates between self-related beliefs and academic achievement and vice versa. Academic motivation can boost self-esteem and academic achievement. Self-related ideas can motivate students to get good marks. This study examined how self-related beliefs affect online English reading. We examined how self-efficacy and self-concept, which construct self-related beliefs, influence academic motivation. Self-efficacy in online reading and reading self-concept are key to understanding self-related ideas influencing reading motivation. The subsequent discussion demonstrates the breadth of motivational theories applicable to reading motivation, particularly in EFL reading. We regard them as essential aspects of the basis of this study.

Motivation for Reading Questionnaire

Recent instruments used to assess motivation for reading are based on the Motivation for Reading Questionnaire (MRQ) by Wigfield and Guthrie (1997), who developed MRQ to predict the amount and breadth of reading for elementary school pupils in grades 3 to 5. Self-efficacy, intrinsic and extrinsic motivation and, learning goals, and social motivation were

discovered from questionnaire items. Self-efficacy is pupils' ability-related beliefs about their reading abilities. Learning goals integrate subjective task values, whereas intrinsic and extrinsic motivation refers to internal and external influences that push children to read. Children read because of social drive. Children read to socialize with friends and family.

Wigfield and Guthrie (1997) identified 11 dimensions behind the three core constructs of the MRQ. Reading efficacy (three items) and reading challenge (five items) are components of the self-efficacy concept. The first indicates the beliefs of students with respect to self-reported reading ability, whereas the second is intended to reveal the internal motivation to read challenging text. Reading Curiosity (six items) and Reading Involvement (six items) are then designed to elicit information about students' intrinsic motivations due to their interests. Afterward, the Importance of Reading (two items) reflects pupils' perceptions of the importance of reading for future benefits. Although Reading Work Avoidance (four items) elucidates the reading motivation of students beyond positive performance goals, Competition in Reading (six items), Recognition for Reading (five items), and Reading for Grades (four items) are components of extrinsic motivation, which elucidates the external drive that motivates students to read. Finally, Social Reasons for Reading (seven items) and Compliance (five items) are factors of social motivation that contribute to students' use of reading as a means of social interaction. Wigfield and Guthrie (1997) viewed the MRQ as a tool that is capable of eliciting information regarding the multidimensionality of fourth- and fifth-grade elementary school students and evaluated the success of a particular intervention on third-grade students.

SDT (Ryan and Deci, 2000) and self-efficacy theory (Bandura, 1977; Bandura, 1982) are two theories that support the development of the MRQ. Expanding the MRQ items for the abovementioned structures in light of current educational and technological developments is seemingly critical to the field of extensive reading.

Assessment Tools for Measuring Motivation for Reading

Despite the claim that reading motivation in adults is most likely driven by intrinsic motivation and self-regulation, reading motivation during childhood may influence that of adults' motivation (Schutte and Malouff, 2007). Moreover, children's reading motivation may impact adult reading motivation. More research is needed to determine adult reading's full potential. The next section discusses reading MRQ questionnaires. We study the interaction between adult's and children's reading motives and a possible online reading component. The study criticizes the MRQ.

The first measurement is the motivation for the online reading questionnaire (MORQ), which omits several MRQ aspects deemed irrelevant for online reading (Forzani et al., 2020). In the case of extensive online reading, where the teacher controls reading, various dimensions, such as recognition and competitiveness, are considered due to LMS reports on websites that display students' reading progress and self-motivation to compete with their classmates. The MORQ comprises five items organized into four dimensions: curiosity, value, self-efficacy, and self-improvement beliefs. Several characteristics of the MRQ, such as recognition, competition, compliance, and avoidance, are omitted due to the concentration of MORQ in online reading. Simultaneously, the social part of motivation is ignored.

The second questionnaire is the Adult Reading Motivation Measurement (ARMM), which is similar to the MRQ in its multidimensionality (Davis et al., 2020). The hierarchical dimensions of the questionnaire enable it to examine various characteristics of reading motivation, particularly in adolescents. This restriction of the questionnaire can also be used to explain the limitations of the MRQ. Both questionnaires cannot distinguish between school subjects, fiction or non-fiction, and digital or paper reading. However, MORQ is distinguished from MRQ in that the MRQ is geared toward secondary school students instead of those in elementary school. Teachers require an instrument for reading assessment that may be used a few times throughout the semester to assist students in developing a sense of competence and proficiency in reading. MORQ and ARMM demonstrate how the present study may modify the MRQ subscales for the current questionnaire.

Out of the abovementioned prominent questionnaires, many researchers developed instruments based on the dimensions of the MRQ. However, the MRQ continues to leave avenues for further exploration. According to Davis et al. (2020), the MRQ features several limitations with respect to utility as an instrument for elucidating the motivation of students for reading, its small sample size, and the proclivity of motivation researchers to replicate it using an abbreviated version with 18 items instead of the original 53 items. In other words, the creators of other instruments identify areas for improvement relative to MRQ and bridge the gap by validating the questionnaire using larger sample sizes and by including the dimensions in their replication. In addition, (Davis et al., 2020) underlined the importance of researchers who are developing measures for reading motivation that apply to printed and online reading. This notion indicates that researchers on reading motivation have begun to pay special attention to the measurement of online reading motivation.

Additionally, Neugebauer and Fujimoto (2020) detailed several criticisms of the intrinsic motivation component of the MRQ as being ambiguous. As many contend, the challenge subscale of the MRQ was separate from other components with respect to intrinsic motivation; others believe that challenge was a precedent part of the motivation that should be excluded from intrinsic motivation. We also noted that Wigfield and Guthrie (1997) contradicted Wigfield and Guthrie (1995)including intrinsic and extrinsic motivations for reading, perceptions of reading efficacy, social aspects of reading, and reading disincentives. Aa 82-item questionnaire was developed to measure each dimension, with several items assessing each dimension. The questionnaire was completed by 105 fourth- and fifth-grade children in southern Maryland. Factor analyses showed that some of the proposed dimensions were clearly defined, whereas others were not. Several of the dimensions

were correlated with children's book reading frequency in a school-based reading program. The dimensions that appear to be the most reliable include Reading Efficacy, Reading Challenge, Curiosity, Aesthetic Enjoyment, Recognition, Social, and Competition. A revised version of the questionnaire based on the statistical analyses was developed. (Contains 48 references and five tables of data. The original version of the Motivations for Reading Questionnaire is attached. because the first, but not the second, included the importance of reading to intrinsic motivation. Thus, developers of instruments should clarify this inconsistency, especially those who intend to develop instruments to measure the motivation for reading among older learners.

Research Question

This study established the validity of the Motivation for Extensive Reading Questionnaire (MERQ) for 558 undergraduate students of three Indonesian universities. The objective was to determine whether the questionnaire's structures adequately characterize the dimensions of university students' motivation for reading. We assumed that the motivation subscales were classified into eight categories (Self-Efficacy, Reading Challenge, Reading Curiosity, Reading Involvement, Importance of Reading, Recognition for Reading, Reading for Grades, and Social Reasons for Reading), including extensive reading principles and TAM.

METHOD

Participants

The study involved 558 students from three universities in Indonesia (Table 1). From the total sample, 204 students were initially instructed to fill in the questionnaire, and the questionnaire results were analysed using the exploratory factor analysis (EFA). The remaining 354 students were instructed to fill in the questionnaire, and the results were analysed using the confirmatory factor analysis (CFA). The students had been actively learning English since they enrolled in their colleges' English education programmes 2.5 years ago. They had been learning English since primary school but had not used it because it is a foreign language in Indonesia.

Active Year of Learning English	Gender		Total
	Male	Female	
3	200	358	558

Table 1: Characteristics of the sample

Instrument

The Motivation for Extensive Reading Questionnaire (MERQ) was developed using a three-step process. Initially, we established a theoretical foundation for our adaptation of the MRQ. Second, we reduced the subscales and items that were less correlated based on the university context in Indonesia and the age level of the participants. Afterward, we added several pertinent items in light of the extensive reading principles of Day (2015) and the TAM by Davis (1989). Finally, we examined the questionnaire as a whole and fitted it to the remaining MRQ constructs relevant to the study context.

Design

The Institutional Review Board (IRB) of the Doctoral School of Education at the University of Szeged officially approved the research. All participants provided their informed consent, indicating that they accepted to participate in the study.

With the assistance of individual instructors, the MERQ was delivered online to 204 students at the three universities. Given that the measure was developed for EFL students, we limited the sample to individuals enrolled in programs that emphasize studying the English language. Students spent 20 minutes in one session supervised by instructors. The sole responsibility of the instructors was to ensure that each student completed the self-reported questionnaire within the time allotted without any interference. The instructors spent time training students on completing the questionnaire and responding to any questions about the items. After that, we ran an EFA on the students' test results. After reducing several insignificant items, we tested the remaining items on another 354 students through an online questionnaire using Google Form for a CFA test.

Data Analysis

Before conducting the multivariate analysis, we conducted a pre-analysis stage by checking the multivariate normality and linearity of the data set (Byrne, 2005). This stage was conducted to check for possible redundancy among the items that may measure the same latent constructs of the proposed scales in the questionnaire using an inter-item correlation matrix (Cohen et al., 2013; Cohen, 1988). In the long run, if we find correlational overlaps among items, the pre-analysis stage may lead to item deletion.

In analysing the questionnaire data, exploratory factor analysis(EFA) was employed to check the dimensionality of the instrument, which was tested on 204 students. We used the Statistical Package for Social Sciences (SPSS 25) at this stage. The assumed subscales was proposed before undertaking the EFA process. EFA was used to check the rotated factor matrix of the data model and displayed possible item deletion due to low factor loadings (i.e., less than 0.5).

To confirm the EFA result, a confirmatory factor analysis (CFA) was performed to test the results obtained from another group of 354 students. We examined the model fit criteria, such as the comparative fit index (CFI), the Tucker–Lewis index (TLI), and the root mean square error of approximation (RMSEA). Verifying the model fit of the data is essential to determine whether the data are plausible (Hair et al., 2018). At the same time, we checked the instrument's reliability using inter-item and composite reliability to determine the consistency of responses toward the items in the instrument. Simultaneously, Discriminant and convergent validity were investigated to ensure whether the items of the same construct build on the construct itself and to explore whether the items of a construct did not build on other constructs.

RESULTS

EFA

To determine the dimensions of the adapted instrument, we conducted EFA. Building on (Hair et al., 2018), we decided to eliminate items with loadings of less than 0.50 because they were unlikely to be significant in loading the factors of the questionnaire. We omitted six items from the questionnaire with loading values of < 0.50. The omitted items included one for Reading Involvement (I feel like I make friends with people in good books) and five for Social Reasons for Reading items (I often read to my brother or my sister; I like to make contact with the authors of my reading materials; I like to help my friends with their schoolwork in reading; I talk to my friends about what I am reading; and My friends and I like to trade things to read). Only one of the previously added items was omitted (I like to make contact with the authors of my reading materials). The rest of the five items were the original ones of the adapted MRQ. In general, we eliminated these items because they did not contribute to the structure of the instrument. As a result, the remaining 44 items belonging to 8 previously proposed sub-scales were suitable for confirmation through CFA (Appendix 1), which is the subsequent sequential step.

CFA

CFA was used to validate the structure of the MERQ, which resulted in a model fit that can be used to explain the fitness of the model. All observed items loaded significantly based on the loading judgments' characteristics (Hair et al., 2018). The standardized estimate of the factor loadings ranged from 0.65 to 0.85 with a significance level of 0.01, which indicates that all items were acceptable. The fit indices of the questionnaire for each factor were examined. Table 2 summarizes the fit indices for each factor. All the items remained in the CFA because the model fitted the data well. However, the RMSEA of some of the partial models exceeds the recommended cut-off values, such as Reading Efficacy > 0.08, Importance of Reading > 0.08, and Recognition for Reading > 0.08. This happens probably because of the small sample size of the study. We expect that in the future, we can add more samples to refine the root mean square error approximation and improve the fitness index of the models. Moreover, other than RMSEA, the other fit indices of the partial models, such as CFI and TLI indicate a good model fit.

Finally, to ensure the fitness of the model, we checked the fit of the model to the structure. Table 3 displays the fit model of the questionnaire. Overall, the CFA test indicated a good model fit. These results indicated that overall, our instrument has a good fit index.

Constructs	Chi-square	*df*	*p* <	CFI	TLI	RMSEA
Reading Efficacy	8.19	2	0.01	0.99	0.97	0.09
Reading Challenge	27.9	9	0.00	0.98	0.97	0.07
Reading Curiosity	121	35	0.00	0.95	0.93	0.08
Reading Involvement	36.9	20	0.01	0.99	0.99	0.04
Importance of Reading	7.85	2	0.02	0.98	0.94	0.12
Recognition for Reading	28.2	5	0.17	0.96	0.93	0.11
Reading for Grades	0.00	0	n/s	1.00	1.00	0.00
Social Reasons for Reading	0.29	2	0.86	1.00	1.01	0.00

Table 2: Goodness of fit of questionnaire subscales

Chi-square	*df*	*p* <	CFI	TLI	RMSEA	Estimator
1604	874	0.001	0.935	0.930	0.048	ML

Table 3: Goodness of fit of the questionnaire

Reliability

The internal consistency reliabilities of the instrument were calculated as a whole and for each factor. Cronbach's alpha and omega coefficients of the instrument as a whole were acceptable (0.97 and 0.98, respectively). At the same time, Cronbach's alpha and omega coefficients for each factor were within acceptable ranges from 0.82 to 0.97 (Table 4), which indicated satisfactory reliabilities. All the factors suggested equal satisfactory reliabilities.

Constructs	CRB	CR
Reading Efficacy	0.97	0.88
Reading Challenge	0.89	0.90
Reading Curiosity	0.91	0.91
Reading Involvement	0.93	0.93
Importance of Reading	0.83	0.83
Recognition for Reading	0.90	0.90
Reading for Grades	0.82	0.82
Social Reasons for Reading	0.90	0.90

Table 4: Internal consistency reliability and composite reliability of each factor of the questionnaire

Validity

To verify the convergent validity of the scale, we ran average variance extracted (AVE). The results indicated that convergent validity was medium, ranging from 0.51 to 0.66. We assume that these medium AVE values are acceptable, because the majority of the composite reliabilities of

the factors exceed 0.60 (Fornell and Larcker, 1981). As a result, the study established convergent validity.

The heterotrait-monotrait (HTMT) ratio was used to determine discriminant validity (Henseler et al., 2015). Table 5 summarizes the results. The values varied between 0.05 and 0.94. Some of the values are more than 0.85, indicating that discriminant validity has been partially established.

Subscales	AVE	RE	RC	RCU	RI	IOR	RFR	RFG	SRFR
Reading Efficacy	0.66		0.87	0.87	0.85	0.88	0.89	0.89	0.14
Reading Challenge	0.60			0.88	0.92	0.86	0.94	0.74	0.24
Reading Curiosity	0.51				0.86	0.88	0.93	0.78	0.31
Reading Involvement	0.62					0.84	0.87	0.78	0.14
Importance of Reading	0.56						0.90	0.76	0.25
Recognition for Reading	0.52							0.72	0.37
Reading for Grades	0.61								0.05
Social Reasons for Reading	0.52								

Table 5: Convergent validity and discriminant validity

DISCUSSION

This study aimed to develop a more precise measure of university students' motivation for reading in the EFL context, with a particular emphasis on students who have undertaken extensive reading intervention. This instrument was created to support extensive online reading, which has increased in popularity recently. According to the literature, the majority of instructors on extensive reading adapted reading motivation scales from the field of psychology. Among well-known motivation questionnaires for reading, many scholars have referred to the MRQ. Originally composed of 11 subscales with a total of 82 items, this questionnaire was then reduced by its creators to 53 items. The 11 subscales are Reading Efficacy, Reading Challenge, Reading Curiosity, Reading Involvement, Reading Importance, Reading Work Avoidance, Reading for Recognition, Reading Competition, Reading for Grades, Social Reasons for Reading, and Compliance. Additionally, a brief version of the MRQ contains only 18 items. For validation, we adapted the questionnaire with 53 items and used 8 of the 11 subscales of the MRQ.

We omitted reading work avoidance, competition for reading, and compliance from the list of MRQ subscales in light of the contextualization of MRQ with the context of Indonesian universities and the extensive online reading concept. Although Davis et al. (2018) proposed that reading work avoidance is a dimension of students' motivation for reading that should be validated, we refrained from using the subscale because we were primarily interested in the probable beneficial effect of extensive online reading on students' motivation for reading. Although we argued that competition in reading is irrelevant for university students, we also propose that compliance is irrelevant for adult learners who are not required to follow their teachers with respect to extensive reading. By eliminating the three subscales, the total number of items was reduced to 34. Afterward, we added other items based on the TAM concept (Davis, 1989) to determine whether students' motivation for reading was susceptible to the perceived ease of use and usefulness of the online virtual library and any other software packages they may use for extensive reading. Consequently, we added items based on the well-established 10 principles of extensive reading (Day, 2015) because the study focuses on extensive online reading derived from the concept of extensive reading. As a result, 16 items were added, leading to0 items on the Motivation for Extensive Reading Questionnaire (MERQ).

According to the above-mentioned theoretical foundation, the initial number of factors that we proposed was eight, with multiple items for each one: Reading Efficacy (4 items), Reading Challenge (6 items), Reading Curiosity (10 items), Reading Involvement (9 items), Reading for Recognition (5 items), Reading for Social Reasons (9 items), the Importance of Reading (4 items), and Reading for Grades (3 items). The eight subscales are based on three underlying constructs: self-efficacy, intrinsic–extrinsic motivation, and social motivation for reading. According to (Wigfield and Guthrie, 1997), self-efficacy consists of reading efficacy, which indicates a belief one can be successful in reading and reading challenges that lead to the enjoyment of comprehending complicated text (e.g., I learn more from reading than most students in the class and I enjoy reading books about people in different countries). In addition, intrinsic motivation denotes the desire to be good at reading (e.g., I read to learn new information about topics that interest me, I find it easier to manage my reading by using online virtual library [e.g., Xreading, ER-Central, and ReadTheory], and it is very important to me to be a good reader). At the same time, extrinsic motivation prefers external drives that push individuals to read (e.g., I am happy when someone recognizes my reading, and Grades are a good way to see how well I am doing in reading). Finally, the social motivation for reading refers to socialization with others (e.g., I often find uninteresting reading materials turn out to be interesting as many people like them and keep talking and discussing them). The number of items of the MERQ was nearly the same as that of the MRQ, with 50 items, which were then tested using EFA.

We used a fixed number of factors in the EFA process because we were confident in the theoretical foundation when adapting the MRQ. We assumed that eight factors would be extracted from the modified questionnaire. Subsequently, eliminating six items after the EFA process increased the instruments' suitability for assessing students' motivation for extensive reading. Given that five of the omitted

items were derived from the subscales measuring social reasons for reading, which were supposed to complement the scale for motivation goal, we can deduce that the social motivation of reading was relatively less reliant on face-to-face interaction (Appendix). Two of the remaining four SRFR items elicited the use of social media among students to read and express what they had recently read in public (e.g., I find it easy to read and post comments on certain issues posted on Instagram, Twitter, or Facebook and Social media increases my reading motivation through a reading challenge from my friends). In addition, three of the deleted items that did not load to the subscales revealed several pieces of information. First, helping friends with schoolwork in reading (I like to help my friends with their schoolwork in reading) was not a social reason for university students to read. In other words, they read extensively beyond the obligatory homework. Next, I talk to friends about what they are currently reading, as if I talk to my friends about what I am reading that appears already represented by social media. Lastly, the item My friends and I like to trade things to read sounded extraneous because the availability of a wide range of online reading materials has provided the students with abundant and accessible reading materials.

The CFA process confirmed the final factor of the proposed questionnaire. The remaining 44 items after EFA loaded significantly between 0.65 and 0.85 in the CFA. Importantly, this calculation did not influence the structure of the questionnaire. However, we checked the model fit indices to determine the questionnaire's overall fitness and individual factors. Although the outputs of the analysis indicated that the questionnaire fit the model well as a whole, RMSEA results of the individual factor check revealed that five factors were outside the fit model, namely, Reading Efficacy (0.09), Reading Challenge (0.07), Reading Curiosity (0.08), Importance of Reading (0.12), and Recognition for Reading (0.11). Given that a badness-of-fit score of 0.06 is considered within the close fit range (Hair et al., 2019) and a score of 0.10 is considered negligible (Shi, Lee and Maydeu-Olivares, 2019), we deemed that Reading Efficacy, Reading Challenge, Reading Curiosity, Importance of Reading, and Recognition for Reading required additional consideration. However, RMSEA tended to decrease with the addition of the indicators of the observed variables (Shi, Lee, and Maydeu-Olivares, 2019); we theoretically exhausted the possibilities of adding indicators in the quest to obtain a perfect model. Thus, we based our absolute fitness model on the overall RMSEA result of the questionnaire, which fit perfectly. Additionally, the CFI and TLI results for individual factors and whole factors were within acceptable ranges of fit at > 0.90. Thus, we infer that our hypothesized model was fit.

The validity check of the questionnaire indicated that items within the same subscales were built on the respective directed latent variables. At the same time, items of different subscales could be distinguished from one another. This fact supports our additional items to the original MRQ and indicates that the current measurement of the motivation of university students for extensive reading in EFL must consider the technological aspect of motivation (Takase, 2007; Pal and Vanijja, 2020; Rafique et al., 2020) and contextualize the ER principles to the items in the questionnaire (Day, 2015). The final structure of the proposed questionnaire was in line with that of the Takase model for reading motivation in the second language, which included online technology, such as the Internet, to reveal the reading motivation of university students. Simultaneously, the final structure is also in line with the questionnaire developed by (Park, 2015), which focused more on Korean EFL students' intrinsic and extrinsic motivation for reading. Following this validation is relatively interesting for the current study in exploring whether students' extrinsic motivation is related to their use of online reading strategies. In addition, despite our modification to the original MRQ, we continued to retain the expectancy-value aspects (Shang, Moss and Chen, 2023) in the form of subjective task value (importance of reading), which may be perhaps represented more by expectancy values as in the questionnaire of (Mori, 2002), which used nearly all components of expectancy-value aspects, intrinsic value, attainment value, extrinsic value, and importance value. The MERQ also confirmed the MORQ and the ARMM using the original MRQ scales in developing the instrument. However, the MERQ differs from both questionnaires in terms of its ability to transfer psychological theories about motivation to the context of EFL reading.

In conclusion, the MERQ was validated to determine the fittest measurement of students' reading motivation after the extensive online reading intervention in the EFL context. In doing so, we reduced the original subscales of the MRQ without altering the remaining items. Moreover, we added several items based on TAM and extensive reading theories but remained attached to the remaining eight original subscales of the MRQ. Specifically, we aimed to contextualize the MRQ with extensive online reading context at the university level. However, the MERQ has its limitations, which are as follows: First, some MERQ parameters must be reassessed due to low RMSEA. Future subscale additions can fix this problem. Second, study samples were limited to third-year college students with considerable reading intervention experience. We may have improved the questionnaire's generalizability if we had more replies from different fields and semesters. Future studies should address this issue. Third, the validation was not followed by investigating gender, age, English competence, and economic and social status disparities in reading motivation. Future research should uncover these discrepancies. Fourth, future research needs to recruit more respondents. University students are adult learners; therefore, involving diverse jobs of the same age range will raise the chance of getting more replies and improve the fitness of questionnaires measuring motivation for extensive reading in Indonesia. Future studies can increase TAM transfer to questionnaire items.

CONCLUSION

The study demonstrated that the subscales of the MRQ, in conjunction with TAM and extensive reading theories, can be used to assess the motivation of students to read following extensive online reading programs. Thus, this study opened possible avenues for future instructors of extensive reading

to use the items in the proposed questionnaire to establish the positive characteristics of motivated students with respect to EFL reading. MERQ is distinguishable from other questionnaires on extensive reading due to its ability to elicit information about students' motivation for reading via online interfaces. Finally, but certainly not least, the questionnaire may provide teachers or instructors with direct feedback regarding their students' motivation for EFL reading.

ACKNOWLEDGMENT

This work was supported by the Research Programme for Public Education Development, Hungarian Academy of Sciences (Grant KOZOKT2021-16), and by the Humanities and Social Sciences Cluster of the Centre of Excellence for Interdisciplinary Research, Development and Innovation of the University of Szeged. Anita Habók is a member of the Digital Learning Technologies Incubation Research Group.

REFERENCES

Anandari, C. L. and Iswandari, Y. A. (2019) 'Extensive reading in Indonesian schools: A successful story', *Teflin Journal*, Vol. 30, No. 2, pp. 137–152. https://doi.org/10.15639/teflinjournal.v30i2/137-152

Bandura, A. (1977) 'Self-efficacy: Toward a Unifying Theory of Behavioral Change', *Psychological Review*, Vol. 84, No. 2, pp. 191–215. https://doi.org/10.1037/0033-295x.84.2.191

Bandura, A. (1982) 'Self-efficacy mechanism in human agency', *American Psychologist*, Vol. 37, No. 2, pp. 122–147. https://doi.org/10.1037/0003-066x.37.2.122

Byrne, B. M. (2005) 'Factor analytic models: Viewing the structure of an assessment instrument from three perspectives', *Journal of Personality Assessment*, Vol. 85, No. 1, pp. 17–32. https://doi.org/10.1207/s15327752jpa8501_02

Chanthap, N. and Wasanasomsithi, P. (2019) 'The effect of integration of a blended learning and extensive reading instructional model on Thai efl undergraduate students' learner autonomy', *LEARN Journal: Language Education and Acquisition Research Network*, Vol. 12, No. 2, pp. 76–96.

Cohen, R. J., Swerdlik, M. E. and Sturman, E. D. (2013) *Psychological Testing and Assessment*, 8th Edition, New York: McGraw Hill.

Cohen, J. (1988) *Statistical Power Analysis for the Behavioral Sciences*, 2nd Edition, New York: Lawrence Erlbaum Associates.

Cote, T. and Milliner, B. (2015) 'Implementing and Managing Online Exensive Reading: Student Performance and Perceptions', *IALLT Journal of Language Learning Technologies*, Vol. 45, No. 1, pp. 70–90. https://doi.org/10.17161/iallt.v45i1.8550

Dakhi, S. and Damanik S. I. (2018) 'Students' Motivation in Reading English Text: A Qualitative Study in EFL Context', *Journal of English Teaching*, Vol. 4, No. 2, pp. 81–92. https://doi.org/10.33541/jet.v4i2.832

Davis, F. D. (1989) 'Perceived usefulness, perceived ease of use, and user acceptance of information technology', *MIS Quarterly: Management Information Systems*, Vol. 13, No. 3, pp. 319–339. https://doi.org/10.2307/249008

Davis, F. D., Bagozzi, R. P. and Warshaw, P. R. (1989) 'User Acceptance of Computer Technology: A Comparison of Two Theoretical Models', *Management Science*, Vol. 35, No. 8, pp. 982–1003. https://doi.org/10.1287/mnsc.35.8.982

Davis, M. H., Tonks, S. M., Hock, M., Wang, W. and Rodriguez, A (2018) 'A review of reading motivation scales', *Reading Psychology*, Vol. 39, No. 2, pp. 121–187. https://doi.org/10.1080/02702711.2017.1400482

Davis, M. H., Wang, W., Kingston, N. M., Hock, M., Tonks, S. M. and Tiemann, G. (2020) 'A computer adaptive measure of reading motivation', *Journal of Research in Reading*, Vol. 43, No. 4, pp. 434–453. https://doi.org/10.1111/1467-9817.12318

Day, R. (2015) 'Extending Extensive Reading.', *Reading in a Foreign Language*, Vol. 27, No. 2, pp. 294–301.

Deci, E. L. and Ryan, R. M. (1985) *Intrinsic Motivation and Self-Determination in Human Behavior*, New York: Springer Science.

Deci, E. L. and Ryan, R. M. (2000) 'The " What " and " Why " of Goal Pursuits : Human Needs and the Self-Determination of Behavior', *Psychological Inquiry*, Vol. 11, No. 4, pp. 37–41. https://doi.org/10.1207/s15327965pli1104_01

Ferrari, J. R., Johnson, J. L. and McGown. W. G. (1995) *Procrastination and task avoidance - Theory, Research, and Treatment*. Edited by C. R. Snyder. Kansas: Springer Science.

Fornell, C. and Larcker, D. F. (1981) 'Evaluating Structural Equation Models with Unobservable Variables and Measurement Error', *Journal of Marketing Research*, Vol. 18, No. 1, pp. 39–50. https://doi.org/10.2307/3151312

Forzani, E., Leu, D. J., Yujia Li, E., Rhoads, C., Guthrie, J. T. and McCoach, B. (2020) 'Characteristics and Validity of an Instrument for Assessing Motivations for Online Reading to Learn', *Reading Research Quarterly*, Vol. 56, No. 4, pp. 761–780. https://doi.org/10.1002/rrq.337

Habók, A., Magyar, A., Németh, M. B. and Csapó, B. (2020) 'Motivation and self-related beliefs as predictors of academic achievement in reading and mathematics: Structural equation models of longitudinal data', *International Journal of Educational Research*, Vol. 103, p. 101634. https://doi.org/10.1016/j.ijer.2020.101634

Hagley, E. T. (2017) 'Extensive graded reading with engineering students : Effects and outcomes', *Reading in a Foreign Language*, Vol. 29, No. 2, pp. 203–217.

Hair, J. F., Babin, B. J., Anderson, R. E. and Black, W. C. (2019) Multivariate Data Analysis, 8th Edition, England: Pearson Prentice.

Hendriwanto and Kurniati, U. (2019) 'Building reading fluency with mobile assisted extensive reading', *International Journal of Interactive Mobile Technologies*, Vol. 13, No. 6, pp. 84–92. https://doi.org/10.3991/ijim.v13i06.9799

Henseler, J., Ringle, C. M. and Sarstedt, M. (2015) 'A new criterion for assessing discriminant validity in variance-based structural equation modeling', *Journal of the Academy of Marketing Science*, Vol. 43, No. 1, pp. 115–135. https://doi.org/10.1007/s11747-014-0403-8

Kim, K. J. (2011) 'Reading motivation in two languages: An examination of EFL college students in Korea', *Reading and Writing*, Vol. 24, No. 8, pp. 861–881. https://doi.org/10.1007/s11145-010-9229-z

Matsuda, S. (2020) 'Sharing Reading Experiences with University Students Using Goodreads', *Journal of Extensive Reading*', pp. 77–86.

Mori, S. (2002) 'Redefining Motivation to Read in a Foreign Language', *Reading in a Foreign Language*, Vol. 14, No. 2, pp. 91–110.

Neugebauer, S. R. and Fujimoto, K. A. (2020) 'Distinct and Overlapping Dimensions of Reading Motivation in Commonly Used Measures in Schools', *Assessment for Effective Intervention*, Vol. 46, No. 1, pp. 39–54. https://doi.org/10.1177/1534508418819793

Pal, D. and Vanijja, V. (2020) 'Perceived usability evaluation of Microsoft Teams as an online learning platform during COVID-19 using system usability scale and technology acceptance model in India', *Children and Youth Services Review*, Vol. 119, p. 105535. https://doi.org/10.1016/j.childyouth.2020.105535

Park. J. (2015) 'Insights into Korean EFL Students' Reading Motivation, Proficiency, and Strategy Use', *English Teaching*, Vol. 70, No.1, pp. 57–74. https://doi.org/10.15858/engtea.70.1.201503.57

Pirih, A. (2019) 'Extensive reading and changes to readingmotivationinefl among slovene primary school pupils', *Journal of Elementary Education*, Vol. 12, No. 4, pp. 291–314. https://doi.org/10.18690/rei.12.4.291-314.2019

Protacio, M. S. (2012) 'Reading motivation: A focus on english learners', *Reading Teacher*, Vol. 66, No. 1, pp. 69–77. https://doi.org/10.1002/trtr.01092

Rafique, H., Almagrabi, A. O., Shamim, A., Anwar, F. and Bashir, A. K. (2020) 'Investigating the Acceptance of Mobile Library Applications with an Extended Technology Acceptance Model (TAM)', *Computers and Education*, Vol. 145, p. 103732. https://doi.org/10.1016/j.compedu.2019.103732

Renandya, W. A. (2007) 'The power of extensive reading', *RELC Journal*, Vol. 38, No. 2, pp. 133–149. https://doi.org/10.1177/0033688207079578

Rezaee, M. and Farahian, M. (2020) 'Promoting university students' receptive skills through extensive reading in multimedia-based instruction', *Journal of Applied Research in Higher Education*, Vol. 13, No. 5, pp. 1464–1489. https://doi.org/10.1108/jarhe-09-2020-0304

Ryan, R. M. and Deci, E. L. (2000) 'Self-Determination Theory and the Facilitation of Intrinsic Motivation, Social Development, and Well-Being', *University of Dar es Salaam Library journal*, Vol. 55, No. 1, pp. 68–78. https://doi.org/10.1037/0003-066x.55.1.68

Schutte, N. S. and Malouff, J. M. (2007) 'Dimensions of reading motivation: Development of an adult reading motivation scale', *Reading Psychology*, Vol. 28, No. 5, pp. 469–489. https://doi.org/10.1080/02702710701568991

Shang, C., Moss, A. C. and Chen, A. (2023) 'The expectancy-value theory: A meta-analysis of its application in physical education', *Journal of Sport and Health Science*, Vol. 12, No. 1, pp. 52–64. https://doi.org/10.1016/j.jshs.2022.01.003

Shi, D., Lee, T. and Maydeu-Olivares, A. (2019) 'Understanding the Model Size Effect on SEM Fit Indices', *Educational and Psychological Measurement*, Vol. 79, No. 2, pp. 310–334. https://doi.org/10.1177/0013164418783530

Shurentsetseg, N. and Nyamsuren, A. (2015) 'Implementing Extensive Reading through Xreading.com for Mongolian EFL Learners'.

Takase, A. (2007) 'Japanese High School Students' Motivation for Extensive L2 Reading', *Reading in a Foreign Language*, Vol. 19, No. 1, pp. 1–18.

Tuominen, H., Niemivirta, M., Lonka, K. and Salmela-Aro, K. (2020) 'Motivation across a transition: Changes in achievement goal orientations and academic well-being from elementary to secondary school', *Learning and Individual Differences*, Vol. 79, p. 101854. https://doi.org/10.1016/j.lindif.2020.101854

Wigfield, A. (1994) 'Expectancy-value theory of achievement motivation: A developmental perspective', *Educational Psychology Review*, Vol. 6, No. 1, pp. 49–78. https://doi.org/10.1007/bf02209024

Wigfield, A. and Eccles, J. S. (2000) 'Expectancy-value theory of achievement motivation', *Contemporary Educational Psychology*, Vol. 25, No. 1, pp. 68–81. https://doi.org/10.1006/ceps.1999.1015

Wigfield, A. and Guthrie, J. T. (1995) 'Dimensions of Children's Motivations for Reading: An Initial Study. Reading Research Report No. 34', *Reading Research Report No. 34*, pp. 9–33.

Wigfield, A. and Guthrie, J. T. (1997) 'Relations of Children's Motivation for Reading to the Amount and Breadth of Their Reading', *Journal of Educational Psychology*, Vol. 89, No. 3, pp. 420–432. https://doi.org/10.1037/0022-0663.89.3.420

Wolgast, A. (2018) 'School students' beliefs about abilities and perspective-taking over time', *Large-Scale Assessments in Education*, Vol. 6, No. 1. https://doi.org/10.1186/s40536-018-0065-y

APPENDIX

EXPLORATORY FACTOR ANALYSIS OF THE MOTIVATION FOR EXTENSIVE READING QUESTIONNAIRE

No.	Code	Items	Loading	Factor
1	RCU4	I read to learn new information about topics that interest me.	0.84	Reading Curiosity
2	RCU7	I read about my hobbies to learn more about them.	0.82	
3	RCU6	I find it easier to read about information I want to know on Google.	0.81	
4	RCU5	I like to read about new things.	0.82	
5	RCU1	If the teacher discusses something interesting, I might read more about it.	0.79	
6	RCU2	I have favourite subjects that I like to read about.	0.77	
7	RCU8	I cannot stop reading a series until I reach its end.	0.76	
8	RCU10	I enjoy reading a series.	0.72	
9	RCU3	I enjoy reading books about people in different countries.	0.69	
10	RCU9	I always choose the reading materials by myself.	0.65	
11	RI7	I find it easier to manage my reading by using online virtual library (Xreading, ER-Central, ReadTheory, etc.).	0.82	Reading Involvement
12	RI1	I enjoy a long, involved story or fiction book.	0.79	
13	RI4	I like mysteries.	0.79	
14	RI9	Reading graded readers in online virtual library (Xreading, ER-Central, ReadTheory, etc.) increases my reading rate.	0.78	
15	RI2	I make pictures in my mind when I read.	0.75	
16	RI3	I read stories about fantasy and make believe.	0.69	
17	RI8	I like to read various topics and genres.	0.69	
18	RI5	I read a lot of adventure stories.	0.67	
19	RC3	I like it when the questions in books make me think.	0.75	Reading Challenge
20	RC5	I like hard, challenging books.	0.75	
21	RC2	If a book is interesting, I don't care how hard it is to read.	0.73	
22	RC1	I usually learn difficult things by reading.	0.72	
23	RC4	If the project is interesting, I can read difficult material.	0.70	
24	RC6	I always want to read reading materials which are slightly above my reading level.	0.60	
25	RFR1	I am happy when someone recognizes my reading.	0.77	Reading for Recognition
26	RFR4	I like to get compliments for my reading.	0.74	
27	RFR3	My friends sometimes tell me I am a good reader.	0.74	
28	RFR2	I like having the teacher say I read well.	0.74	
29	RFR5	I always wait for my teacher to report our reading progress.	0.70	
30	SRFR9	I often find that uninteresting reading materials turn out to be interesting because many people like them and keep talking and discussing them.	0.71	Social Reasons for Reading
31	SRFR8	I find it easy to read and post comments on certain issues posted in Instagram, Twitter, or Facebook.	0.65	
32	SRFR6	I like to tell my family about what I am reading.	0.65	
33	SRFR7	Social media increases my reading motivation through reading challenge from my friends.	0.59	
34	IOR1	It is very important to me to be a good reader.	0.80	Importance of Reading
35	IOR2	Compared to other activities, it is very important to me to be a good reader.	0.76	
36	IOR3	I don't mind getting bad reading scores as long as I love reading.	0.77	
37	IOR4	I feel something is missing from my life if I do not read any time in a day.	0.73	
38	RFG2	Grades are a good way to see how well I am doing in reading.	0.81	Reading for Grade
39	RFG3	I look forward to finding out my reading grade.	0.82	
40	RFG1	I read to improve my grades.	0.78	
41	RE1	I know that I will do well in reading next year.	0.78	Self-Efficacy
42	RE3	I learn more from reading than most students in the class.	0.72	
43	RE2	I am a good reader.	0.73	
44	RE4	I can read any reading materials.	0.62	

SELF-REPORTED ZOOM EXHAUSTION AND FATIGUE LEVELS AMONG PHYSICAL EDUCATION TEACHER EDUCATION STUDENTS IN A STATE UNIVERSITY IN THE PHILIPPINES

Julius Ceazar G. Tolentino
John Paul P. Miranda✉

Don Honorio Ventura State University, Philippines

✉ jppmiranda@dhvsu.edu.ph

ABSTRACT

The study determined the self-reported Zoom exhaustion and fatigue (ZEF) levels of physical education teacher education (PETE) students in the Philippines who are attending a flexible learning setup (i.e., synchronous and asynchronous learning) offered by their university as a solution to continue classes during the coronavirus disease (COVID)-19 pandemic. Utilizing a cross-sectional comparative research design, the study indicates that PETE students were very tired and exhausted both in general and visually after a series of video conferencing within a semester. Furthermore, they were socially-, motivationally-, and emotionally- moderately tired and exhausted. It was observed that students in the PETE program experienced moderate to very tired and exhausted levels of fatigue whenever they participated in synchronous online classes, such as attendance at a video teleconferencing platform. Exploration of how specific mental aspects relate to their general health with regard to their culture and habits is worth exploring, either for students, teachers, or professionals in general.

KEYWORDS

Self-reported, Zoom exhaustion and fatigue, synchronous and asynchronous learning, teacher education students

HOW TO CITE

Tolentino J. C. G., Miranda J. P. P. (2024) 'Self-Reported Zoom Exhaustion and Fatigue Levels among Physical Education Teacher Education Students in a State University in the Philippines', *Journal on Efficiency and Responsibility in Education and Science*, vol. 17, no. 3, pp. 205–213. http://dx.doi.org/10.7160/eriesj.2024.170303

Article history
Received
December 22, 2022
Received in revised form
August 27, 2023
Accepted
May 23, 2024
Available on-line
September 30, 2024

Highlights

- *PETE students experienced moderate to very high levels of fatigue, especially during synchronous online classes, with visual fatigue being the most reported issue.*
- *Visual fatigue was linked to extended exposure to devices used for attending virtual classes and contributes to students' exhaustion.*
- *Male PETE students exhibited higher level of exhaustion compared to their female counterpart.*

INTRODUCTION

The coronavirus disease (COVID-19 pandemic) has posed multiple challenges in a plethora of sectors all around the world. The education sector is no exception, as this was massively affected by the global health crisis. For this reason, a major paradigm shift in the delivery of instruction to learners worldwide prompted educational leaders to take affirmative actions and, therefore, requires solutions that would enable the continuity of learning. While closures of educational institutions were frequent at the height of the pandemic, transitioning to a distance learning approach was deemed necessary.

The transition from a traditional classroom setting to a virtual learning modality has led to the utilization of existing and emerging technologies that may essentially provide solutions to connect to learners and continuously deliver instruction (Henritius et al., 2019). With this, the use of video conferencing platforms, for one, became a common means to teach learners synchronously despite location or time zone differences. This allows students and teachers to communicate in real-time, various ways of teaching can be administered, and learning can ultimately occur (Rasouli et al., 2020). For instance, the utilization and support of Microsoft Teams, Zoom, Cisco Webex, and Google Meet, among others, often create lively and interactive learning environments even in a virtual world (Iannizzotto et al., 2020). Furthermore, these educational application tools generate a more feasible and viable method to

continuously pursue education amidst the challenges brought by the global health crisis.

However, despite the advantageous reasons for the use of video conferencing tools for teaching and learning, scholars started to investigate an emerging phenomenon of possible fatigue or exhaustion caused by periodic attendance to virtual meetings, which is referred to in contemporary studies as video conferencing or Zoom fatigue (Fauville et al., 2021). The need to investigate this phenomenon, particularly among college students, is of vital interest to consider mechanisms that would lead to the more effective delivery of learning processes that will not compromise their welfare in terms of emotional, mental, social, and general states. Looking into this dimension of learning in a virtual ecology would lead to a well-structured learning plan that is affirmative of students' needs and will maximize learning outcomes.

LITERATURE REVIEW

COVID-19, Remote Learning, and Videoconferencing in Higher Education

At the onset of the COVID-19 pandemic, many educational institutions shifted from in-person to online learning (Alvarez, Abel, 2020; K. A. Bird et al., 2022; Lim et al., 2022). As mentioned by Rotas and Cahapay (2020), the impact of COVID-19 on online and remote learning, in general, is substantial. Bird et al. (2022) suggested that teachers with a considerable amount of teaching experience with the same course for a long time do not even mitigate the negative effects of this sudden shift to virtual learning. the pandemic has brought the need for more video calls for learning or work-related matters. Studies conducted in 2021 found that the increase in screen time during the pandemic was significantly higher than before (Ganne et al., 2021; Pandya & Lodha, 2021).

Meanwhile, Asurion (2019) reported that, on average, Americans check their devices, particularly smartphones, 96 times a day - a 20% increase even before the pandemic occurred. Onn the other hand, Filipinos spent at least 10 hours and two minutes daily in 2019, mostly browsing the Internet or visiting social media platforms (Gonzales, 2019; Balita, 2024). In another study, Nagata et al. (2022) reported that screen time among children has doubled during the pandemic as most of them rely on their devices to conduct various activities (e.g., entertainment, socialization, and education). Similar findings have also shown a drastic increase in total screen time during the pandemic (Pandya & Lodha, 2021). This is alarming, partly because it was reported that an increase in screen time often leads to higher eye strain among students, particularly those attending online classes (Ganne et al., 2021; Sundarasen et al., 2020). Moreover, increased screen time often negatively impacts people's physical and mental well-being (Pandya & Lodha, 2021).

Remote learning adversely affects students and teachers in general (Bautista et al., 2021). Rotas and Cahapay (2020) mentioned that the sudden shift to online learning has also impacted teachers. The uncertainty of when the pandemic will end also increases anxiety in higher education (Jung et al., 2021). These things are attributed to making online teaching and learning even more challenging for both teachers and students. Despite this, Bolatov et al. (2021) believed that the shift to online learning not only helped to lessen the spread of COVID-19 but also provided some positive benefits to the students' mental health. This is supported by a recent study indicating that medical students in Saudi Arabia received the sudden shift to online learning (Khalil et al., 2020). On the contrary, the abrupt shift from online to in-person learning has been found to negatively affect student academic performance (K. A. Bird et al., 2022).

Furthermore, Gonzales-Ramirez et al. (2021) claimed that remote learning makes students more exhausted. They also believed that it has far-reaching implications for the students, both mentally and physically (Sundarasen et al., 2020). On the other hand, Li et al. (2021) suggested that people's life satisfaction and mental health are also at risk when they overuse social media and the internet to compensate for many things they cannot do, especially during lockdowns (e.g., lack of physical interaction).

Students' Weariness Towards Online Learning

During the pandemic, there has been an increase in gadget usage for both online learning and entertainment. This has caused the prevalence of digital eye strain among students. This is true particularly when they are compared to the general public (Ganne et al., 2021). Moreover, students are sometimes tended or forced to learn and are tempted to multitask despite attending online classes (Alibudbud, 2021; Baticulon et al., 2021), mainly when this is done at home (Baticulon et al., 2021).

Several studies have also indicated that barriers to online learning also contribute to the physical and psychological state of the students (Baticulon et al., 2021; Ortega et al., 2022; Peper et al., 2021; Sundarasen et al., 2020). For example, poor communication related to their schooling contributes to the students' weariness during the conduct of online classes (Baticulon et al., 2021; Peper et al., 2021; Sundarasen et al., 2020). Frustration due to the lack of necessities such as food, conducive space, limited access to devices, and a reliable internet connection is part of why students feel more exhausted than ever (Alvarez, 2020; Baticulon et al., 2021; Rotas & Cahapay, 2020; Sundarasen et al., 2020). Increased workload, and other activities unrelated to learning were also seen to contribute or to students' overall fatigue and exhaustion (Peper et al., 2021). Another issue related to student weariness is that in an online learning setup, students are most of the time forgotten or unable to do physical activities they used to do in an in-person classroom setting. Common physical activities that require movement, such as walking, running, and standing, are considered important and a usual part of a person's daily life much more so with student life. Peper and Lin reported in 2021 that students who do physical activities significantly increase their subjective energy and increase their attention levels.

An increase in social isolation is also seen among students during the period of online learning. It is perceived as one of the contributing factors to student weariness (Li et al., 2021; Peper et al., 2021). The lack of social interaction among students, teachers, and with their peers often contributes to their physical and mental exhaustion (Peper et al., 2021).

In connection, De Oliveira Kubrusly Sobral et al. (2022) found that students attending hybrid methodology tend to have a higher frequency of wanting to be alone after attending a video conference. Furthermore, it was found in the study by Martínez-Líbano et al. (2021) that Chilean social science students' exhaustion during the COVID-19 pandemic showed that their study stipulates that the students had higher levels of exhaustion and believed that their mental health deteriorated during the pandemic. The abrupt shift to online learning makes burnout and other negative mental symptoms prevalent among students (Bolatov et al., 2021). Other challenges that contribute to this phenomenon include pressure to concentrate during attendance to online learning while the threat of the COVID-19 pandemic is ongoing (Peper et al., 2021), overloaded activities (Rotas & Cahapay, 2020), and lack of control over various issues (i.e., technical issues) related to online learning (Peper et al., 2021).

Despite this, several studies have recommended countering student weariness in online learning. For example, Alibudbud (2021) suggested that regular breaks and avoiding multitasking are necessary to prevent burnout when attending online learning and to improve concentration among students attending such classes. Furthermore, limiting screen time is also a good option to potentially address this issue (Ganne et al., 2021). Alternatively, Räisänen et al. (2018) believed that identifying students' profiles is imperative to help students who need support in their learning. The key findings in their study, for example, suggested that students who already have self-regulated problems paired with high levels of peer learning and peer support tend to have higher study-related exhaustion. Another option is the intervention that lawmakers and policymakers can provide to address this issue, which is crucial (Ganne et al., 2021). Their intervention is needed as the challenges and problems related to student weariness may persist even after the pandemic subsides (De Oliveira Kubrusly Sobral et al., 2022).

In light of existing literature and studies, there is a need to further expand the knowledge on how students, particularly those who specialize in a movement-dominated discipline like physical education in the Philippines, perceive and assess their levels of exhaustion in multiple spectra of fatigue when attending classes via a video teleconferencing platform. This research documented the fatigue levels of pre-service physical educators in a Philippine-based public university and determined whether constructs underlying this emerging type of fatigue in contemporary times had statistically significant differences when their demographic profiles were considered. The study was an attempt to uncover this phenomenon, which may eventually serve as a data-driven and empirical basis for developing appropriate modalities for students, either in distance or blended learning. Through this, more proactive and tangible programs for curriculum delivery and instruction may be institutionalized toward a well-rounded learning experience for pre-service teachers in physical education.

THEORETICAL BACKGROUND

The theoretical framework of this study explores the potential interconnections between sex, class level, and Zoom exhaustion and fatigue (ZEF), aiming to shed light on their collective influence. Incorporating sex and class-level variables into the study objectives serves to address a pivotal research context. Existing literature has indicated that various demographic factors, including sex and class level (Dacillo et al., 2022; Purba et al., 2022; Fauville et al., 2023; Oducado et al., 2022; Salim et al., 2022; Usta Kara & Esroy, 2022), could contribute to shaping individuals' experiences of ZEF. Notably, gender has been implicated in Zoom-related fatigue studies (George et al., 2022; Ratan et al., 2021; Shockley et al., 2021). Emerging research also suggests potential disparities in how individuals perceive and navigate challenges within online interactions based on these factors (Dacillo et al., 2022; Purba et al., 2022; Ratan et al., 2021; Usta Kara & Esroy, 2022). In the context of the rapid expansion of virtual communication platforms like Zoom, comprehending the intersections between these variables and ZEF holds substantial academic and practical significance. This study, by delving into the potential impacts of sex and class level on ZEF, particularly among students engaged in active tasks, seeks to contribute to a nuanced comprehension of the intricate dynamics underpinning individuals' fatigue during virtual engagements.

This investigation into the potential effects of ZEF, particularly in educational settings, based on sex and class level could offer a more comprehensive understanding of its impact on individual students. Both sex and class level may exert influence by shaping students' experiences, stressors, coping mechanisms, and overall mental well-being. Research has suggested sex-based variations in coping mechanisms among students (Christiansen et al., 2022; Grace, 2019; Graves et al., 2021; Mahmoud et al., 2015; Scott-Young et al., 2020), potentially leading to different manifestations of fatigue and stress. As men and women are distinct in their psychological makeup and socialization patterns, societal expectations may engender divergent responses (L. J. Bird et al., 2023; Cislaghi & Heise, 2020; Ellemers, 2018; Grace, 2019; Newsome et al., 2016). Class-level, likewise, has been linked to varying workloads and expectations that could contribute to Zoom-related fatigue (Bare et al., 2023; Bird et al., 2023; Grace, 2019; Labrague, 2013; Rotas & Cahapay, 2020). For instance, earlier studies underscore that first-year students grappling with a transition phase and novel learning environments may necessitate greater support (Blair, 2016; Honkimäki & Kálmán, 2012; Mahmoud et al., 2015; Maymon et al., 2019; Meehan & Howells, 2018; Nyar, 2021). Concurrently, class-level disparities may correlate with Zoom-related fatigue, as students at different academic stages may possess distinct focuses, possibly impacting susceptibility to burnout, fatigue, stress, and related outcomes (Bird et al., 2023; Little et al., 2021; Mahmoud et al., 2015; Nyar, 2021).

AIMS OF THE STUDY

The researchers aimed to report the levels of ZEF as perceived by students in a pre-service teacher education (PETE) institution situated in a state university in Pampanga, Philippines. Specifically, the following research questions were answered:

1. How may the self-reported levels of ZEF among PETE students be described in terms of (a) emotional fatigue, (b) motivational fatigue, (c) social fatigue, (d) visual fatigue, and (e) general fatigue?

2. Are there statistically significant differences in the individual ZEF levels of the respondents when grouped according to their sexes and class level?

METHODOLOGY

Research Design

The study adopted a comparative cross-sectional survey design to collect data on the self-reported levels of ZEF among PETE students. A cross-sectional approach allows for collecting data from multiple respondents at a single point in time, providing a snapshot of their experiences in a particular phenomenon, the Zoom exhaustion fatigue in this context. Apart from the descriptive nature of the design, it particularly centred on establishing a thorough analysis of the differences in the ZEF levels of PETE students when grouped according to their sex and class levels.

Respondents

The study was conducted in a teacher education institution in a public university in Pampanga, Philippines. For the whole academic year 2021 - 2022, the university offered full online learning, particularly employing a combination of synchronous and asynchronous sessions as a modality to deliver lecture and laboratory courses. The courses are offered either as a three-hour lecture or as a five-hour laboratory. There were 37,398 students in the whole university enrolled at the time of the study. The study utilized a purposive random sampling technique among 555 students enrolled in the physical education teacher education program. The said student groups are of interest as their program demands actual and physical demonstration of skills, which may seem to be challenging in a virtual context. Using the list provided by the university, with a 99% confidence level and a five percent margin of error, a minimum of 303 respondents are needed for the study. To reduce the attrition, 350 prospective respondents were invited to answer the short electronic and self-administered survey sent to their respective institutional email accounts the day before the end of the academic year to immediately assess their perceived fatigue level concerning the online modality employed for their learning.

Respondents' Demographics

As seen in Table 1, 303 respondents were included in the study. Their age ranged from 19 to 21 years old. Based on the total respondents, 196 are females (64.7%), and 107 are males (35.3%). Furthermore, almost a third of them are either first-year (n = 94) or fourth-year students (n = 92).

Variables	Frequency	Percentage
Age (mean ± *SD*)	20 ± *1.626*	
Sex		
Male	107	35.3%
Female	196	64.7%
Year Level		
First Year	94	31.0%
Second Year	61	20.1%
Third Year	56	18.5%
Fourth Year	92	30.4%
TOTAL	303	100.0%

Table 1: Respondent's profile

As illustrated in Figure 1, 40% of the respondents participated in a video conference about once a day before the conduct of this study (n = 124). Furthermore, roughly a third of the respondents said their average daily participation in video conferences was at least once a day. More than half said these video conferences lasted over an hour (Figure 2). Meanwhile, when respondents were asked about the average interval between each video conference within the day, a third of them mentioned that the gap was more than an hour (Figure 2).

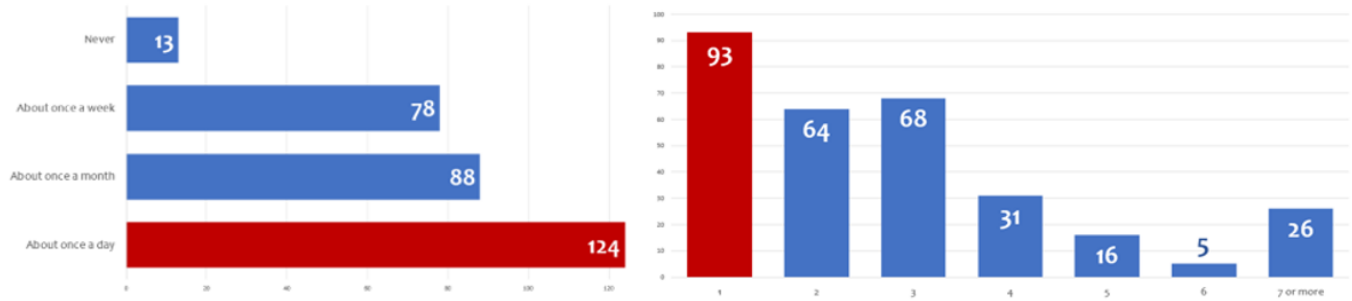


Figure 1: Average length (left) of each video conference, including their interval (right)

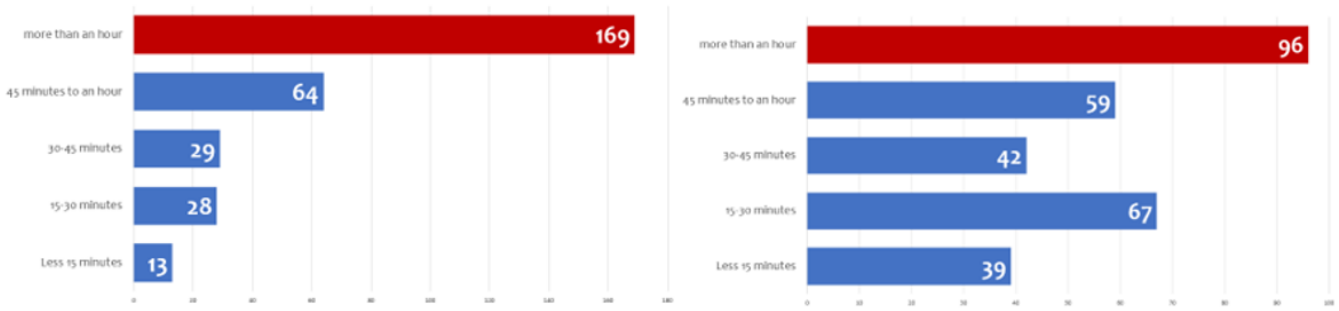


Figure 2: Overall (left) and daily average (right) participation in video conferences

Instrument

The primary instrument utilized in this study was the Zoom Exhaustion and Fatigue (ZEF) scale, developed by Fauville et al. (2021). No modifications were made to the instrument, and it was utilized without any translation, validity checks, or additional reliability testing for this particular study. The instrument's prior validation and reliability testing by Fauville provided a foundation for its applicability. Within the context of this study, ZEF denoted the fatigue arising from frequent participation in videoconferencing, regardless of the platform. This 15-item instrument was structured around five distinct dimensions: general, visual, social, motivational, and emotional aspects. Each dimension encompassed three questions, yielding a total of 15 items. The ZEF utilized English as a medium of instruction in its administration (hence, translation was no longer needed). This is applicable in the context of the Philippines because English is one of the country's two official languages by virtue of the 1987 Constitution (along with Filipino), which is deemed used for "purposes of communication and instruction," and this cascades from basic to higher education.

The dimensions of the ZEF instrument were crafted to evaluate various dimensions of exhaustion and fatigue specifically attributed to Zoom interactions. The instrument's reliability was substantiated by Fauville et al. (2021), who reported robust indices such as Cronbach's alpha (ranging from .82 to .90) and composite reliability (ranging from .83 to .90). These indices surpassed the conventional threshold of .70, underscoring the instrument's dependable reliability.

The collected data underwent analysis to extract meaningful insights. Descriptive statistics, including mean, standard deviation, and percentage, were computed to provide an overview of the data distribution. A normality test was conducted to ascertain the data's adherence to normal distribution utilizing IBM Statistical Package for Social Sciences (SPSS) version 25.0. Subsequently, non-parametric tests were applied, specifically the Mann-Whitney and Kruskal-Wallis tests. These tests were selected to explore potential variations in the self-reported ZEF levels based on two variables: respondents' sex and class level. These non-parametric tests were chosen due to the distribution characteristics of the ZEF data. The interpretation of the analysis outcomes drew upon the insights provided by Table 2. This table detailed the Likert scale employed, its corresponding range, and the verbal interpretation associated with each mean score.

Scale	Range Value	Verbal Interpretation
5	4.50 - 5.00	Extremely tired and exhausted
4	3.50 - 4.49	Very tired and exhausted
3	2.50 - 3.49	Moderately tired and exhausted
2	1.50 - 2.49	Slightly tired and exhausted
1	1.00 - 1.49	Not at all tired and exhausted

Table 2: Five-point Likert rating scale and its interpretation

Ethical Considerations

All respondents were assured that they could discontinue answering the survey at any given time and that all the data collected from them were strictly confidential and solely intended for this study. Furthermore, all the data will be destroyed one year after the study is conducted in compliance with the existing data privacy laws in the Philippines.

RESULTS AND DISCUSSION

Self-reported Levels of ZEF among PETE Students

This study examined the self-reported ZEF levels of PETE students attending synchronous and asynchronous classes in a public university in the Philippines. The results indicated that students are mostly exhausted across ZEF constructs (i.e., general, visual, motivational, social, and emotional). Figure 5 indicates the respondents' ZEF-reported levels. It shows that PETE students are very tired and exhausted both in the general sense ($\bar{x}$ = 3.50, SD = 0.97) and visually ($\bar{x}$ = 3.64, SD = 0.99) after a series of video conferencing within a semester. Furthermore, PETE students are moderately tired and exhausted in terms of social ($\bar{x}$ = 3.45, SD = 0.93), motivational ($\bar{x}$ = 3.47, SD = 0.92), and emotional ($\bar{x}$ = 3.31, SD = 0.99) domains. This is consistent with De Oliveira Kubrusly Sobral et al. (2022) findings, where half of the students surveyed experienced Zoom fatigue. Another reason for this fatigue is because of higher screen time among

students, particularly during the COVID-19 pandemic, when most students rely on their devices to do almost everything (Nagata et al., 2022). This is supported by previous studies where they posited that this kind of fatigue will persist even after the pandemic (De Oliveira Kubrusly Sobral et al., 2022; Nagata et al., 2022). Several studies have offered some kind of treatment to solve this. For one, Alibudbud (2021) recommended that regularly providing a short break every thirty minutes and avoiding multitasking while attending online classes can improve student concentration. This is supported by previous studies indicating that performing some physical activities, even for a minute, can increase positive mental well-being (Peper et al., 2021). Moreover, revisiting and revising the course activities and outcomes to adjust these recommendations are necessary to avoid student burnout (Alibudbud, 2021). In addition, lifestyle modification and self-imposed limitations to the use of digital media are seen to have positive effects on better health and well-being (Ganne et al., 2021; Li et al., 2021; Pandya & Lodha, 2021). Providing adequate support to teachers and promoting mental health training are also seen as important factors in fostering a healthy school environment. This intervention is seen to have a direct and indirect impact on reducing fatigue and exhaustion among students (Alibudbud, 2021; Bautista et al., 2021).

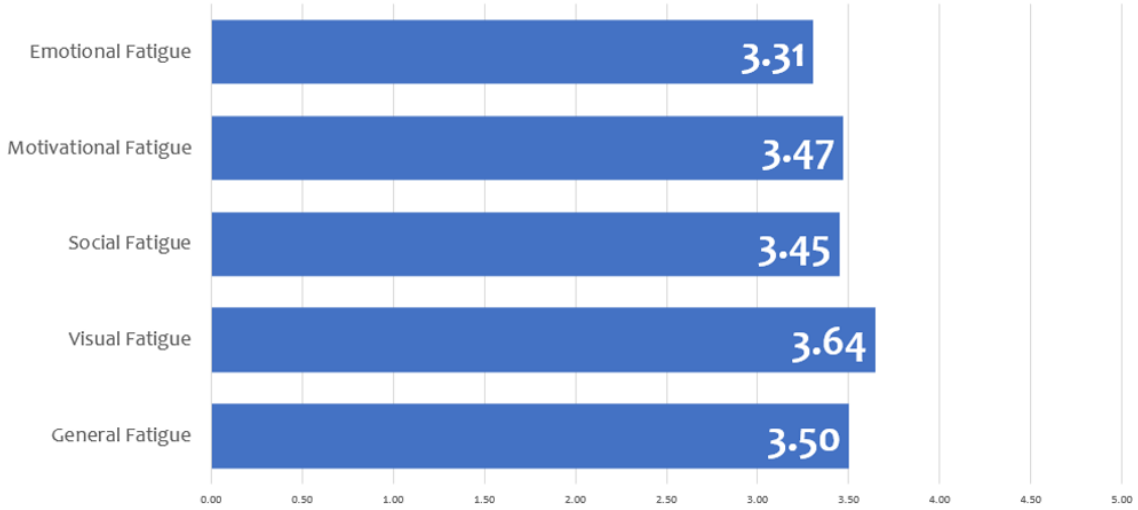


Figure 3: Self-reported levels of ZEF among PETE students

Table 3 shows that there are no statistically significant differences in the self-reported individual ZEF levels of PETE students when grouped according to their sex and year level. All PETE students' self-reported ZEF levels when grouped according to sex are the same in terms of their general ($U = 10{,}263.5$; $p = 0.757$), visual ($U = 9{,}287.5$; $p = 0.097$), social ($U = 9{,}442$; $p = 0.149$), motivational ($U = 9{,}288$; $p = 0.097$), and emotional exhaustion and fatigue ($U = 9{,}160$; $p = 0.067$). This is contrary to previous results by other studies that suggest women, in particular, have higher levels of Zoom-related fatigue and exhaustion compared to their counterparts (Purba et al., 2022; Fauville et al., 2023; Oducado et al., 2022; Usta Kara & Esroy, 2022). When grouped according to their year level, PETE students also reported higher levels of general ($x^2 = 5.975$; $p = 0.113$), visual ($x^2 = 3.652$; $p = 0.302$), social ($x^2 = 0.231$; $p = 0.972$), motivational ($x^2 = 3.753$; $p = 0.289$), and emotional exhaustion and fatigue ($x^2 = 2.088$; $p = 0.554$).

	ZEF by sex (*df* = 1)				
	General	Visual	Social	Motivational	Emotional
Mann-Whitney *U*	10,263.5	9,287.5	9,442.0	9,288.0	9,160.0
p Value	0.757	0.097	0.149	0.097	0.067
Z Score	-0.309	-1.660	-1.444	-1.659	-1.834
	ZEF by Year Level (*df* = 3)				
Kruskal-Wallis H (x^2)	5.975	3.652	0.231	3.753	2.088
p Value	0.113	0.302	0.972	0.289	0.554

Table 3: Test of difference on PETE student's zoom fatigue

CONCLUSION AND RECOMMENDATIONS

The study was conceived to assess the physical education teacher education students' self-reported ZEF levels and determine whether differences exist between males and females and the class level they were in at the time of the study. It was concluded that by and large, students in the PETE program experienced moderate to very tired and exhausting levels of fatigue whenever they participated in synchronous online classes, such as attendance to a video teleconferencing platform. Evidence derived from the study also indicated visual fatigue with the highest mean score, indicating that students feel very tired or exhausted when

their eyes are exposed to a certain extent in the device they use to attend virtual synchronous classes. This is coherent with the findings of other studies that other than attending virtual classes, students are also exposed to extended periods as they usually use their devices for many things other than learning (e.g., socialization and entertainment).

The findings of this study may have direct implications for the physical education teacher education program and, therefore, cannot be generalized to other teacher education students. The context may vary, considering that PETE students have relatively more movement-based courses than the other education programs, given the nature of physical education as a performance-based program. Methodologically, the context was also limited to the fact that the cross-sectional survey was facilitated as a design of the study. While the design may establish comparisons between variables, it cannot determine causality associated with the phenomenon. Moreover, it cannot track temporal changes since the data collection was limited to a specific point in time only, particularly at the height of the pandemic. At the same time, classes were transitioned to full online modality. Also, cohort effects were likewise viewed as potential confounding variables that were not included in the study as age-associated differences, because of historical and social contexts, may influence changes in the results, hence making the results particular to the context of the PETE program only.

Many aspects of the ZEF still need further exploration. For example, due to the late implementation and resumption of online learning in the Philippines, many previous studies conducted might not apply to the country, not to mention that the country has very different economic, infrastructural, and cultural aspects that can also affect the ZEF levels of the students. Moreover, the country also implemented a very different and more flexible type of learning that tried to accommodate all types of learners. The same might be true when similar studies are conducted on teachers in different regions of the country. Further studies are also warranted to understand how ZEF affects teachers, particularly in those in public and private schools. An exploration of how specific mental aspects relate to their general health with regard to their culture and habits is worth exploring, either for students, teachers, or professionals in general.

ACKNOWLEDGMENT

This study was supported by the author's individual affiliations.

REFERENCES

Alibudbud, R. (2021) 'On online learning and mental health during the COVID-19 pandemic: Perspectives from the Philippines', *Asian Journal of Psychiatry*, Vol. 66, p. 102867. https://doi.org/10.1016/j.ajp.2021.102867

Alvarez, Abel, J. (2020) 'The phenomenon of learning at a distance through emergency remote teaching amidst the pandemic crisis', *Asian Journal of Distance Education*, Vol. 15, No. 1, pp. 144–153. https://doi.org/https://doi.org/10.5281/zenodo.3881529

Asurion (2019) *Americans check their phones 96 times a day (Press release)*, [Online], Available: https://www.asurion.com/press-releases/americans-check-their-phones-96-times-a-day/#:~:text=Despite%20our%20attempts%20to%20curb,%20tech%20care%20company%20Asurion1 [10 December 2023]

Balita C. (2024) 'Average daily time spent using online media in the Philippines in 3rd quarter 2021, by activity', *Statista*, [Online], Available: https://www.statista.com/statistics/803812/daily-time-spent-using-online-media-by-activity-philippines/

Bare, M. A. D., Castro, F. M. F., Quimio, J. L. G. and J, M. J. (2023) 'Perceived Effect of Academic Stress to the Performance of Students during Online Class', The 12th Annual International Conference on Industrial Engineering and Operations Management, pp. 3693–3704. https://doi.org/10.46254/an12.20220673

Baticulon, R. E., Sy, J. J., Alberto, N. R. I., Baron, M. B. C., Mabulay, R. E. C., Rizada, L. G. T., Tiu, C. J. S., Clarion, C. A. and Reyes, J. C. B. (2021) 'Barriers to Online Learning in the Time of COVID-19: A National Survey of Medical Students in the Philippines', *Medical Science Educator*, Vol. 31, No. 2, pp. 615–626. https://doi.org/10.1007/s40670-021-01231-z

Bautista, Angelito P., J., Bleza, D. G., Buhain, C. B. and Balibrea, D. M. (2021) 'School Support Received and the Challenges Encountered in Distance Learning Education by Filipino Teachers during the Covid-19 Pandemic', *International Journal of Learning, Teaching and Educational Research*, Vol. 20, No. 6, pp. 360–385. https://doi.org/10.26803/ijlter.20.6.19

Bird, K. A., Castleman, B. L. and Lohner, G. (2022) 'Negative Impacts From the Shift to Online Learning During the COVID-19 Crisis: Evidence From a Statewide Community College System', *AERA Open*, Vol. *8*, p. 23328584221081220. https://doi.org/10.1177/23328584221081220

Bird, L. J., McCabe, M., Lim, Y. Y. and Cornish, K. (2023) 'Prevalence and correlates of subjective cognitive concerns in Australian university students during the COVID-19 pandemic', *Frontiers in Psychology,* Vol. 13. https://doi.org/10.3389/fpsyg.2022.1094497

Blair, A. (2016) 'Understanding first-year students' transition to university: A pilot study with implications for student engagement, assessment, and feedback', *Politics*, Vol. 37, No. 2, pp. 215–228. https://doi.org/10.1177/0263395716633904

Bolatov, A. K., Seisembekov, T. Z., Askarova, A. Z., Baikanova, R. K., Smailova, D. S. and Fabbro, E. (2021) 'Online-Learning due to COVID-19 Improved Mental Health Among Medical Students', *Medical Science Educator*, Vol. 31, No. 1, pp. 183–192. https://doi.org/10.1007/s40670-020-01165-y

Christiansen, D. M., McCarthy, M. M. and Seeman, M. V. (2022) 'Where Sex Meets Gender: How Sex and Gender Come Together to Cause Sex Differences in Mental Illness', *Frontiers in Psychiatry,* Vol. 13. https://doi.org/10.3389/fpsyt.2022.856436

Cislaghi, B. and Heise, L. (2020) 'Gender norms and social norms: differences, similarities and why they matter in prevention science', *Sociology of Health & Illness*, Vol. 42, No. 2, pp. 407–422. https://doi.org/10.1111/1467-9566.13008

Dacillo, M. J. F., Dizon, J. K. M., Ong, E. J. T., Pingol, A. M. L. and Cleofas, J. V. (2022) 'Videoconferencing fatigue and online student engagement among Filipino senior high school students: A mixed methods study', *Frontiers in Education,* Vol. 7. https://doi.org/10.3389/feduc.2022.973049

De Oliveira Kubrusly Sobral, J. B., Lima, D. L. F., Lima Rocha, H. A., de Brito, E. S., Duarte, L. H. G., Bento, L. B. B. B. and Kubrusly, M. (2022) 'Active methodologies association with online learning fatigue among medical students', *BMC Medical Education*, Vol. *22*, No. 1, p. 74. https://doi.org/10.1186/s12909-022-03143-x

Ellemers, N. (2018) 'Gender Stereotypes', *Annual Review of Psychology*, Vol. 69, No. 1, pp. 275–298. https://doi.org/10.1146/annurev-psych-122216-011719

Fauville, G., Luo, M., Queiroz, A. C. M., Bailenson, J. N. and Hancock, J. (2021) 'Zoom Exhaustion & Fatigue Scale', *Computers in Human Behavior Reports*, Vol. 4, p. 100119. https://doi.org/10.1016/j.chbr.2021.100119

Fauville, G., Luo, M., Queiroz, A. C. M., Lee, A., Bailenson, J. N. and Hancock, J. (2023) 'Video-conferencing usage dynamics and nonverbal mechanisms exacerbate Zoom Fatigue, particularly for women', *Computers in Human Behavior Reports*, Vol. 10, p. 100271. https://doi.org/10.1016/j.chbr.2023.100271

Ganne, P., Najeeb, S., Chaitanya, G., Sharma, A. and Krishnappa, N. C. (2021) 'Digital Eye Strain Epidemic amid COVID-19 Pandemic – A Cross-sectional Survey', *Ophthalmic Epidemiology*, Vol. *28*, No. 4, pp. 285–292. https://doi.org/10.1080/09286586.2020.1862243

George, J., Mirsadikov, A., Nabors, M. and Marett, K. (2022) 'What do Users Actually Look at During 'Zoom' Meetings? Discovery Research on Attention, Gender and Distraction Effects', *Proceedings of the 55th Hawaii International Conference on System Sciences*, pp. 4779–4787. https://doi.org/10.24251/hicss.2022.582

Gonzales, G. (2019) 'Filipinos spend most time online, on social media worldwide – report', *Rappler*, [Online], Available: https://www.rappler.com/technology/222407-philippines-online-use-2019-hootsuite-we-are-social-report/

Gonzalez-Ramirez, J., Mulqueen, K., Zealand, R., Silverstein, S., Mulqueen, C. and BuShell, S. (2021) 'Emergency Online Learning: College Students' Perceptions During the COVID-19 Pandemic', *College Student Journal*, Vol. 55, No. 1, pp. 29–46.

Grace, M. K. (2019) 'Parting ways: Sex-based differences in premedical attrition', *Social Science & Medicine*, Vol. 230, pp. 222–233. https://doi.org/10.1016/j.socscimed.2019.04.030

Graves, B. S., Hall, M. E., Dias-Karch, C., Haischer, M. H. and Apter, C. (2021) 'Gender differences in perceived stress and coping among college students', *PLOS ONE*, Vol. 16, No. 8, p. e0255634. https://doi.org/10.1371/journal.pone.0255634

Henritius, E., Löfström, E. and Hannula, M. S. (2019) 'University students' emotions in virtual learning: a review of empirical research in the 21st century', *British Journal of Educational Technology*, Vol. *50*, No. 1, pp. 80–100. https://doi.org/10.1111/bjet.12699

Honkimäki, S. and Kálmán, O. (2012) '*Approaches to Transition Support for First Year Students in Higher Education', in* Tynjälä, P., Stenström, M. L. and Saarnivaara M.(ed) *Transitions and Transformations in Learning and Education,* Amesterdam: Springer. https://doi.org/10.1007/978-94-007-2312-2_15

Iannizzotto, G., Nucita, A., Fabio, R. A., Caprì, T. and Lo Bello, L. (2020) 'Remote Eye-Tracking for Cognitive Telerehabilitation and Interactive School Tasks in Times of COVID-19', *Information*, Vol. 11, No. 6, p. 296. https://doi.org/10.3390/info11060296

Jung, J., Horta, H. and Postiglione, G. A. (2021) 'Living in uncertainty: the COVID-19 pandemic and higher education in Hong Kong', *Studies in Higher Education*, Vol. 46, No. 1, pp. 107–120. https://doi.org/10.1080/03075079.2020.1859685

Khalil, R., Mansour, A. E., Fadda, W. A., Almisnid, K., Aldamegh, M., Al-Nafeesah, A., Alkhalifah, A. and Al-Wutayd, O. (2020) 'The sudden transition to synchronized online learning during the COVID-19 pandemic in Saudi Arabia: a qualitative study exploring medical students' perspectives', *BMC Medical Education*, Vol. 20, No. 1, p. 285. https://doi.org/10.1186/s12909-020-02208-z

Labrague, L. (2013) 'Stress, Stressors, and Stress Responses of Student Nurses in a Government Nursing School', *Health Science Journal*, Vol. 7, No. 4, pp. 424–435.

Li, J., Zhou, L., Van Der Heijden, B., Li, S., Tao, H. and Guo, Z. (2021) 'Social Isolation, Loneliness and Well-Being: The Impact of WeChat Use Intensity During the COVID-19 Pandemic in China', *Frontiers in Psychology*, Vol. 12. https://doi.org/10.3389/fpsyg.2021.707667

Lim, L. T. S., Regencia, Z. J. G., Dela Cruz, J. R. C., Ho, F. D. V, Rodolfo, M. S., Ly-Uson, J. and Baja, E. S. (2022) 'Assessing the effect of the COVID-19 pandemic, shift to online learning, and social media use on the mental health of college students in the Philippines: A mixed-method study protocol', *PLOS ONE*, Vol. 17, No. 5, p. e0267555. https://doi.org/10.1371/journal.pone.0267555

Little, M., Lateef-Vaksvik, A., Yu, B. and Kunkonlakarn, P. (2021) *Engaging Allard Law Students in Virtual Wellbeing Programming*, [Online], Available: https://open.library.ubc.ca/collections/18861/items/1.0398397 [9 December 2023]

Mahmoud, J. S. R., Staten, R. "Topsy," Lennie, T. A. and Hall, L. A. (2015) 'The Relationships of Coping, Negative Thinking, Life Satisfaction, Social Support, and Selected Demographics With Anxiety of Young Adult College Students', *Journal of Child and Adolescent Psychiatric Nursing*, Vol. 28, No. 2, pp. 97–108. https://doi.org/10.1111/jcap.12109

Martínez-Líbano, J., Yeomans, M. M., GonzálezCampusano, N. and Campos Flores, E. (2021) 'Emotional Exhaustion and Mental Health in a Sample of Chilean Social Science Students During the COVID-19 Pandemic', *PSOCIAL*, Vol. 7, No. 2.

Maymon, R., Hall, N. C. and Harley, J. M. (2019) 'Supporting first-year students during the transition to higher education: The importance of quality and source of received support for student well-being', *Student Success*, Vol. 10, No. 3, pp. 64–75. https://doi.org/10.5204/ssj.v10i3.1407

Meehan, C. and Howells, K. (2018) 'What really matters to freshers?': evaluation of first year student experience of transition into university', *Journal of Further and Higher Education*, Vol. 42, No. 7, pp. 893–907. https://doi.org/10.1080/0309877x.2017.1323194

Nagata, J. M., Cortez, C. A., Cattle, C. J., Ganson, K. T., Iyer, P., Bibbins-Domingo, K. and Baker, F. C. (2022) 'Screen Time Use Among US Adolescents During the COVID-19 Pandemic: Findings From the Adolescent Brain Cognitive Development (ABCD) Study', *JAMA Pediatrics*, Vol. 176, No. 1, pp. 94–96. https://doi.org/10.1001/jamapediatrics.2021.4334

Newsome, J., Vaske, J. C., Gehring, K. S. and Boisvert, D. L. (2016) 'Sex Differences in Sources of Resilience and Vulnerability to Risk for Delinquency', *Journal of Youth and Adolescence*, Vol. 45, No. 4, pp. 730–745. https://doi.org/10.1007/s10964-015-0381-2

Nyar, A. (2021) 'The 'Double Transition' for First-Year Students: Understanding the Impact of Covid-19 on South Africa's First-Year University Students', *Journal for Students Affairs in Africa*, Vol. 9, No. 1, pp. 77–92. https://doi.org/10.24085/jsaa.v9i1.1429

Oducado, R. M. F., Dequilla, M. A. C. V. and Villaruz, J. F. (2022) 'Factors predicting videoconferencing fatigue among higher education faculty', *Education and Information Technologies*, Vol. 27, No. 7, pp. 9713–9724. https://doi.org/10.1007/s10639-022-11017-4

Ortega, H. C., Castro, R., Tolentino, J. C., Pusung, D. S. and Abad, R. (2022) 'The hidden curriculum in a Filipino pre-service physical educators' virtual ecology', *Edu Sportivo: Indonesian Journal of Physical Education,* Vol. 27, No. 7, pp. 25–40. https://doi.org/10.25299/es:ijope.2022.vol3(1).8851

Pandya, A. and Lodha, P. (2021) 'Social Connectedness, Excessive Screen Time During COVID-19 and Mental Health: A Review of Current Evidence', *Frontiers in Human Dynamics*, Vol. 3. https://doi.org/10.3389/fhumd.2021.684137

Peper, E., Wilson, V., Martin, M., Rosegard, E. and Harvey, R. (2021) 'Avoid Zoom Fatigue, Be Present and Learn', *NeuroRegulation*, Vol. 8, No. 1, pp. 47–56. https://doi.org/10.15540/nr.8.1.47

Purba, D. S., Sijabat, F. and Amila. (2022) 'Stress, Zoom Fatigue, and Learning Outcomes of Bachelor of Nursing Students in The Covid-19 Pandemic', *International Journal of Nursing and Health Services (IJNHS)*, Vol. 5, No. 6, pp. 551–558.

Räisänen, M., Postareff, L., Mattsson, M. and Lindblom-Ylänne, S. (2018) 'Study-related exhaustion: First-year students' use of self-regulation of learning and peer learning and perceived value of peer support', *Active Learning in Higher Education*, Vol. 21, No. 3, pp. 173–188. https://doi.org/10.1177/1469787418798517

Rasouli, J. J., Shin, J. H., Than, K. D., Gibbs, W. N., Baum, G. R. and Baaj, A. A. (2020) 'Virtual Spine: A Novel, International Teleconferencing Program Developed to Increase the Accessibility of Spine Education During the COVID-19 Pandemic', *World Neurosurgery*, Vol. 140, pp. e367–e372. https://doi.org/10.1016/j.wneu.2020.05.191

Ratan, R., Miller, D. B. and Bailenson, J. N. (2021) 'Facial Appearance Dissatisfaction Explains Differences in Zoom Fatigue', *Cyberpsychology, Behavior, and Social Networking*, Vol. 25, No. 2, pp. 124–129. https://doi.org/10.1089/cyber.2021.0112

Rotas, E. E. and Cahapay, M. B. (2020) 'Difficulties in Remote Learning: Voices of Philippine University Students in the Wake of COVID-19 Crisis', *Asian Journal of Distance Education*, Vol. 15, No. 2, pp. 147–158. https://doi.org/10.5281/zenodo.4299835

Salim, J., Tandy, S., Arnindita, J. N., Wibisono, J. J., Haryanto, M. R. and Wibisono, M. G. (2022) 'Zoom fatigue and its risk factors in online learning during the COVID-19 pandemic', *Medical Journal of Indonesia*, Vol. 31, No. 1, pp. 13–19 https://doi.org/10.13181/mji.oa.225703

Scott-Young, C. M., Turner, M. and Holdsworth, S. (2020) 'Male and female mental health differences in built environment undergraduates', *Construction Management and Economics*, Vol. 38, No. 9, pp. 789–806. https://doi.org/10.1080/01446193.2020.1748213

Shockley, K. M., Gabriel, A. S., Robertson, D., Rosen, C. C., Chawla, N., Ganster, M. L. and Ezerins, M. E. (2021) 'The fatiguing effects of camera use in virtual meetings: A within-person field experiment', *Journal of Applied Psychology*, Vol. 106, No. 8, pp. 1137–1155. https://doi.org/10.1037/apl0000948

Sundarasen, S., Chinna, K., Kamaludin, K., Nurunnabi, M., Baloch, G. M., Khoshaim, H. B., Hossain, S. F. and Sukayt, A. (2020) 'Psychological Impact of COVID-19 and Lockdown among University Students in Malaysia: Implications and Policy Recommendations', *International Journal of Environmental Research and Public Health*, Vol. 17, No. 17, p. 6206. https://doi.org/10.3390/ijerph17176206

Usta Kara, I. and Esroy, E. G. (2022) 'A New Exhaustion Emerged with COVID-19 and Digitalization: A Qualitative Study on Zoom Fatigue', *OPUS Toplum Araştırmaları Dergisi*, Vol. 19, No. 46, pp. 365–379. https://doi.org/10.26466/opusjsr.1069072

Full research paper

THE IMPACT OF THE LEARNING ENVIRONMENT ON SELF-EFFICACY AND ACHIEVEMENT GOALS OF ISRAELI PRE-SERVICE TEACHERS

Marcel Amasha

Tel Hai Academic College, Israel

✉ marcelam@telhai.ac.il

ABSTRACT

Advancing public education by improving the skills and knowledge of its teachers is a major challenge. The teacher-training phase shapes not only skills and abilities but also perceptions of pre-service teachers regarding their educational and teaching goals. We examined a hypothetical theoretical model that explains how pre-service teachers' perceptions of their own learning environment affects belief in their self-efficacy in teaching, and how this shape their achievement goals in teaching as future educators. The study included 278 pre-service teachers studying at all five colleges in our country that offer teacher training programs. Existing questionnaires were adapted to the study population and underwent structure validation. The hypothesized structural model was deemed a good fit for the data and was able to explain 35% of variance in the mastery goals of pre-service teachers, 24% of variance in performance-approach goals, and 65% of variance in performance-avoidance goals. The structural model shows that perception of the learning environment has a strong and significant impact on teaching ability and the achievement goals of pre-service teachers. Fostering a constructivist learning environment in teacher training colleges may increase belief in self-efficacy in teaching and enable pre-service teachers to adopt teaching control goals.

KEYWORDS

Achievement goals, learning environment, pre-service teachers, self-efficacy, teaching colleges

HOW TO CITE

Amasha M. (2024) 'The Impact of the Learning Environment on Self-efficacy and Achievement Goals of Israeli Pre-service Teachers', *Journal on Efficiency and Responsibility in Education and Science*, vol. 17, no. 3, pp. 214–222. http://dx.doi.org/10.7160/eriesj.2024.170304

Article history
Received
November 24, 2022
Received in revised form
August 16, 2023
Accepted
August 26, 2024
Available on-line
September 30, 2024

Highlights

- *Pre-service teachers' perceptions of their own learning environment affect belief in their self-efficacy.*
- *Belief in self-efficacy can shape pre-teachers' achievement goals as future educators.*
- *A constructivist learning environment in teacher training colleges may increase belief in self-efficacy.*

INTRODUCTION

Teacher Training and Goal Theory

Teacher motivation towards orientation-achievement goals is critical since these directly influence both teaching and learning processes at large (Butler & Shibaz, 2008; Fasching et al., 2010; Retelsdorf et al., 2010). In the same respect, the orientation of achievement goals also confers particular repercussions on the dimensions of teachers' own professional development (Laine & Gegenfurtner, 2013; Minnaert et al., 2011; Segers & Gegenfurtner, 2013; Volet, 2013).

The prior assumption's main dimension and domain of application are the field of education at large, with its numerous aspects, multifaceted contexts, and particularly its strata, ranging from basic education institutions (Polychroni et al., 2012) to middle school, secondary school and up to the academic institutions (Bipp & Spinath, 2012) and teacher training (Nitsche et al., 2013). Butler's (2007) goal orientation theory comes up an effective scrutinizing tool via which achievement goals of pre-service teachers can be closely examined and reformulated while still at the preliminary training stage.

The primary conceptualization of achievement goals was a dichotomous model that set apart mastery and performance goals. Mastery goals are mainly concerned with the inner value of learning, the effective and constructive employment and implementation of efforts, the systematic development and emergence of mastery and skills, and enriching and reinforcing the learning of new skills and techniques. Conversely, performance goals have an obvious ad tangible

distinctive feature that is focusing on the individual's sense of self-esteem, determined by how one conceives their own capacity to perform. Individuals determine their capacities and capabilities in contrast with others, with the goal of exhibiting and providing tangible proof of their more distinguished abilities (Ames, 1992; Nicholls et al., 1985).

The predisposition in research, however, over the years in connection with teachers' achievement goals was the employment of the tri-chotomous model (Cho & Shim, 2013; Daniels et al., 2013; Kucsera et al., 2011; Van Daal et al., 2014). According to the previously mentioned model model, teachers' achievement goals are mastery goals, the yearning to gain and build professional skills (performance-approach goals), and the aspiration to exude a high-level capacity in comparison with the other teachers (performance-avoidance goals).

Based on the finding of prominent figures in the field, namely; (Aarts & Elliot, 2012; Elliot & Fryer, 2008; Gollwitzer & Oettingen, 2012), achievement goals are in reality *ability-dominant* goals that directly influence and may guide achievement behavior. These goals are mostly situational, contextual and attainable (Hagenauer & Hascher, 2010; Minnaert et al., 2011).

There is a well-established connection between aims, motivation, self-efficacy, and perseverance against difficulties (Elliot & Church, 1997; Patrick, Ryan & Kaplan, 2007). For instance, mastery goals have a positive correlation with self-efficacy (Nitsche et al., 2011), and may help to reduce teacher burnout, and positively impact teachers tendencies to seek helping hand in developing their capabilities (Butler, 2007). However as (Papaioannou & Christodoulidis, 2007) pointed out, this is not applicable for all kinds of goals, and while there is a positive connection between job satisfaction and mastery goals among teachers, no such connection is detected with performance goals, whether approach or avoidance. In addition to the prior, (Cho & Shim, 2013) discovered medium-positive relation between teaching efficacy on one hand, and mastery-approach goals on the other hand, and low-positive correlation between teaching efficacy and performance-approach goals, but no correlation whatsoever between teaching efficacy and performance-avoidance goals. Similarly, (Yildizlli, 2019) exuded no correlation between performance-approach goals and teachers' self-efficacy and burnout.

Mastery and performance goals promote numerous behaviors and methods of teaching. For instance, teachers who are in fact characterized by mastery goals aspire to advance their professional skills in teaching, and to provide support and feedback to their students. They motivate students to ask questions, examine, and scrutinize current and general situations; encouraging them to acquired better-shaped thinking skills. In comparison, teachers who are more focused on performance goals, aspire to exude transcending teaching abilities or *even* to veal inferior ones (Butler & Shibaz, 2008; Retelsdorf et al., 2010). In the same context, pre-service teachers who endorse mastery goals within the classroom environment are in fact more focused on their individual progress; by actively implementing *cognitive and metacognitive self-guided learning techniques* (Liu et al., 2019).

According to (Butler, 2007), teachers' achievement goals forego their teaching activities; consequently, the ultimate production of pre-service teachers as future educators may be anticipated and influenced by formulating their objectives, while still within the training phase. Besides, the learning atmosphere has a powerful and notable impact on the motives and achievement objectives that student teachers endorse (Kaplan & Maehr, 2007; Yıldızlı et al., 2016). Therefore, an environment that stresses the significance of effort and investment, possession of skills, individual growth, and proper assimilation of school assignments is expected to profoundly impact students to pursue mastery goals (Gonida et al., 2009. contrariwise, when the emphasis within the teaching environment is on grades, exterior consolidation, and social comparison, it might be more probable that students will endorse performance goals alternatively (Meece et al., 2006).

The social-cognitive approach to education perceives the learning atmosphere as a vital background to enhancing self-efficacy in the current learning effort (Ames & Archer, 1988; Bell & Kozlowski, 2002; Zimmerman, 1990; Zimmerman & Martinez-Pons, 1992). Within the same respect of teacher-training programs, the learning atmosphere also possesses a vital role in the development of self-efficacy of the actively involved pre-service teachers (Romi & Leyser, 2006), thus influencing which type of achievement goals prospective teachers endorse (Deemer, 2004; Wolters & Daugherty, 2007). Numerous researches on this subject have established a positive correlation between self-efficacy and mastery goals among students (Bong, 2001; Gerhardt & Brown, 2006), where students with a more notable sense of self-efficacy are more likely to endorse mastery goals than those students with less prominent efficacy.

One prominent element of teacher training is the learning atmosphere within the classroom, that directly impacts the achievement goals endorsed by the concerned student teacher (Kaplan & Maehr, 2007). A class that stresses mastery and learning goals, the significance of abtaining skills, investing effort into the assigned tasks, assimilating the school assignments, the individual's personal growth and progress, creating a learning atmosphere where it is highly probable that student teachers will endorse such objectives, inspiring them to implement and embloy efficient learning techniques, undertaking more difficult tasks, perseverance despite obvious and constant challenges (Gonida et al., 2009). Adversely, a classroom environment whose emphasis is on the importance of grades, external reinforcement and social comparison, will probably lead its students to adopt performance goals instead (Meece et al., 2006).

Teachers' self-efficacy is elucidated as their confidence in their own personal qualifications to employ particular teaching and learning tasks within the classroom walls, and to push on their students' achievements (Dellinger, 2001). One's beliefs concerning self-efficacy are the results and direct production of a cognitive and meta-cognitive process that depend on four sources: the performances and individual experiences of the individual; experties based on the observation of other individuals' behavior; verbal persuasion; and physiological and emotional reactions. The priors are expected to impact people's confidence in their ability to realize their full potential (Bandura, 1997; Chen & Usher, 2013; Usher & Pajares, 2008). Fruitful experiences at an early phase of training could assist pre-service teachers to cope better when they commence their first year of actual teaching, while early substandard experience

during training might dishearten pre-service teachers from the teaching career (Hoy & Spero, 2005).

Pre-service teachers' belief in their teaching self-efficacy as well as the trait of their achievement goals, may also guarantee successful management throughout their career with the fast rythm of innovations in professional knowledge that calls for uninterrupted adaptation of work techniques, starategies, and mechanisms. Thereupon, pre-service teachers' self-efficacy has been established to be linked to career commitment (Han et al., 2016; Klassen & Chiu, 2011), compulsion to obtaining the teaching degree (Pfitzner-Eden, 2016), and positively connected to students' academic achievement (Klassen & Tze, 2014).

Despite the fact some studies have established that self-efficacy in teaching varies according gender and year of study (Hoy & Spero, 2005), other researches found that the correlation was not of statistical significance (Kass & Miller, 2015). Cho & Shim (2013) arrived at the finding that female participants were more motivated than their male peers by mastery goals, while male participants were more motivated by performance goals. Besides, other studies found that, generally speaking, the achievement goals of pre-service teachers lessened in the course of their teacher training years (Fasching et al., 2010).

In conclusion, and based on then previously examined and analyzed literature review, it may be deduced that the experience of pre-service teachers during their studies, and their interaction with college lecturers regarding teaching, learning and assessment, may impact their self-efficacy in teaching and in the achievement goals they design for themselves and pursue as future teachers. Teacher-training colleges play a vital role in developing the self-efficacy of student teachers and in forming and formulating the achievement goals that will stand out in the classroom. The learning environment positively stresses teaching goals, increasing the self-efficacy of pre-service teachers, and reinforcing their propensity to set those goals. Therefore, the study focused on the following sole research question:

- To what extent can the achievement goals of pre-service teachers be predicted based on their perceptions of their college learning environment and their belief in self-efficacy in teaching?

METHODS

Study Participants

The sample included 278 pre-service teachers (231 female, 47 male) from five teacher-training colleges randomly sampled from all teacher-training colleges in our country. About 16% of the participants were first-year students, 41% second-year, 26.6% third-year, and 16.5% were in their fourth year, studying toward their degree in education. After receiving approval from the ethics committees of the five colleges, we arrived at the colleges and distributed the questionnaires to students who agreed to participate in the study. The data presented in this paper were collected from those completed questionnaires.

Research Instruments

The present research, conducted according to a quantitative-correlative approach, was based on one questionnaire of demographic data and three other questionnaires which had been translated into Hebrew and then back-translated into English as a control measure to ensure that the translation was true to the source.

Background data questionnaire: This questionnaire included the following variables: gender, age, study year, study subject.

'Preservice teacher's perception of the college learning environment' questionnaire: This questionnaire was comprised of 28 statements from the College and University Classroom Environment Inventory (CUCEI) and the Course Experience Questionnaire (CEQ). The two questionnaires have been considered reliable on the subject, as they examine experiencing the learning environment over a long period of time (Aldridge and Fraser, 2000). Items were rated on a Likert scale between 1 (lowest) and 5 (highest). Final scores were calculated using averages of the items included in each factor.

To check the validity of the structure of our questionnaire, which was adapted to the current study from the two questionnaires, exploratory factor analysis (EFA) was used in the first stage, and confirmatory factor analysis (CFA) in the second stage.

In the first stage, half of the research sample was selected randomly for the exploratory factor analysis (EFA). Four statements that had loadings of lower than 0.4 or that had a high loading on two or more factors at the same time (cross loading) were omitted. The exploratory analysis yielded a structure of six factors: good instruction, learning assignments, skill development, academic environment, traditional lecture, and appropriate assessment. The exploratory factor analysis succeeded in explaining 63.25% of the explained variance of the questionnaire.

In the second stage, a confirmatory factor analysis (CFA) was conducted on the remaining half of the sample. Its results fit the structure obtained in the exploratory factor analysis. The final questionnaire on the perceptions of the pre-service teachers regarding the learning environment at the college (and in the classroom) included 24 statements, divided into six different factors. its goodness-of-fit indexes were as follows: $\chi^2 = 404.586$, $df = 245$, $P = .000$; $\chi^2/df = 1.651$; SRMR = 0.053; CFI = .941 RMSEA = 0.048 (0.040, 0.057). The loadings on the variable of learning environment ranged from 0.36 to 0.90.

Achievement goals questionnaire: the preservice teachers' achievement goals were measured using the Butler's Goal Orientations for Teaching scale (2007) questionnaire which assesses the achievement goals of practicing teachers, adjusted for the present research population. The original questionnaire included four indices, of which only three were used in this study: mastery goals index, performance-approach goals index, performance-avoidance goals index. The fourth index—the goal of avoiding work—was found to be irrelevant to the present research and therefore those questions were not included in our questionnaire.

Each index in the questionnaire utilized in the present study was composed of four items, redefined to suit pre-service teachers. For example, the statement "I feel that a successful teaching day is when something occurs in the classroom that makes me want to deepen my professional knowledge" was replaced with "As a future teacher, I would feel successful if something occurred in the classroom that made me want to deepen my professional knowledge." the internal consistency estimates of reliability (Cronbach's alpha) scores for the different statements in each of the three indices of the current study ranged from $\alpha = 0.7$ to $\alpha = 0.78$.

Items were rated on a Likert scale between 1 (lowest) and 5 (highest). Final scores were calculated using averages of the items included in each factor.

Confirmatory factor analysis was used to examine whether the questionnaire structure was a good fit to the data of the present research population. $\chi^2 = 43.505$, $df = 24$, $P = .009$; $\chi^2/df = 1.813$; SRMR = 0.043; CFI = .979 RMSEA = 0.054 (0.027, 0.079).The factor loadings of the different achievement goals ranged from 0.53 to 0.92.

Teaching self-efficacy questionnaire: the current study used the Scale for Teacher Self-Efficacy (STSE) questionnaire. This questionnaire was developed for the actual teacher population and is also valid for the pre-service teaching population. The questionnaire includes 12 statements with response options from1-5, ranging from "not at all sure of my ability to do" to "completely sure of my ability to do" (Pfitzner-Eden et al., 2014). The questionnaire represents three dimensions: 1. Self-efficacy in implementing teaching strategies, 2. Self-efficacy in classroom management, and 3. Self-efficacy in the ability to involve students in learning.

Sample statements:

1. To what extent do you feel able to provide an alternative explanation or example when students are confused?
2. To what extent do you feel able to persuade students to follow classroom procedures?
3. To what extent do you feel you are able to help students think critically? (Develop critical thinking)

Confirmative factor analysis was used to examine the fit of the questionnaire structure to the data of the present research population. The structure of the three factors was confirmed. The goodness-of-fit indices of the final model obtained were $\chi^2 = 50.388$, $df = 31$; $P = .015$; $\chi^2/df = 1.625$; SRMR = 0.048; CFI = .978; RMSEA = 0.048 (0.021, 0.071). The loadings on the variable of self-efficacy in teaching ranged from 0.12 to 0.90, where the loadings of classroom management and ability to engage students in learning were high (0.8 and 0.9, respectively) compared with the low loading (0.12) of the factor of efficacy in employing teaching strategies.

Statistical Analysis

Data were entered and analyzed using SPSS version 28. First, descriptive statistics were produced using means and standard deviations for all variables. Reliabilities of the scales were evaluated by Cronbach Alpha, while their validities were estimated by Confirmatory Factor Analysis. Correlations between variables were assessed using Pearson correlations.

To assess the relationship between the independent variables and the dependent variable (self efficacy), path analysis using Structural Equation Modeling (SEM) was conducted. The following indices were used to evaluate the model: chi-squared, which is acceptable when the value is not significant; the goodness of fit index (GFI), the comparative fit index (CFI), and the non-normed fit index (NNFI), (adequate values - above 0.90, excellent fit - above 0.95); and the root mean square error of approximation (RMSEA) (adequate values - less than 0.08, excellent fit - less than 0.06) (Arbuckle, 2013). SEM was tested using AMOS software. Level of significance (p-value) was 5%.

RESULTS

Descriptive Statistics

The present research examined how pre-service teachers perceived their college learning environment, their belief in their teaching self-efficacy, and how these affected the prediction of their achievement goals as future teachers.

Table 1 presents the descriptive statistics of the six factors of the learning environment in teacher-training colleges. The results presented in Table 1 indicate that pre-service teachers agreed to a moderate degree that the essential components of the learning environment existed at the teacher-training colleges. They also agreed only to a relatively low degree that good instruction took place at the colleges at which they studied.

	Number of items	Reliability score (α)	Mean	Standard deviation
Student perceptions regarding good instruction in the college	4	0.83	2.06	1.07
Student perceptions regarding development of skills	4	0.85	3.16	.77
Student perceptions regarding academic environment	4	0.78	3.16	.82
Student perceptions regarding quality of assessment	3	0.77	3.16	.83
Student perceptions regarding quality of study assignments	3	0.78	2.96	.80
Student perceptions regarding traditional instruction at the college	5	0.63	3.38	.97

Table 1: Description of Factors of Perceptions of the Learning Environment Among Pre-service Teachers (Scale of 1-5, where 1 is lowest and 5 highest)

Table 2 presents the descriptive statistics of the three factors of self-efficacy in teaching of the pre-service teachers. The results presented in Table 2 indicate that the pre-service teachers' beliefs in their self-efficacy in teaching were at a moderate level. The level of their self-efficacy in employing teaching strategies was low-moderate and the level of their self-efficacy in managing the classroom and engaging their students in learning was moderate-high.

Table 3 presents the descriptive statistics of the factors of achievement goals in teaching of the preservice teachers. The results indicate that declared achievement goals of the preservice teachers as future teachers were relatively high regarding mastery goals and performance-approach goals, while the mean for performance-avoidance goals was of a moderate level.

	Number of items	Reliability score (α)	Mean	Standard deviation
Efficacy in teaching strategies	4	0.77	2.68	.60
Efficacy in classroom management	3	0.74	3.63	.68
Efficacy in engaging students	3	0.78	3.65	.74

Table 2: Factors of Self-Efficacy in Teaching of Preservice Teachers (Scale of 1-5, where 1 is lowest and 5 highest)

	Number of items	Reliability score (α)	Mean	Standard deviation
Mastery goals	3	0.77	3.70	.67
Performance-approach goals	4	0.74	3.94	.78
Performance-avoidance goals	3	0.78	3.10	.77

Table 3: Factors of Achievement Goals in Teaching of Preservice Teachers (Scale of 1-5)

Pearson Correlations between Study Variables

Table 4 presents the Pearson correlations between the main study variables. Efficacy in teaching strategies were positively correlated to learning environment factors ($.237 < r < .461$). In addition, efficacy in classroom management was positively related with learning environment factors (besides with good instruction in the college) ($.231 < r < .370$). similarly, efficacy in engaging students was positively related with learning environment factors (besides with good instruction in the college) ($.231 < r < .366$). Finally, Efficacy in engaging students was positively related with learning environment factors (besides with good instruction in the college) ($.168 < r < .225$).

Mastery goals were positively related to learning environment factors besides with good instruction in the college ($.226 < r < .311$), efficacy in classroom management ($r = .600$, $p < .01$) and Efficacy in engaging students ($r = .577$, $p < .01$).

Performance approach goals was negatively related to the student perceptions regarding good instruction in the college ($r = -.218$, $p < .01$), but positively related to the other learning environment factors ($.132 < r < .225$), and also positively related to efficacy in classroom management ($r = .346$, $p < .01$) and efficacy in engaging students ($r = .185$, $p < .01$).

Positively related to learning environment factors besides with good instruction in the college ($.226 < r < .311$), efficacy in classroom management ($r = .600$, $p < .01$) and Efficacy in engaging students ($r = .577$, $p < .01$).

In addition, a positive, significant correlation was found between learning environment at the college and self-efficacy in teaching ($r = .38$; $p < .001$), mastery goals for teaching ($r = .269$; $p < .001$), performance-approach goals ($r = .224$; $p < .001$), and performance-avoidance goals ($r = .317$; $p < .001$). The more positive the pre-service teacher's perception of the learning environment, the higher his/her self-efficacy in teaching, as well as the level of his/her achievement goals in teaching. Positive correlations were also found between self-efficacy in teaching and mastery goals for teaching ($r = .522$; $p < .001$), performance-approach goals in teaching ($r = .275$; $p < .001$), and performance-avoidance goals in teaching ($r = .294$; $p < .001$). The greater the self-efficacy in teaching of the preservice teachers, the higher their achievement goals. The correlations between the different achievement goals indicated a positive correlation between mastery goals for teaching and performance-approach goals ($r = .425$; $p < .001$), and between mastery goals for teaching and performance-avoidance goals ($r = .339$; $p < .001$) the higher the mastery goals, the higher the performance-approach and performance-avoidance goals. A positive, strong, and statistically significant correlation was found between performance-approach and performance-avoidance goals ($r = .778$; $p < .001$). The higher the performance-approach goals, the higher the performance-avoidance goals.

	1	2	3	4	5	6	7	8	9	10	11
Student perceptions regarding good instruction in the college											
Student perceptions regarding development of skills	.120*										
Student perceptions regarding academic environment	.127*	.648**									
Student perceptions regarding quality of assessment	.323**	.675**	.659**								
Student perceptions regarding quality of study assignments	.167**	.631**	.698**	.672**							
Student perceptions regarding traditional instruction at the college	.146*	.515**	.541**	.678**	.601**						
Efficacy in teaching strategies	.461**	.237**	.326**	.450**	.386**	.292**					
Efficacy in classroom management	-.042	.366**	.302**	.231**	.370**	.285**	.089				
Efficacy in engaging students	-.027	.225**	.205**	.168**	.242**	.195**	.082	.464**			
Mastery goals	-.100	.300**	.311**	.206**	.350**	.226**	.078	.600**	.577**		
Performance-approach goals	-.218**	.264**	.288**	.170**	.272**	.208**	.010	.414**	.321**	.388**	
Performance-avoidance goals	.001	.137*	.173**	.132*	.225**	.175**	.049	.346**	.185**	.303**	.474**

Table 4: Pearson correlations between the main study variables

Structural Model of Achievement Goals of Pre-service Teachers

The independent variables in the model were self efficacy and learning environment while the dependent variables were mastery goals, performance-approach goals, and performance-avoidance goals. Results showed acceptable goodness of fit indices χ^2 = 1207.95, *df* = 697, *P* = 0.001; χ^2/df = 1.733; SRMR = 0.061; CFI = 0.90; RMSEA = 0.05 (0.04, 0.06). The model examined data while controlling for background variables (gender; age, studying year). (see Figure 1: Pre-service achievement goals model).

The structural model showed that self-efficacy in teaching had a strong positive direct effect on the mastery goals of the pre-service teachers (β = .89, p < .001), and also a moderate positive direct effect on the performance-approach of the pre-service teachers (β = .32, p < .01). No significant direct effect was found between self-efficacy in teaching and avoidance goals (β = -.05, p = .89).

Hence, the preservice teachers who were characterized by high self-efficacy in teaching set mastery goals for themselves and aspired to develop their teaching abilities. However, they were also likely to set performance-approach but not performance-avoidance goals for themselves.

In addition, results also showed that college learning environment has a strong positive effect on goal avoidance (β = .88, p < .001), a moderate positive effect on goal-approach (β = .20, p < .01). No significant effect was found between college learning environment and mastery goals (β = .04, p = .75).

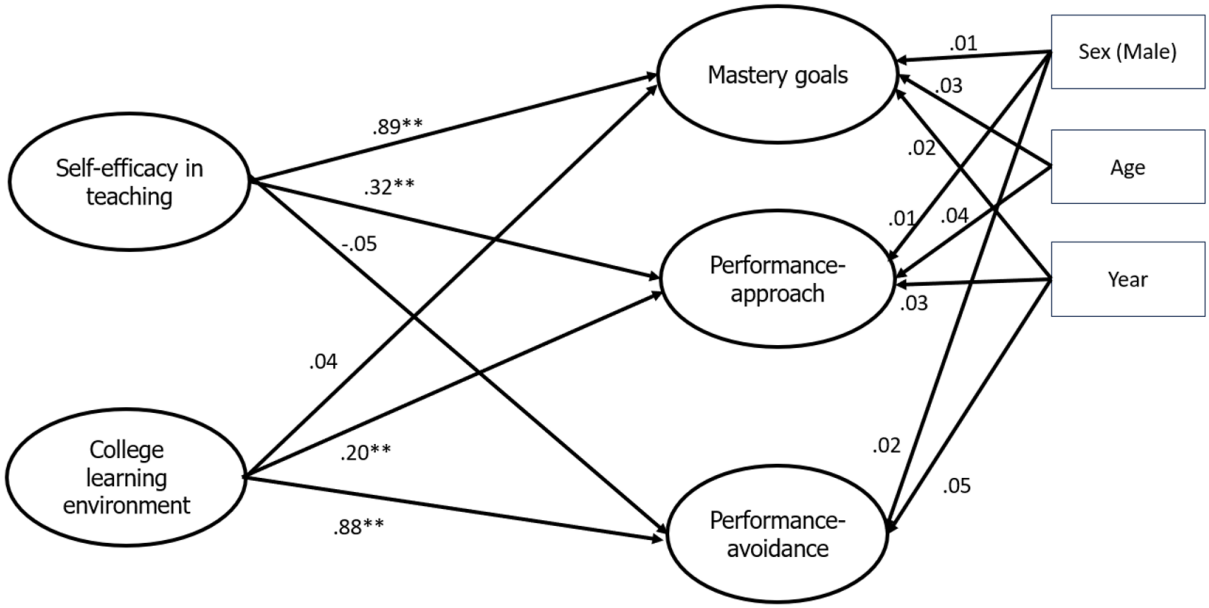


Figure 1: Relationship between self-efficacy in teaching and college learning environment with mastery goals, performance approach and performance avoidance

DISCUSSION

The research findings refer to a notable, direct, and positive effect of teaching self-efficacy in teaching on mastery goals and performance approach goals. This result highlights the self-efficacy in teaching in determining the mastery goals of pre-service teachers are particularly significant. These findings are in line with other studies on the same domain (Bong, 2001; Cho & Shim, 2013; Gerhardt & Brown, 2006; Yildizlli, 2019). It is obvious, based on the reached findings, that self-efficacy in teaching in fact makes it possible for pre-service teachers to approach a risky pattern to enhance their teaching capabilities, skills, methods, and techniques, and achieve more profound educational targets in the process of teaching, rather than satiating with formal instruction that sets apart achievement goals and the desire to exude one's abilities in teaching.

In addition to the prior, this study's final model indicated that the learning environment in teacher-training colleges plays a critical role in forming and reforming the performance goals of pre-service teachers. These findings are compatible with other studies that have indicated an obvious connection between learning environment, self-efficacy and achievement goals (Alkharusi, 2009; Elliot & Church, 1997; Nie & Lau, 2010; Urdan & Midgley, 2003; Yıldızlı et al., 2016).

Mastery goals for teaching are more challenging to achieve using the *top–down* teaching methods, which is characterized by classical -old–school- instruction, controlled mainly by the lecturers, without permitting the pre-service teachers' genuine practice and the opportunity to develop necessary teaching skills and strategies. Even though some teaching intervention by lecturers is indeed vital, the results of this study suggest a bottom–up teaching approach, characterized by a *constructivist* learning atmosphere that involves more engagement of the pre-service teachers in the training process at large. The prior approach is expected to enable the participants to independently construct their knowledge and actively create teaching strategies and techniques that would better advance their adoption of mastery skills rather than performance goals.

A constructivist learning atmosphere is mainly concerned with learning, dialogue in instruction, significant learning, and employment of alternative evaluation tools in training teachers. Such environment can provide pre-service teachers with numerous opportunities for actual and genuine teaching skills in the early stages of their studies. To enhance the learning

environment by changing its distinctive feature into a more constructivist one, the academic institution should pay close attention to six components:

1. Improving the instructional quality of lecturers: it is necessary in this respect to set an example of *the effective teacher figure*, who comprehends the students' challenges, illustrates the educational material in a proper and effective manner, and possesses the necessary skills to break the barriers and the rigidity of the curriculum that create a sense on boredom and monotony.
2. Providing more opportunities for pre-service teachers to develop general skills such as: conducting proper work plans, creating problem solving methods, possessing the capacity to deal with emerging problems, acquiring communication skills, and properly engaging in teamwork. For example, Israel's Ministry of Education is currently integrating pre-service teacher into schools with the aim of providing such opportunities.
3. Creating a more intellectually provoking academic environment, to raise the desire to learn.
4. Improving the current evaluation methods and promoting diverse evaluation tools that would stress high-level thinking, as well as constructive feedback for the pre-service teachers on the tasks they are assigned.
5. Investing more tangible effort into the structure of classroom assignments and activities, so that they are clear, engrossing, and in direct relevance to the pre-service teachers.
6. Promoting a variety of instruction methods, skills, and techniques, and shifting from classical, lecturer-centered instruction to instruction focused on the pre-service teachers themselves.

It is necessary to note that performance goals in particular are not undesirable in pre-service teachers, since they will need to assist their students achieve the standards designed and agreed upon by the education system. Certain studies have exuded that participants may possess multi-faceted goals concurrently (Levy et al., 2004; Yildizli, 2020). However, our study's structural model shows that the main emphasis should be attributed to mastery goals and development of skills which may, in turn, increase performance goals, demonstration of abilities, and the meeting of required standards of the college learning environment.

The findings of the current study suggest that there is pressing need to enhance the various components of self-efficacy of pre-service teachers, as effectively raising their self-efficacy at a premature phase of their training process is more likely to assist them cope better in their first year of teaching, a time at which they are most in peril and predisposed to be adversely influenced by substandard and undesirable experiences (Hoy & Spero, 2005).

Limitations

Together with the promising results from this study, there are some limitations that future research may address. First, although the model fit the data, alternative models may fit the data as well and should be tested in future research. Second, replication studies are necessary to confirm the results of the current study and to add to their generalizability. Third, a qualitative approach would likely aid in the interpretation of the significant effects revealed in the current study. Finally, the effect of the research variables on other outcome variables such as achievement in theoretical courses and in practicum, and emotional and social variables that characterize pre-service teachers throughout the training period, could also evaluate the efficacy of the training process in teacher-training colleges.

CONCLUSIONS

Learning environments have a significant effect on the adoption of achievement goals, both mastery goals and performance approach goals. our study shows a need for improving the learning environment at our country's teacher-training colleges. A learning environment more consistent with constructivist ideas would enable greater teaching self-efficacy and encourage pre-service teachers to set not only performance goals but also mastery goals for themselves.

ACKNOWLEDGEMENT

No funding was received for this study.

REFERENCES

Aarts, H. and Elliot, A. (2012) '*Goal-directed behavior*'. Psychology Press. https://doi.org/10.4324/9780203869666

Aldridge, J. and Fraser, B. (2000) 'A cross-cultural study of classroom learning environments in Australia and Taiwan', *Learning Environments Research*, Vol. 3, No. 2, pp. 101–134. https://doi.org/10.1023/a:1026599727439

Alkharusi, H. (2009) 'Classroom assessment environment, self-efficacy, and mastery goal orientation: A causal model', *INTI Journal: Special Issue on Teaching and Learning*, pp. 104–116.

Ames, C. (1992) 'Classrooms: Goals, structures, and student motivation', *Journal of Educational Psychology,* Vol. 84, No. 3, pp. 261–271. https://doi.org/10.1037/0022-0663.84.3.261

Ames, C. and Archer, J. (1988) 'Achievement goals in the classroom: Students' learning strategies and motivation processes', *Journal of Educational Psychology*, Vol. 80, No. 3, pp. 260–267. https://doi.org/10.1037/0022-0663.80.3.260

Arbuckle, J. L. (2013) IBM SPSS AMOS 22 Users' Guide, New York: IBM Corp.

Bandura, A. (1997) *Self-efficacy: The exercise of control*, New York: W.H. Freeman and Company.

Bell, B. S. and Kozlowski, S. W. J. (2002) 'Goal orientation and ability: Interactive effects on self-efficacy, performance, and knowledge', *The Journal of Applied Psychology*, Vol. 87, No. 3, pp. 497–505. https://doi.org/10.1037/0021-9010.87.3.497

Bipp, T., Steinmayr, R. and Spinath, B. (2012) 'A functional look at goal orientations: Their role for self-estimates of intelligence and performance', *Learning and Individual Differences*, Vol. 22, No. 3, pp. 280–289. https://doi.org/10.1016/j.lindif.2012.01.009

Bong, M. (2001) 'Between-and within-domain relations of academic motivation among middle and high school students: Self-efficacy, task-value and achievement goals', *Journal of Educational Psychology*, Vol. 93, No. 1, pp. 23–34. https://doi.org/10.1037/0022-0663.93.1.23

Butler, R. (2007) 'Teachers' achievement goal orientations and associations with teachers' help seeking: Examination of a novel approach to teacher motivation', *Journal of Educational Psychology*, Vol. 99, No. 2, pp. 241–252. https://doi.org/10.1037/0022-0663.99.2.241

Butler, R. and Shibaz, L. (2008) 'Achievement goals for teaching as predictors of students' perceptions of instructional practices and students' help seeking and cheating', *Learning and Instruction*, Vol. 18, No. 5, pp. 453–467. https://doi.org/10.1016/j.learninstruc.2008.06.004

Chen, J. A. and Usher, E. L. (2013) 'Profiles of the sources of science self-efficacy', *Learning and Individual Differences*, Vol. 24, pp. 11–21. https://doi.org/10.1016/j.lindif.2012.11.002

Cho, Y. and Shim, S. S. (2013) 'Predicting teachers' achievement goals for teaching: The role of perceived school goal structure and teachers' sense of efficacy', *Teaching and Teacher Education*, Vol. 32, pp. 12–21. https://doi.org/10.1016/j.tate.2012.12.003

Daniels, L. M., Frenzel, A. C., Stupnisky, R. H., Stewart, T. L. and Perry, R. P. (2013) 'Personal goals as predictors of intended classroom goals: Comparing elementary and secondary school pre-service teachers', *British Journal of Educational Psychology*, Vol. 83, No. 3, pp. 396–413. https://doi.org/10.1111/j.2044-8279.2012.02069.x

Deemer, S. (2004) 'Classroom goal orientation in high school classrooms: revealing links between teacher beliefs and classroom environments', *Educational Research*, Vol. 46, No. 1, pp. 73–90. https://doi.org/10.1080/0013188042000178836

Dellinger, A. M. B. (2001) '*A study of the measurement and sources of teachers' self and collective efficacy beliefs in professional learning environments*', PhD dissertation, [Online], Available: http://digitalcommons.lsu.edu/cgi/viewcontent.cgi?article=1335&context=gradschool_disstheses

Elliot, A. J. and Church, M. A. (1997) 'A hierarchical model of approach and avoidance achievement motivation', *Journal of Personality and Social Psychology*, Vol. 72, No. 1, pp. 218–232. https://doi.org/10.1037/0022-3514.72.1.218

Elliot, A. J. and Fryer, J. W. (2008) 'The goal construct in psychology', in Shah J. Y and Gardner, W. L. (ed.), *Handbook of motivation science*, New York: The Guilford Press.

Fasching, M. S., Dresel, M., Dickhäuser, O. and Nitsche, S. (2010) 'Goal orientations of teacher trainees: Longitudinal analysis of magnitude, change and relevance', *Journal for Educational Research Online*, Vol. 2, No. 2, pp. 9–33. https://doi.org/10.25656/01:4573

Gerhardt, M. W. and Brown, K. G. (2006) 'Individual differences in self-efficacy development: The effects of goal orientation and affectivity', *Learning and Individual Differences*, Vol. 16, No. 1, pp. 43–59. https://doi.org/10.1016/j.lindif.2005.06.006

Gollwitzer, P. M. and Oettingen, G. (2012) 'Goal pursuit', in Ryan, R. M. (ed.) *The Oxford handbook of human motivation*, Oxford: Oxford University Press.

Gonida, E. N., Voulala, K. and Kiosseoglou, G. (2009) 'Students' achievement goal orientations and their behavioral and emotional engagement: Co-examining the role of perceived school goal structures and parent goals during adolescence', *Learning and Individual Differences*, Vol. 19, No. 1, pp. 53–60. https://doi.org/10.1016/j.lindif.2008.04.002

Hagenauer, G. and Hascher, T. (2010) 'Learning enjoyment in early adolescence', *Educational Research and Evaluation,* Vol. 16, No. 6, pp. 495–516. https://doi.org/10.1080/13803611.2010.550499

Han, J., Yin, H. and Wang, W. (2016) 'The effect of tertiary teachers' goal orientations for teaching on their commitment: The mediating role of teacher engagement', *Educational Psychology*, Vol. 36, No. 3, pp. 526–547. https://doi.org/10.1080/01443410.2015.1044943

Hoy, A. W. and Spero, R. B. (2005) 'Changes in teacher efficacy during the early years of teaching: A comparison of four measures', *Teaching and Teacher Education*, Vol. 21, No. 4, pp. 343–356. https://doi.org/10.1016/j.tate.2005.01.007

Kaplan, A. and Maehr, M. L. (2007) 'The contributions and prospects of goal orientation theory', *Educational Psychology Review*, Vol. 19, No. 2, pp. 141–184. https://doi.org/10.1007/s10648-006-9012-5

Kass, E. and Miller, E. (2015) 'Teacher-education programs and teacher trainees' sense of professional efficacy', *Journal of Pedagogy, Pluralism, and Practice*, Vol. 7, No. 1, pp. 128.

Klassen, R. M. and Chiu, M. M. (2011) 'The occupational commitment and intention to quit of practicing and pre-service teachers: Influence of self-efficacy, job stress, and teaching context', *Contemporary Educational Psychology*, Vol. 36, No. 2, pp. 114–129. https://doi.org/10.1016/j.cedpsych.2011.01.002

Klassen, R. M. and Tze, V. M. (2014) 'Teachers' self-efficacy, personality, and teaching effectiveness: A meta-analysis', *Educational Research Review*, Vol. 12, pp. 59–76. https://doi.org/10.1016/j.edurev.2014.06.001

Kucsera, J. V., Roberts, R., Walls, S., Walker, J. and Svinicki, M. (2011). 'Goal orientation towards teaching (GOTT) scale', *Teachers and Teaching*, Vol. 17, No. 5, pp. 597–610. https://doi.org/10.1080/13540602.2011.602212

Laine, E. and Gegenfurtner, A. (2013) 'Stability or change? Effects of training length and time lag on achievement goal orientations and transfer of training', *International Journal of Educational Research*, Vol. 61, pp. 71–79. https://doi.org/10.1016/j.ijer.2013.03.014

Levy, I., Kaplan, A. and Patrick, H. (2004) 'Early adolescents' achievement goals, social status, and attitudes towards cooperation with peers', *Social Psychology of Education*, Vol. 7, No. 2, pp. 127–159. https://doi.org/10.1023/b:spoe.0000018547.08294.b6

Liu, J., Xiang, P., McBride, R. and Chen, H. (2019) 'Psychometric properties of the Cognitive and Metacognitive Learning Strategies Scales among preservice physical education teachers: A bifactor analysis', *European Physical Education Review*, Vol. 25, No. 3, pp. 616–639. https://doi.org/10.1177/1356336x18755087

Meece, J. L., Anderman, E. M. and Anderman, L. H. (2006) 'Classroom goal structure, student motivation, and academic achievement', *Annual Review of Psychology*, Vol. 57, No. 1, pp. 487–503. https://doi.org/10.1146/annurev.psych.56.091103.070258

Minnaert, A., Boekaerts, M., De Brabander, C. and Opdenakker, M. C. (2011) 'Students' experiences of autonomy, competence, social relatedness and interest within a CSCL environment in vocational education: The case of commerce and business administration', *Vocations and Learning*, Vol. 4, No. 3, pp. 175–190. https://doi.org/10.1007/s12186-011-9056-7

Nicholls, J. G., Patashnick, M. and Nolen, S. B. (1985) 'Adolescents' theories of education', *Journal of Educational Psychology*, Vol. 77, No. 6, pp. 683–692. https://doi.org/10.1037/0022-0663.77.6.683

Nie, Y. and Lau, S. (2010) 'Differential relations of constructivist and didactic instruction to students' cognition, motivation, and achievement', *Learning and Instruction*, Vol. 20, No. 5, pp. 411–423. https://doi.org/10.1016/j.learninstruc.2009.04.002

Nitsche, S., Dickhäuser, O., Fasching, M. S. and Dresel, M. (2011) 'Rethinking teachers' goal orientations: Conceptual and methodological enhancements', *Learning and Instruction*, Vol. 21, No. 4, pp. 574–586. https://doi.org/10.1016/j.learninstruc.2010.12.001

Papaioannou, A. and Christodoulidis, T. (2007) 'A measure of teachers' achievement goals', *Educational Psychology*, Vol. 27, No. 3, pp. 349–361. https://doi.org/10.1080/01443410601104148

Patrick, H., Ryan, A. M. and Kaplan, A. (2007) 'Early adolescents' perceptions of the classroom social environment, motivational beliefs, and engagement', *Journal of Educational Psychology*, Vol. 99, No. 1, pp. 83–98. https://doi.org/10.1037/0022-0663.99.1.83

Pfitzner-Eden, F. (2016) 'Why do I feel more confident? Bandura's sources predict preservice teachers' latent changes in teacher self-efficacy', *Frontiers in Psychology*, Vol. 7, p. 1486. https://doi.org/10.3389/fpsyg.2016.01486

Polychroni, F., Hatzichristou, C. and Sideridis, G. (2012) 'The role of goal orientations and goal structures in explaining classroom social and affective characteristics', *Learning and Individual Differences,* Vol. 22, No. 2, pp. 207–2017. https://doi.org/10.1016/j.lindif.2011.10.005

Retelsdorf, J., Butler, R., Streblow, L. and Schiefele, U. (2010) 'Teachers' goal orientations for teaching: Associations with instructional practices, interest in teaching, and burnout', *Learning and Instruction*, Vol. 20, No. 1, pp. 30–46. https://doi.org/10.1016/j.learninstruc.2009.01.001

Romi, S. and Leyser, Y. (2006) 'Exploring inclusion preservice training needs: a study of variables associated with attitudes and self-efficacy beliefs', *European Journal of Special Needs Education*, Vol. 21, No. 1, pp. 85–105. https://doi.org/10.1080/08856250500491880

Segers, M. and Gegenfurtner, A. (2013) 'Transfer of training: New conceptualizations through integrated research perspectives', *Educational Research Review*, Vol. 8, pp. 1–4. https://doi.org/10.1016/j.edurev.2012.11.007

Urdan, T. and Midgley, C. (2003) 'Changes in the perceived classroom goal structure and pattern of adaptive learning during early adolescence', *Contemporary Educational Psychology*, Vol. 28, No. 4, pp. 524–551. https://doi.org/10.1016/s0361-476x(02)00060-7

Usher, E. L. and Pajares, F. (2008) 'Sources of self-efficacy in school: Critical review of the literature and future directions', *Review of Educational Research*, Vol. 78, No. 4, pp. 751–796. https://doi.org/10.3102/0034654308321456

Van Daal, T., Donche, V. and De Maeyer, S. (2014) 'The impact of personality, goal orientation and self-efficacy on participation of high school teachers in learning activities in the workplace', *Vocations and Learning*, Vol. 7, No. 1, pp. 21–40. https://doi.org/10.1007/s12186-013-9105-5

Volet, S. (2013) 'Extending, broadening, and rethinking existing research on transfer of training', *Educational Research Review,* Vol. 8, pp. 90–95. https://doi.org/10.1016/j.edurev.2012.11.005

Wolters, C. A. and Daugherty, S. G. (2007) 'Goal structures and teachers' sense of efficacy: Their relation and association to teaching experience and academic level', *Journal of Educational Psychology*, Vol. 99, No. 1, pp. 181–193. https://doi.org/10.1037/0022-0663.99.1.181

Yildizli, H. (2019) 'Structural relationships among teachers' goal orientations for teaching, self-efficacy, burnout, and attitudes towards teaching', *Journal on Efficiency and Responsibility in Education and Science*, Vol. 12, No. 4, pp. 111–125. https://doi.org/10.7160/eriesj.2019.120402

Yildizli, H. (2020) 'Classroom assessment practices and student goal orientations in mathematics classes', *Journal of Qualitative Research in Education*, Vol. 8, No. 1, pp. 294–323. https://doi.org/10.14689/issn.2148-2624.1.8c.1s.13m

Yildizli, H., Saban, A. and Baştuğ, M. (2016) 'Öğretmeye yönelik hedef yönelimi ölçeği'nin Türkçe'ye uyarlanması [Adaptation of goal orientations for teaching Scale into Turkish] ', *Elementary Education Online*, Vol. 15, No. 4, pp. 1254–1267. https://doi.org/10.17051/io.2016.62664

Zimmerman, B. J. and Martinez-Pons, M. (1992) 'Perceptions of efficacy and strategy use in the self-regulation of learning', in Schunk, D. H. and Meece, J. L. (ed.) *Student perceptions in the classroom*, Mawah, New Jersey: L. Erlbaum.

Zimmerman, B. J. (1990) 'Self-regulated learning and academic achievement: An overview', *Educational Psychologist*, Vol. 25, No. 1, pp. 3–17. https://doi.org/10.1207/s15326985ep2501_2

Full research paper

GENDER DIFFERENCES IN SCHOOL ACHIEVEMENT AND ATTITUDES TOWARDS MOTIVATION IN SECONDARY ECONOMIC EDUCATION

Jaromír Novák[1]✉
Kateřina Berková[2]
Andrea Kubišová[3]
Dana Kolářová[4]

[1]Competence Centre, Faculty of Economics and Finance, University of Economics in Bratislava, Slovakia

[2]Department of Economic Teaching Methodology, Faculty of Finance and Accounting, Prague University of Economics and Business, Czech Republic

[3]Department of Mathematics, College of Polytechnics Jihlava, Czech Republic

[4]Department of Hospitality, University College Prague, Czech Republic

✉ jaromir.novak@euba.sk

ABSTRACT

This study examined the school achievement and attitudes of Czech and Slovak secondary school students (n = 572; age: 17-19) towards motivation in economic subjects. The aim was to analyse the factors of students' motivation, their relations with selected teacher's competences, and students' school achievement with regard to their gender. The data were obtained by using a questionnaire and analyzed by Mann-Whitney U-test, Pearson, and Spearman correlation coefficients. Girls performed better than boys in all the subjects analysed. At the same time, for all factors of motivation with significant gender differences, their motivation was stronger compared to boys. School achievement and attitudes towards selected motivation factors are correlated. Some factors motivating boys with better achievement motivate girls with worse achievement.

KEYWORDS

Gender differences, school achievement, motivation, teacher personality, economic education, secondary schools

HOW TO CITE

Novák J., Berková K., Kubišová A., Kolářová D. (2024) 'Gender Differences in School Achievement and Attitudes towards Motivation in Secondary Economic Education', *Journal on Efficiency and Responsibility in Education and Science*, vol. 17, no. 3, pp. 223–236. http://dx.doi.org/10.7160/eriesj.2024.170305

Article history
Received
January 29, 2023
Received in revised form
February 3, 2024
Accepted
August 27, 2024
Available on-line
September 30, 2024

Highlights

- *School achievement and attitudes towards selected motivation factors are correlated.*
- *Gender differences were found in attitudes towards motivation in economic subjects.*
- *Girls perform better than boys in all of the analysed school subjects.*
- *Girls spend more time per week studying both economic subjects compared to boys.*

INTRODUCTION

Students' school achievement depends primarily on the performance of the students. The theory assumes that this is determined mainly by their ability, motivation, and learning climate, including the teaching methods chosen and their fit with students' learning styles. Apart from the relevant abilities of students, their self-regulation, self-efficacy, and fear of examinations can be considered as the best predictors of school achievement (Pintrich and de Groot, 1990). However, the cited study does not consider motivational beliefs that the curriculum is useful and interesting as a sufficient guarantee of school achievement. On the contrary, the OECD (2010) infers from the PISA results that students may be motivated to study because of the relevance of the subject to their future education or career. The study mentions that students' attitudes towards the subject should be related to school success, but it is not clear if educational success brings about better attitudes or if a positive attitude towards subjects is a key factor motivating students to learn.

The big challenge for teachers is how to make the best use of all these attributes. Successful teachers see knowledge and skills as something built up over time, so they move away from traditional, direct teaching to methods that stimulate active learning. In this context, motivation occupies a significant role in students' learning. Many studies have examined factors with an impact on motivation and school achievement (e.g., Coetzee et al., 2020; Demirbas and Demirkan, 2007; Giota and Bergh, 2021; Marmeleira et al., 2020; Owoseni et al., 2020; Silva-Arias et al., 2020; Zhang and Wang, 2020). Statistically significant differences in motivation and school achievement in terms of

gender, school type, subject, age, year of study, learning style, and socio-economic background have been found, but they have not always been confirmed in a truly consistent manner, and in some studies, they have not been demonstrated at all. Moreover, the studies usually analyze school achievement in traditional subjects such as mathematics, languages, or science. Therefore, it is interesting to examine whether the results would be similar in the field of secondary economic education, which is an example of vocational education where students' motivation to study might be different from that in the subjects forming their general education.

To effectively manage motivation in the school environment and promote active learning and thus school achievement, it is necessary to know students well and understand the differences that may arise between different students and their groups, for example, due to gender differences. This consideration is supported by an OECD study (2009:8), according to which "*educational policy has to take into account the existence of gender differences in performance to be effective in promoting quality student outcomes and equity*."

LITERATURE REVIEW

Gender differences in students' school performance

Gender differences are reflected in a variety of ways in the field of education. The findings of major international studies and comparisons (e.g., PISA) show that there are large gender differences in students' performance and attitudes in different countries and that they also vary according to the areas tested. Although the OECD itself sees the main reason for studying gender differences as being able to understand the origins of gender inequalities better, it also sees the possibility of improving average student performance through such research and better understanding how students learn as an important reason (OECD, 2009).

In relation to school performance, the majority of studies tend to confirm some advantages of girls over boys, although these differences may vary according to educational areas, etc. Gender differences in students' performance in language, mathematics, and science are the most frequently examined. Particularly valuable for making general statements are the results of meta-analyses summarizing the findings of a large number of studies (e.g., a generalizing study on school performance by Voyer and Voyer, 2014; or a meta-analysis focusing on school performance in mathematics by Lindberg et al., 2010). The former meta-analysis is based on a large sample of heterogeneous studies examining gender differences in school performance. It found a statistically significant difference in school performance in terms of grades in favor of girls in all educational domains examined, with the largest differences in favor of girls shown in language and the smallest in mathematics. The fact that mathematics is the least distinct and unambiguous area in terms of gender differences is confirmed by studies that found no statistically significant differences in mathematics performance by gender at all (e.g., Agyei and Eyiah-Bediako, 2008; Lindberg et al., 2010).

The study by Eriksson et al. (2020) confirms the existence of gender differences in learning outcomes in subjects based on abstract-visual thinking. Coetzee et al. (2020) also confirmed that girls perform better than boys (especially in language, only very slightly better in mathematics). A similar conclusion was reached in a longitudinal study of about 70,000 English children, which found that girls performed statistically significantly better in all subjects studied except physics (Deary et al., 2007).

A number of other studies have examined gender differences in school achievement in selected samples of learners, with mixed findings. The study by Marmeleira et al. (2020), while confirming better grades for girls than boys in several subjects (Portuguese, Philosophy, Mathematics, Foreign Language, Biology/Geology, and Psychology), did not find statistically significant gender differences in other subjects (Physics/Chemistry, Geography and History). A specific educational area is physical education, in which this study and several other studies have confirmed gender differences in performance in the opposite direction, i.e., boys perform better in this subject. Demirbas and Demirkan (2007:345) confirmed that *"scores of females were higher in artistic and fundamental courses and in the semester academic performance scores"* but, conversely, *"performance scores of males were higher in technology-based courses"*, supporting the idea that gender differences may differ depending on the subject of study. Therefore, we decided to examine gender differences in secondary economic education, represented in our study by two major school subjects: economics and accounting.

Key factors influencing students' school achievement from a gender perspective

School achievement is determined by a number of factors. Many studies emphasize that teachers and their professional characteristics are important factors influencing students' academic success (Chetty et al., 2014; Rivkin et al., 2005; Rockoff, 2004; Fung et al., 2017). Other factors include students' abilities and motivation, as well as fear of exams and stage fright (e.g., Pintrich and de Groot, 1990), which are usually more evident in girls than boys (Chmelárová et al., 2018). This conclusion is also supported by a study whose results are significant because it simultaneously examined gender differences in school achievement in the same sample, failing to confirm lower school achievement for girls compared to boys (Núñez-Peña et al., 2016). Therefore, it can be assumed that other factors influence school achievement in a more significant way. At the same time, increased fear of exams might be associated with a more responsible approach to studying and higher levels of self-discipline. This reasoning is supported by the results of an analysis of the goal orientation of secondary school students (Erdem-Keklik and Keklik, 2014), which revealed significant gender differences. Within the chosen structure of factors of goal orientation, girls' goal orientation was particularly confirmed to be more oriented towards not failing, which, although increasing their fear of exams, also contributed to their better school performance through a more responsible approach to studying. According to Giota and Bergh (2021), girls have a stronger Social responsibility, as well as Mastery and a greater orientation towards Future than boys.

Student motivation is considered one of the crucial factors determining the effectiveness of the educational process (Fontana, 2014). Teaching is more effective if teachers "*reinforce students' self-concept and confidence in their own abilities and show students that it is worth investing time and effort in learning because it is an investment with ample returns*" (Habók et al., 2020:10). Eriksson (2020) studied students' motivation to learn mathematics in 50 countries and found that gender differences in interest in the subject are related to differences in student achievement. The factors that influence the intensity of motivation are numerous. In general, the two most fundamental characteristics of a teacher's influence on students have been shown to have a positive effect: emotional intelligence and teacher's self-efficacy (Valente et al., 2020). The content of the issues studied, the students' strictness about themselves, and their expectations from their studies cannot be overlooked (Berková et al., 2018). An individual's expectations and their fulfilment can be considered important for the intensity of motivation. Research by Perelygina et al. (2020) also shows that increasing students' motivation depends on the motivation of the teacher to perform their profession. This is reflected in their teaching style, choice of teaching methods, and emotional intelligence in relation to the profession (Krpálek and Krpálková Krelová, 2016).

Higher student motivation usually leads to greater learning activity, which increases learning effectiveness. Several studies have shown that girls are usually more engaged in the learning process than boys. If teachers want to encourage students to be active, they should adjust teaching and learning methods and tools to their gender. For girls, support and assistance aimed at a better understanding of the curriculum covered have a better effect on increasing engagement; for boys, clearer structuring of learning activities has a greater impact (Bru et al., 2021). According to the study by Lebid and Shevchenko (2020), there are also gender differences in the perception of interactive learning methods - they were perceived more favorably by girls than boys in the sample studied, both in terms of their effectiveness and appropriateness. This also points to the importance of tailoring the choice of appropriate teaching methods to the gender of the students.

This study analyses gender differences in school achievement and students' attitudes towards motivation in secondary economic education, where these relations are relatively under-researched. A research study (Berková et al., 2018) examined interdisciplinary differences in attitudes towards motivation, but it focused on students in higher economic education. Students in Finance and Management, and Applied Computer Science consider as the most motivating factors when the teacher leads them to think, accepts unusual ideas and discusses them, gives more than facts, and uses real-life examples. The research found that students of technical majors had greater demands for studying than students of economic majors, who were more indifferent to studying.

An OECD study (2014) confirmed a significant positive relation between numeracy and financial literacy. Economic education is based on many mathematical principles and logic theory and is characterized by the following cognitive procedures - analysis, synthesis, deduction, and induction. The influential English economist J. M. Keynes stated that economics is a tool of thinking to make correct conclusions relevant in relation to the economic problem being solved (Jurečka et al., 2013), which emphasizes the connection between economics, thinking, and drawing correct conclusions. One of the main subjects of secondary economic education is accounting, which is not only based on economic theories and methods but also uses mathematical operations and the theory of logic. Thus, accounting functions in interactions are underpinned by alternative mathematical-analytical models (Watts and Zimmerman, 1978; Fields et al., 2001). Since economic courses demand critical thinking skills, reading comprehension should also be encouraged as they are interrelated skills (Wiliam, 2011). Stoet and Geary (2013) have found that in the area of language learning, the average girl always outperforms the average boy in reading. From the above, it can be deduced that language and mathematical skills are important prerequisites for studying economic subjects. In this study, the researchers also examined these questions in the field of secondary economic education, where they tested the validity of these relationships.

The purpose and objectives of the study

The research on the relation between school achievement and students' motivation focuses on gender differences, as several studies suggest (Coetzee et al., 2020; Demirbas and Demirkan, 2007; Giota and Bergh, 2021; Marmeleira et al., 2020; Owoseni et al., 2020; Silva-Arias et al., 2020; Zhang and Wang, 2020) that neither academic achievement nor motivation have the same tendency in education in terms of gender. The main purpose of this study is to examine gender differences in school achievement and the attitudes of students from Czech and Slovak secondary schools towards motivation in the profile secondary school subjects of economics and accounting.

Motivation is examined by students' self-assessment of how they believe a particular factor in the course might motivate them. Nine motivation factors are included in the model, which was inspired by the studies of Berková et al., 2018 and Perelygina et al., 2020. To better portray the attitudes towards motivation, students' attitudes towards the personality of the teacher of economics and accounting were further studied and split into the following professional competencies: (a) *expertise*; (b) *clear explanation*; (c) *responsiveness and empathy*; (d) *humanity, student's trust towards the teacher*; (e) *ability to motivate* (Akin and Kurbanoglu, 2011). To make it more contextual, the model also takes into account gender differences in time spent studying at home. The learning outcomes in economics and accounting are compared with those in mathematics and mother tongue in order to demonstrate the association of gender differences in economic subjects with differences in school achievement in the area of reading and mathematics (Stoet and Geary, 2013) and also to confirm the assumed relation between numeracy and financial literacy (OECD, 2014). Thus, the paper adds to the literature with new findings in relation to the teaching of economic subjects in secondary schools.

The main objectives of this study are:

- to analyse and compare the main findings of the existing studies about gender differences among students in their school achievement with respect to the key aspects of learning, mainly focusing on the relation between students' school achievement and their motivation to learn,
- to expand knowledge in the area of gender differences between high school students already confirmed for traditional high school subjects such as mathematics, sciences, languages, and the like by researching analogous relations within the scope of economic school subjects,
- to analyse whether and how these aspects are influenced by teachers of economic subjects and their personality.

In view of the previous studies that led to different results, the current study focuses on these two research questions: Do secondary school students differ in the key aspects of learning (motivation, preparation time, school achievement) from a gender perspective? Is there a link between these aspects in the level of the two profile secondary school subjects of economics and accounting?

The following hypotheses arise from the research questions and the main objectives of the study:

- **H1:** Students' attitudes towards motivation in the case of economics and accounting differ by gender.
- **H2:** Students' attitudes towards the teacher in the case of the economics and accounting courses differ by gender.
- **H3:** There are gender differences in the time taken by students to prepare for the subjects of economics and accounting.
- **H4:** There are gender differences in students' school achievement in the subjects of economics and accounting.
- **H5:** From a gender perspective, there is a correlation between preparation time and school success in the case of economics and accounting.
- **H6:** School success in the subjects of economics, accounting, mathematics, and mother tongue is correlated with gender.
- **H7:** There is a correlation between students' attitudes towards motivation and school achievement in the case of economics and accounting in terms of gender.

MATERIALS AND METHODS

Participants

The research involved secondary schools from the Czech Republic and Slovakia that provide quality economic education. The sample of secondary schools was chosen deliberately, the main criterion being close cooperation with both research institutes. The cooperation involves providing teaching training for future teachers of economic subjects. The selected secondary schools implement analogous study programs with a similar curriculum and, as faculty training schools, guarantee quality teaching of economic subjects. Teachers' approach to teaching is innovative, and they use not only traditional but mainly modern teaching methods, making their students' learning more experience-based.

A total of 7 secondary schools from the capital cities of both countries participated in the research. 577 students in their final (fourth) and pre-final (third) years were involved in the main research from these secondary schools. 572 questionnaires out of 577 were valid. 5 respondents were excluded from the total observation due to incomplete questions or ambiguous answers. The sample of students was drawn using purposive sampling. This method was adopted to obtain responses from upper-year students who have more experience with economic education and can assess their attitudes better. The research sample consisted of 392 girls (i.e., 68.5%) and 180 boys (i.e., 31.5%), which corresponds to the normal ratio of boys and girls at this type of school in both countries. The respondents studying the third year (total: 288) and the fourth year (total: 284) of secondary school were all 17-19 years old.

Data and instruments

The research was conducted as a quantitative study using the questionnaire method. The questionnaire was constructed as a non-standardized questionnaire, and the selection of variables was inspired by several national and international studies. The questionnaire was distributed to secondary schools and among students in hard copy from April to June 2018. Data was collected by the researchers personally, thus ensuring a higher return guarantee. Respondents were advised to answer as objectively as possible, with the understanding that the data collected were subject to a high degree of confidentiality and anonymity and would only be used for research purposes. The questionnaire consisted of several separate sections.

In the first section, students expressed their attitudes towards motivation in the subjects of economics and accounting. Motivation was expressed by 9 factors (Berková et al., 2018; Perelygina et al., 2020). The students expressed their attitudes towards each factor, separately for the subject of economics and the subject of accounting:

- Factor 1: the exposition emphasizes the curriculum's applicability to real-life situations.
- Factor 2: the exposition is mainly focused on thorough practicing of the basic curriculum with simple examples.
- Factor 3: the teacher just explains the curriculum from the textbook.
- Factor 4: the teacher engages the students with frequent questions, allowing for discussion.
- Factor 5: the teacher uses examples from real life and case studies, and the problems raised are jointly solved and discussed.
- Factor 6: the teacher sets the task and lets the students work independently and individually.
- Factor 7: the teacher guides students to think about the curriculum, teaching them to critically evaluate information.
- Factor 8: the teacher accepts unusual ideas, evaluates their applicability, and shows how to learn from possible mistakes.
- Factor 9: the teacher is flexible and answers all questions, but does not deviate from the topic, communicating "something more" than just the facts.

A Likert scale was used to measure attitudes towards motivation in the subjects of economics and accounting, on which students expressed how the factors listed would motivate them. These were ordinal values: -2 strongly demotivating; -1 moderately demotivating; 0 don't care; 1 moderately motivating; 2 strongly motivating.

In the second section, students expressed their attitudes towards the personality of the teacher of economics and accounting. The teacher's personality was described by the following competences (Alharbi et al., 2020; Berková and Krejčová, 2016; Perelygina et al., 2020):

- Competence 1 – expertise
- Competence 2 – clear explanation
- Competence 3 – responsiveness, empathy
- Competence 4 – humanity, the student's trust in the teacher
- Competence 5 – ability to motivate students

Again, a Likert scale was used to measure student attitudes towards the teacher, on which students expressed their evaluation of the teacher from their perspective, i.e., the validity of a statement describing each characteristic accurately. These were ordinal values: from -2 to 2, with -2 expressing a low rating, 0 I don't care; and 2 a high rating.

In the third section of the questionnaire, students reported their weekly preparation time for the subjects of economics and accounting (each separately). Weekly preparation time was described using the following scale:

- More than 4 hours per week (5)
- 3 - 4 hours per week (4)
- 1 - 2 hours per week (3)
- Less than 60 minutes per week (2)
- I don't prepare for this subject at home at all (1)

The time variation was expressed in a scale of 1 - 5, i.e., again with ordinal values.

The fourth section related to the school achievement of students primarily in the subjects of economics and accounting, by means of the annual grade obtained by students at the end of their second year of study, with a range of 1 - excellent; 2 - very good; 3 - good; 4 - sufficient; 5 - insufficient. For better comparison with the results of other studies, the researchers also examined grades in the subjects of mathematics and mother tongue (Czech or Slovak). Accounting is based on mathematical operations and is based on the theory of logic (Fields et al., 2001). Accounting and economics use specific terminology, so it is important that students have a well-developed mother tongue to understand the text correctly and are able to analyse it and draw conclusions. For this reason, a combination of these four subjects was chosen to find gender differences and correlations at the level of all four subjects (Berková et al., 2020). School achievement represents an ordinal variable.

The fifth section included student demographic information, such as gender, high school name, city, and state. Respondents consented to the processing of their personal information. For the purposes of this research, any data that could be used to identify the respondent were encrypted or removed.

The validity of the questionnaire was established through a focus group research method conducted with 8 students in face-to-face mode (Langfeldt and Kyvik, 2011). It was one focal group of 90 minutes with two moderators - researchers. Due to the positive outcome of the content validation, no further methods were resorted to. The reliability of each subscale related to one construct was measured by calculating Cronbach's alpha. The results for each construct are shown below:

- Motivational factors for the subject of economics 0.523
- Motivational factors for the subject of accounting 0.630
- Teacher competences in the subject of economics 0.843
- Teacher competences in the subject of accounting 0.867
- Preparation time of the pupil 0.715
- School success rate in subjects 0.818

Appendix to this paper provides more detailed results of the Cronbach's alpha when the variable would be removed.

Data analysis

The original data obtained from the questionnaire survey are of several types. The variables expressing the descriptive characteristics of the respondents, i.e., high school name, grade, city and state are nominal variables, of which the gender variable is a dichotomous variable and is used as a sorting factor to conduct all comparative analyses. The data contain mostly numerical ordinal variables, i.e., student motivational factors and teacher professional competencies are expressed on a five-level Likert scale from -2 to 2, and educational achievement is represented by the final grade at the end of Year 2. Both types are commonly treated as numerical variables, the level is described by an arithmetic mean (Jamieson, 2004). Since these traits do not meet the requirement of normality (verified by the Shapiro-Wilk test), but meet the requirement of homogeneity of variances (verified by the Levene's test), the Mann-Whitney U test was selected from the two-sample tests to assess the hypotheses 1, 2 and 4. The Mann-Whitney U test is a non-parametric statistical test used to compare two independent groups or samples. It is often used when the assumptions for a t-test, such as normal distribution or equal variances, are not met.

This test is useful when the data is ordinal or skewed and cannot be assumed to follow a normal distribution. It is commonly employed in various fields, including social sciences, healthcare, and business research. The Mann-Whitney U test compares the distribution of values between the two groups and determines whether there is a significant difference between them. It is particularly effective in analyzing ranked or ordinal data, where the values are sorted in a particular order. To perform the Mann-Whitney U test, you need two independent samples from each group you want to compare. These samples can be of different sizes and can have unequal variances. The test assesses whether the observations from one group tend to be higher or lower than those from the other group. The test generates a U statistic and p-value as the output. The U statistic represents the rank-sum of the observations for one group relative to the other. The p-value indicates the significance of the difference between the two groups. A significant p-value indicates that there is a significant difference between the groups being compared (Norman, 2010).

Variables expressing preparation time are verbally expressed ordinal variables, frequencies were used to process them.

Gender differences in students' preparation time within the subjects studied (hypothesis 3) were tested using the Chi-square test used for categorical data. For tables larger than 2 x 2, the Chi-squared test can be used if at most 20% of the expected frequencies are less than 5 and none are less than 2. Our original data did not meet this assumption, so we merged the categories 3-4 hours; more than 4 hours into one: more than 3 hours.

To establish the correlation relationships between the variables in the case of testing hypotheses 5 - 7, a correlation matrix was constructed; the tables in the Results section show only part of it. The field inside the body of the table always contains the value of Pearson's correlation coefficient r, which is used for numerical signs; in the case of ordinal signs for students' preparation time for the course, Spearman's correlation coefficient rho was used (Jamieson, 2004; Norman, 2010). Statistical analysis was performed using SPSS software.

Null hypotheses tested at 5% significance level:

- H_{0-1}: Students' attitudes towards motivation in the case of economics and accounting do not differ by gender.
- H_{0-2}: Students' attitudes towards the teacher in the case of the economics and accounting course do not differ by gender.
- H_{0-3}: There are no gender differences in the time taken by students to prepare for the subjects of economics and accounting.
- H_{0-4}: There are no gender differences in students' school achievement in the subjects of economics and accounting.
- H_{0-5}: From a gender perspective, there is no correlation between preparation time and school success in the case of economics and accounting.
- H_{0-6}: School success in the subjects of economics, accounting, mathematics and mother tongue is not correlated in terms of gender.
- H_{0-7}: There is no correlation between students' attitudes towards motivation and school achievement in the case of economics and accounting in terms of gender.

For the purposes of the analysis, the abbreviations of the variables that appear in the explanations in the Results section:

- **F1 – F9:** motivation factors 1–9
- **C1 – C5:** competences describing the teacher's personality
- **ECO:** Economics
- **ACC:** Accounting
- **MAT:** Mathematics
- **MOT:** mother tongue (Czech and Slovak)
- **SA:** school achievement
- **TIME:** preparation time

RESULTS

Descriptive analysis of the examined variables

Gender differences were first identified using descriptive analysis methods. Differences were analysed in terms of factors that could motivate students in the learning process of economic subjects, in terms of students' evaluation of selected pedagogical competences of the teacher, in terms of school achievement in the subjects at the end of the second year of study and in terms of time spent per week on economics and accounting. The values were obtained using the average (Tables 1, 2, 3). Differences are shown by gender and subject.

Variable	Economics		Accounting	
	Boys (*n*=180)	Girls (*n*=392)	Boys (*n*=180)	Girls (*n*=392)
Attitudes towards motivation (from -2 to 2)				
F1	0.950	1.054	0.850	1.066
F2	0.250	0.268	0.239	0.217
F3	-0.850	-0.913	-0.856	-0.936
F4	1.072	1.122	0.928	1.036
F5	1.311	1.304	0.989	1.125
F6	-0.311	-0.217	-0.311	-0.079
F7	0.944	0.985	0.789	0.944
F8	1.217	1.311	1.178	1.247
F9	1.228	1.357	1.033	1.334
Teacher's evaluation (from -2 to 2)				
C1-Expertise	1.400	1.194	1.456	1.490
C2-Explanation	1.322	1.107	1.006	1.003
C3-Empathy	0.811	0.599	0.800	0.783
C4-Trust	0.944	0.890	0.994	0.949
C5-Motivator	0.894	0.702	0.717	0.819

Table 1: Descriptive statistics (ordinal variables), 2018

In economics and accounting, both boys and girls consider the following motivational factors to be the most important. These are the teacher's instructions on how to think (F7), accepting unusual ideas and analyzing mistakes (F8), and telling more than the facts (F9). None of the factors was rated negatively, i.e., as having a demotivating effect. Therefore, the factors can be considered important for increasing students' motivation in the subjects.

For the girls, the teacher's coherent way of speaking, their flexible reactions and readiness for the lesson are essential for

increasing motivation. This finding was identified at the level of both subjects. Boys perceive motivational factors differently. From their point of view, in the subject of economics, it is most important that the teacher uses practical examples, case studies with problem elements, so that they can propose a solution and discuss its correctness. In the subject of accounting, they find it most motivating when the teacher accepts their unusual ideas and can explain why a given suggestion can or cannot be put into practice, so they can learn from their own mistakes. Conversely, the least motivating factor, which creates negative attitudes in students, is explaining the curriculum using only the textbook (F3). This result appears to be true for both economics and accounting, and both boys and girls agree. Less significant, but also exclusively negative attitudes were identified in the situation when the teacher just assigns tasks and lets the students work independently (F6).

These results correspond with the student evaluation of selected pedagogical competences of teachers of economics and accounting. Both boys and girls assigned the highest values to the teacher's expertise and ability to explain the curriculum. Other competencies (empathy and confidence) averaged below 1 for both genders. A positive finding is that teacher's personality has a rather positive effect on students' motivation. Boys are more motivated to study economics by teacher's personality than girls. Conversely, to study accounting, the teacher motivates girls more than boys.

To better understand the relationship between motivation and school achievement, students' grades in economic subjects and the related subjects of mathematics and mother tongue were analysed. In terms of students' school achievement, differences were identified between boys and girls in all the subjects studied. At the end of the second year of study, on average, girls had better learning outcomes than boys (Table 2).

Variable	Boys (*n*=180)		Girls (*n*=392)	
	Mean	SD	Mean	SD
School Achievement (1-5)				
Economics	2.650	0.936	2.135	0.938
Accounting	2.611	1.038	2.097	1.027
Mathematics	2.900	1.025	2.454	1.053
Mother tongue	2.689	0.828	2.343	0.829

Table 2: Descriptive statistics (school achievement), 2018

The analysis of preparation time shows that the majority of boys and girls spend less than 1 hour or 1-2 hours per week preparing for both economics and accounting. Girls devote more time to preparation for both subjects (Table 3).

Variable	Economics		Accounting	
	Boys (*n*=180)	Girls (*n*=392)	Boys (*n*=180)	Girls (*n*=392)
Weekly preparation time				
I am not preparing at all (1)	18.89	12.50	18.33	5.10
Less than 1 hour (2)	40.00	31.63	40.56	37.76
1-2 hours (3)	30.56	30.87	27.22	32.14
3-4 hours (4)	9.44	20.41	11.11	19.90
More than 4 hours (5)	1.11	4.59	2.78	5.10

Table 3: Relative frequencies (weekly preparation time in %), 2018

Gender differences in the researched aspects in the subjects of economics and accounting (hypotheses 1-4)

Using Mann Whitney U test, statistical gender differences at the 5% level of significance were examined in respondents' attitudes towards motivational factors, teacher's personality and school achievement in the subjects of economics and accounting. Further, gender differences in weekly preparation time were also found using Chi-square test. Table 4 shows the results of the statistical tests.

Attitudes towards motivation

In the subject of economics, no statistically significant gender differences were found, i.e., boys and girls perceived the motivating or demotivating effect of the methods of teaching economics in the same way. In the subject of accounting, there are more statistically significant gender differences. Factor F1 (teacher emphasizes the applicability of the curriculum to real situations) is motivating for both boys and girls, but the intensity of motivation is perceived more by girls (Table 1). A significant difference was also found in the case of F6 (teacher sets a task, letting students work independently). For both boys and girls, this factor is demotivating, but more intensity was found for boys (Table 1). Another significant difference is found for factor F7 (teacher guides students to think critically), where again greater intensity of the motivating effect was found for girls compared to boys.

In all cases of significant gender differences, girls assigned higher positive values (i.e., greater intensity of motivation) or lower intensity of demotivation to motivational factors than boys. The researchers reject the hypothesis H_{0-1} with 95% reliability for F1, F6, F7 and F9 for the subject of accounting.

Variable	Economics (*p*)	Accounting (*p*)
Attitudes towards motivation		
F1	0.107	0.008
F2	0.787	0.948
F3	0.615	0.482
F4	0.409	0.152
F5	0.754	0.115
F6	0.380	0.039
F7	0.321	0.019
F8	0.132	0.268
F9	0.073	0.003
Teacher's evaluation		
C1-Expertise	0.038	0.679
C2-Explanation	0.065	0.978
C3-Empathy	0.087	0.871
C4-Trust	0.501	0.671
C5-Motivator	0.239	0.363
Weekly preparation time	<0.001	<0.001
School achievement	<0.001	<0.001

Table 4: Mann Whitney U test and Chi-square test – significance of gender differences, 2018

Attitudes towards the personality of the teacher of economics and accounting

In the case of the subject of economics, a statistically significant gender difference in attitudes was found only for the teacher's expertise (Table 3). According to the average values reported in Table 1, boys perceive the economics teacher as a greater expert than girls. In the case of the subject of accounting, girls and boys rate the personality of the teacher equally – no significant gender differences were found for any of the five pedagogical competencies examined. However, for the subject of economics, the opposite result holds true. Girls assigned lower values to all teacher competencies than boys. The researchers reject hypothesis H_{0-2} with 95% reliability in the case of the economics teachers' expertise.

Preparation time and school achievement in the subjects

Statistically significant differences were found in weekly preparation time for each subject, with girls spending more time studying the subject than boys (Table 3). A weak relationship was found between the variables. Cramer's V is 18% in the case of economics and 23.5% in the case of accounting. The researchers reject the hypothesis H_{0-3} with 95% reliability in the case of both economics and accounting.

Significant gender differences were found in school achievement in both economics and accounting at the end of the second year of study. In both subjects, girls performed better than boys (Table 2). This result corresponds with the results shown in the students' attitudes towards motivational factors, whereby boys perceive the motivational intensity of factors at a lower level than girls and, therefore, may have a looser attitude towards their studies compared to girls. This may be related to their poorer grades and the lesser amount of time they devote to preparation for the subjects (Tables 2 and 3). The researchers reject hypothesis H_{0-4} with 95% reliability in the case of economics and accounting.

Correlations between school achievement and weekly preparation time (hypotheses 5-6)

In order to better portray the relationship between preparation time and students' school achievement in economic subjects, a correlation analysis using Spearman's correlation coefficient *rho* was performed at the 5% significance level. The results are illustrated in the correlation matrix (Table 5).

Girls		SA2_ECO	SA2_ACC	Boys		SA2_ECO	SA2_ACC
Time_ECO	*rho*	0.006	0.090	Time_ECO	*rho*	0.015	0.045
	p	0.906	0.077		*p*	0.841	0.551
Time_ACC	*rho*	0.007	0.147	Time_ACC	*rho*	0.023	-0.018
	p	0.892	0.003		*p*	0.760	0.808

Table 5: Spearman's correlation coefficient for school achievement and weekly preparation time (hypothesis 5), 2018

The correlation analysis revealed a statistically significant association between school achievement and preparation time in the subject of accounting for girls ($p = 0.003$). For this group of variables, the dependency ratio (r) was found to be 0.147, i.e., a 15% direct dependence, indicating that girls with lower achievement spend more time on preparation for the subject of accounting. However, the correlation is weak. This result corresponds with the girls' attitude towards motivational factors in accounting. It has been shown that girls would be motivated by the teacher's ways of explanation that would help them understand the subject

matter. At the same time, girls rated accounting teachers better than economics teachers in terms of pedagogical competence. Other correlations between variables were not statistically significant. The researchers reject the hypothesis H_{0-5} at 95% significance level in the case of girls in the subject of accounting.

Furthermore, at the 5% significance level, Pearson's correlation coefficient r was used to determine the connections in school achievement between economics, accounting, mathematics, and mother tongue. The results are shown in the correlation matrix (Table 6).

	Girls				Boys		
	SA2_ECO	SA2_ACC	SA2_MAT		SA2_ECO	SA2_ACC	SA2_MAT
SA2_ACC	0.589	-	-	SA2_ACC	0.647	-	-
	<0.001	-	-		<0.001	-	-
SA2_MAT	0.505	0.591	-	SA2_MAT	0.638	0.620	-
	<0.001	<0.001	-		<0.001	<0.001	-
SA2_MOT	0.488	0.451	0.356	SA2_MOT	0.407	0.411	0.338
	<0.001	<0.001	<0.001		<0.001	<0.001	<0.001

Table 6: Pearson's correlation coefficient for school achievement among subjects (hypothesis 6), 2018

Correlations were found for both boys and girls for all combination pairs of subjects ($p < 0.001$). The correlation of school achievement between subjects is stronger for boys than for girls. For girls, the strongest correlation is between grades in accounting and mathematics ($r = 0.591$) and grades in accounting and economics ($r = 0.589$). Significant gender differences and higher intensity of motivation were found for girls for factors that led to an understanding of the curriculum due to the teacher's helpful explanation. In accounting, cognitive operations are based on mathematical laws and the theory of logic. This may explain the less strong degree of dependence of performance between these subjects. For boys, the strongest correlations were found between grades in accounting and economics ($r = 0.647$), grades in economics and mathematics ($r = 0.638$), and grades in accounting and mathematics ($r = 0.620$). The correlations are weaker for both girls and boys when taking the mother tongue into account, which is consistent with the fact that, cognitively, economic subjects are more similar to mathematics. Hypothesis H_{0-6} is rejected at 95% reliability for both boys and girls for all pairs of subjects.

Correlation between attitudes towards motivational factors and school achievement in economics and accounting (hypothesis 7)

Pearson's correlation coefficient r was used to determine, at the 5% significance level, how educational achievement at the end of the second year of study in the subjects of economics and accounting is related to boys' and girls' attitudes towards motivation in these subjects (Tables 7 and 8).

		F1	F2	F3	F4	F5	F6	F7	F8	F9
SA2_ECO Boys	*r*	-0.028	-0.149	-0.150	-0.106	-0.220	-0.084	-0.088	-0.071	0.015
	p	0.714	0.045	0.045	0.155	0.003	0.262	0.242	0.346	0.841
SA2_ECO Girls	*r*	-0.108	0.101	0.100	-0.039	-0.165	-0.045	-0.175	-0.133	-0.147
	p	0.033	0.046	0.047	0.440	0.001	0.375	<0.001	0.008	0.004

Table 7: Pearson's correlation coefficient for attitudes towards motivational factors and school achievement in economics (hypothesis 7), 2018

The correlation between learning outcomes and students' perceptions of motivational factors in the subject of economics shows a predominantly negative relationship for significant correlations (Table 7). This means that, to a greater extent, students with better school results (i.e., lower grades) assigned higher values to the motivational factors to express the intensity of the motivational effect. Gender differences are evident in the case of economics, as girls have more significant correlations between variables than boys, or they differ in direct or indirect dependence. It should be noted that the strength of the positive or negative dependence between the variables is weak.

In the case of boys, a negative correlation between school achievement and factors F2, F3, and F5 was found at the 5% significance level. This means that boys with better grades have a positive attitude towards motivation when the teacher explains the basic curriculum without deeper understanding, which can result in a good grade. Such a tendency was also shown in the case where the teacher only explained the textbook curriculum and in the case where he used practical examples and the student had to suggest solutions to problems. The strongest correlation was found for F5 ($r = -0.220$).

In the case of girls, both positive and negative relationships between school achievement and perceived motivational factors in the subject of economics were found at the 5% significance level. A negative relationship was found between school achievement and the following factors: F1, F5, F7-F9. This means that girls with better achievement have a positive attitude towards motivation in cases where the teacher presents the curriculum in a way that the student understands the relevance to real-life situations, uses real-life examples, where the student suggests solutions to problems, where the teacher discusses mistakes with the students, where the teacher communicates more than facts and leads the students to think critically. A positive relationship was shown between learning outcomes and factors F2 and F3. That is, girls with poorer school results (i.e., higher grades)

have a positive attitude towards motivation when the teacher explains the basic curriculum without deeper understanding using the textbook, which can result in a good grade.

A major gender difference was identified in the subject of economics, where factors F2 and F3 are motivating for girls with poorer school results and, on the contrary, for boys with better results.

		F1	F2	F3	F4	F5	F6	F7	F8	F9
SA2_ACC	*r*	-0.317	-0.103	-0.122	-0.233	-0.212	-0.217	-0.196	-0.212	-0.177
Boys	*p*	<0.001	0.167	0.102	0.002	0.004	0.003	0.009	0.004	0.018
SA2_ACC	*r*	-0.281	0.080	0.101	-0.152	-0.232	-0.109	-0.171	-0.272	-0.195
Girls	*p*	<0.001	0.112	0.045	0.003	<0.001	0.030	<0.001	<0.001	<0.001

Table 8: Pearson's correlation coefficient for attitudes towards motivational factors and school achievement in accounting (hypothesis 7), 2018

Gender differences in the subject of accounting are to a lesser extent than in the subject of economics. The correlation between learning outcomes and students' perception of motivational factors again shows a negative relationship between the variables except for factor F3 in the case of girls. The degree of dependence of the variables is relatively weak.

In the case of boys, a negative correlation between school achievement and factors F1, F4-F9 was found at the 5% significance level. This means that boys with better grades have a positive attitude towards motivation when the teacher explains the practicality of the curriculum and uses simple examples but gives space for questions and discussion, uses practical examples; when the student has to suggest solutions to problems, when he/she discusses mistakes with students, when he/she communicates more than facts and leads students to think critically. On the other hand, this relationship also applies to the case where the teacher only assigns individual work tasks.

In the case of girls, a negative correlation between school achievement and the perception of the same motivational factors as boys was found at the 5% level of significance. The level of the dependent variables is higher for boys. Positive dependence between the variables was found in the case of F3 for girls. This means that positive attitudes towards motivation in the case where the teacher only assigns individual work tasks are more likely for girls with poorer grades. The main gender difference in this tendency was found in the subject of accounting (similar to the subject of economics).

The researchers reject the hypothesis $H_{0\text{-}7}$ at 95% reliability for both girls and boys for both subjects.

DISCUSSION

The study examined gender differences in school achievement and attitudes of secondary school students towards motivation in economic subjects. For the actual analysis, 572 valid questionnaires were used. To understand the context better, the variables were supplemented with additional data on attitudes towards the teacher of the subject, on the time students spend studying economic subjects, and on learning outcomes in mathematics and mother tongue, which are essential for the development of economic thinking and understanding of information (OECD, 2014; Wiliam, 2011; Watts and Zimmerman, 1978). The study based its research approach on several similar studies that have described significant gender differences not only in attitudes towards motivation and school achievement, but also examined attitudes towards teacher's personality and correlations between variables (Alharbi et al., 2020; Berková et al., 2018; Perelygina et al., 2020). The researchers reject the null hypotheses 3, 4, 6, and 7. The researchers reject the null hypotheses 1 and 5 for the subject of accounting. They reject the null hypothesis 2 in the case of the economics teacher's expertise.

Gender differences in school achievement in economic subjects

Significant gender differences were found in school achievement in both economics and accounting. The results showed that girls perform better than boys in both subjects i.e., a similar relation exists in the economic subjects as has been shown in other educational domains (Deary et al., 2007; Voyer and Voyer, 2014; Eriksson et al., 2020; Coetzee et al., 2020). The results are consistent with the research of Pintrich and de Groot (1990), which points to several important factors influencing learning outcomes. These include self-discipline, which is also confirmed by Duckworth and Seligman (2006), noting that girls perform better on average in these behavioral skills. Other factors are confidence in one's own performance and fear of exams (Erdem-Keklik and Keklik, 2014), which is higher for girls than for boys (Chmelárová et al., 2018). These factors seem to be reflected in the result of this study, which showed significant differences in the amount of preparation time i.e., that girls spend more time per week studying both subjects compared to boys. The reasons for these findings can also be supported by other research. Boys, compared to girls, are more likely to express aggressive behavior and display more developmental difficulties and negative attitudes towards learning (Lansford et al., 2012), suggesting that girls generally adapt more easily to the school environment and have stronger social responsibility (Giota and Bergh, 2021).

There was a direct relationship between school achievement and preparation time in the case of accounting, $r = 0.147$ (girls with lower achievement spend more time on the preparation for accounting). This finding may shed a different light on the assumption made in the previous paragraph, where girls' more responsible approach to their studies was associated with their better performance compared to boys. Anyway, for boys, these relationships were not found at all in accounting, and in the case of economics, the link between school achievement and preparation time was not found for either girls or boys. A different result in accounting can be illustrated by the study of Fields et al. (2001), which discusses the theory of accounting and the approaches underlying the discipline and relies more on the theory of logic. The existence of significant gender

differences in subjects aimed at developing abstract-visual thinking is also supported by Eriksson et al. (2020). This may be the reason why a significant correlation was found in the case of girls with poorer grades in that subject. The above relationships were further explored through correlation analysis of school achievement between economic subjects, mathematics, and mother tongue. This idea is based on the results of PISA (OECD 2010), which showed a relationship between numeracy and financial literacy. Furthermore, the idea is supported by research verifying the relationships between language skills (especially reading comprehension) and the development of critical thinking (Wiliam, 2011), which is important for studying economics. This study found correlations in the case of school achievement between accounting, economics, mathematics, and mother tongue. Girls perform better than boys in all the subjects studied, which is consistent with the study of Marmeleira et al. (2020) in the case of Mathematics and Foreign Language.

Gender differences in attitudes towards motivation in economic subjects

Significant gender differences were found in attitudes towards motivation in both subjects. Girls attributed a higher intensity of motivation (or a lower intensity of demotivation) to the factors assessed than boys. In the case of accounting, the significant factors that elicited a greater motivational effect for girls than boys included several modes of explanation – the applicability of the curriculum to everyday situations, if the teacher guides students to think critically about the topic and understand the content, if the teacher communicates more than facts. In this last factor, the same significant gender difference was also found in the case of economics. The results of this study are in line with relatively recent research (Bru et al., 2021) that has shown that the choice of teaching methods and resources has an impact on increasing girl's activity and engagement in the classroom. Specifically, it supports and assists in a better understanding of the curriculum. What really motivates girls in accounting is the teacher's supportive approach. For boys, activity in the classroom can be increased by a clearer structure of learning activities. Significant gender differences in attitudes towards motivation in accounting were also found when the teacher let students work independently and only corrected inconsistencies. Such a process was demotivating for both girls and boys, but boys perceived it more negatively. The results are consistent with evidence of the impact of activating teaching methods on attitudes towards motivation (Lebid and Shevchenko, 2020).

Significant differences were also found between girls and boys in their perception of the teacher's personality, which was defined in this research by five pedagogical competencies – expertise, clear explanation, empathy, trust, and ability to motivate. A statistically significant gender difference was found in the case of the expertise of the teacher of economic subjects, with boys rating the teacher as more expert than girls. These differences can again be explained by the nature of the two subjects and the approaches underlying both disciplines (Watts and Zimmerman, 1978). Girls may perceive the subject as more difficult and may have more respect for it, which was reflected in their ratings of the teacher's expertise and also in their attitudes towards motivational factors. The personality of the teacher plays an important role in the teaching process. Also, the motivation of the students to study the subject depends on the motivation of the teacher to perform the profession (Perelygina et al., 2020).

A number of studies have found that school achievement and motivation are related, with the teacher and the way they prepare students for professional life being significant factors (Chetty et al., 2014; Rivkin et al., 2005; Rockoff, 2004; Fung et al., 2017). This study found that students' school achievement in economic subjects is significantly related to their attitudes towards motivation and also to their learning outcomes in cognitively related subjects such as mathematics and their mother tongue. Boys and girls with better performance are motivated by the teacher's approaches that keep students' attention during the lesson and lead to a deeper understanding of the curriculum and its relevance to real-life situations. This tendency was found in the subjects of economics and accounting. Another interesting tendency is shown in both subjects, which was demonstrated only in the case of girls with poorer grades. This group has a positive attitude towards motivation in the cases where the teacher explains the basic curriculum using the textbook, which can result in a good grade. In accounting lessons, this group is also motivated by individual work.

CONCLUSION

The study highlighted the importance of studying the relationship between school achievement and students' attitudes towards motivation in secondary economic education in the context of gender. The limitation of this research is mainly in its focus on students' attitudes, i.e., their perceptions of motivation factors in studying economic subjects. The selected motivation factors focused on the attractiveness of the teacher's delivery of the lesson and the ways of holding students' attention. Thus, the examined motivation factors did not include all the attributes of motivation. The research also focused on students' attitudes towards the teachers of economics and accounting. Again, it was about students' perception, and some teacher competences were selected. The focus was on expertise, clear explanations, empathy, trust, and the ability to motivate. In this context, the research was not exhaustive, and leaves space for exploring other pedagogical, psychological and didactic aspects related to the links between student motivation, teacher's personality and school achievement. It would also be interesting to examine school achievement in terms of students' intellectual level and real abilities, which also determine it greatly. Achievement expressed in terms of summative assessment (grade) also does not guarantee a complex assessment of school performance, but is only one of its possible indicators.

The research results in this study broaden the theoretical assumptions about students' school achievement in the context of gender in the area of secondary school economic education. In line with the OECD study (2009), this type of research can help teachers of economics and accounting to improve their student performance. The study discovered several aspects that teachers can use in their teaching practice to differentiate

their approaches to students as well as their choice of teaching and learning methods according to students' gender, thus motivating them better to become active learners of economics and accounting, which should consequently lead to better school achievement of boys and girls in these subjects. Besides these practical implications, the study also contributes to the theoretical understanding of gender differences and their impact on relations between the approaches and teaching and learning methods that teachers prefer and their students' school achievement (e.g., Bru et al., 2021; Lebid and Shevchenko, 2020) by implying how methods of teaching and learning can be better adjusted to boys' and girls' learning styles and how teachers' behavior in class in suitably differentiated ways can also help to motivate girls and boys better.

ACKNOWLEDGEMENT

This article was supported by the the Faculty of Finance and Accounting in the framework of institutional support of Prague University of Economics and Business under Grant IP100040; and Cultural and Educational Grant Agency of the Ministry of Education, Science, Research and Sport of the Slovak Republic under Grant 025EU-4/2023.

REFERENCES

Agyei, D. D. and Eyiah-Bediako, F. (2008) 'Gender differences in mathematics performance: Some fresh insights', *Gender & Behaviour*, Vol. 6, No. 1, pp. 1519–1531. https://doi.org/10.4314/gab.v6i1.23401

Alharbi, M. S., Almatham, A., Alsalouli, M. S. and Hussein, H. B. (2020) 'Mathematics Teachers' Professional Traits that Affect Mathematical Achievement for Fourth-grade Students according to the TIMSS 2015 Results: A Comparative Study among Singapore', *International Journal of Educational Research*, Vol. 104, p. 101671. https://doi.org/10.1016/j.ijer.2020.101671

Akin, A. and Kurbanoglu, I. N. (2011) 'The Relationships between Math Anxiety, Math Attitudes, and Self-efficacy: A Structural Equation Model', *Studia Psychologica*, Vol. 53, No. 3, pp. 263–273.

Berková, K., Krejčová, K. and Novák, J. (2020) 'Students' Attitudes to Selected Subjects in Secondary Economic Education', *Problems of Education in the 21st Century*, Vol. 78, No. 1, pp. 9–23. https://doi.org/10.33225/pec/20.78.09

Berková, K., Borůvková, J. and Lízalová, L. (2018) 'Motivation of Students of Economic and Technical Study Programmes as a Tool of Competitiveness of Universities and Colleges: Empirical Study', *Journal on Efficiency and Responsibility in Education and Science*, Vol. 11, No. 4, pp. 72–77. https://doi.org/10.7160/eriesj.2018.110401

Berková, K. and Krejčová, K. (2016) 'Effect of teachers´ abilities on students´ motivation with varying levels of intellectual abilities in the economics', *Journal on Efficiency and Responsibility in Education and Science,* Vol. 9, No. 3, pp. 81–87. https://doi.org/10.7160/eriesj.2016.090304

Bru, E., Virtanen, T., Kjetilstad, V. and Niemiec, C. P. (2021) 'Gender Differences in the Strength of Association between Perceived Support from Teachers and Student Engagement', *Scandinavian Journal of Educational Research*, Vol. 65, No. 1, pp. 153–168. https://doi.org/10.1080/00313831.2019.1659404

Chetty, R., Friedman, J. and Rockoff, J. (2014) 'Measuring the impacts of teachers II: Teacher value-added and student outcomes in adulthood', *American Economic Review*, Vol. 104, No. 9, pp. 2633–2679. http://dx.doi.org/10.1257/aer.104.9.2633

Chmelárová, Z., Čonková, A. and Šoltés, E. (2018) 'Strach publičnych vystuplenij u učenikov i studentov v podrostkovom i junošeskom vozraste. (Fear of Public Speaking among Students in their Teenage Years)', *Psychological Science and Education*, Vol. 23, No. 4, pp. 70–79. https://doi.org/10.17759/pse.2018230407

Coetzee, D., Pienaar, A. and van Wyk, Y. (2020) 'Relationship between academic achievement, visual-motor integration, gender and socio-economic status: North-West Child Health Integrated with Learning and Development study', *South African Journal of Childhood Education*, Vol. 10, No. 1, p. a646. https://doi.org/10.4102/sajce.v10i1.646

Deary, I. J., Strand, S., Smith, P. and Fernandes, C. (2007) 'Intelligence and educational achievement', *Intelligence*, Vol. 35, No. 1, pp. 13–21. https://doi.org/10.1016/j.intell.2006.02.001

Demirbas, O. O. and Demirkan, H. (2007) 'Learning styles of design students and the relationship of academic performance and gender in design education', *Learning and Instruction*, Vol., 17, No. 3, pp. 345–359. https://doi.org/10.1016/j.learninstruc.2007.02.007

Duckworth, A. L., Seligman, M. E. (2006) 'Self-discipline gives girls the edge: gender in self-discipline, grades, and achievement test scores', *Journal of Educational Psychology*, Vol. 98, No. 1, pp. 198–208. https://doi.org/10.1037/0022-0663.98.1.198

Erdem-Keklik, D. and Keklik, I. (2014) 'High School Students´ Achievement Goals: Assessing Gender, Grade Level and Parental Education', Çukurova *University Faculty of Education Journal*, Vol. 43, No. 1, pp. 63–73. https://doi.org/10.14812/cufej.2014.005

Eriksson, K. (2020) 'Gender Differences in the Interest in Mathematics Schoolwork Across 50 Countries', *Frontiers in Psychology,* Vol. 11, 578092. https://doi.org/10.3389/fpsyg.2020.578092

Eriksson, K., Björnstjerna, M. and Vartanova, I. (2020) 'The Relation Between Gender Egalitarian Values and Gender Differences in Academic Achievement', *Frontiers in Psychology*, Vol. 11, p. 236. https://doi.org/10.3389/fpsyg.2020.00236

Fields, T. D., Lys, T. Z. and Vincent, L. (2001) 'Empirical Research on Accounting Choice', *Journal of Accounting and Economics*, Vol. 31, No. 1–3, pp. 255-307. https://doi.org/10.1016/S0165-4101(01)00028-3

Fontana, D. (2014) *Psychologie ve* školní *praxi*. (Psychology in School Educational Process), Praha: Portál.

Fung, D., Kutnick, P., Mok, I., Leung, F., Pok-Yee Lee, B., Mai, Y. Y. and Tyler, M. T. (2017) 'Relationships between teachers' background, their subject knowledge and pedagogic efficacy, and pupil achievement in primary school mathematics in Hong Kong: An indicative study', *International Journal of Educational Research*, Vol. 81, pp. 119–130. https://doi.org/10.1016/j.ijer.2016.11.003

Giota, J. and Bergh, D. (2021) 'Adolescent Academic, Social and Future Achievement Goal Orientations: Implications for Achievement by Gender and Parental Education', *Scandinavian Journal of Educational Research*, Vol. 65, No. 5, pp. 831–850. https://doi.org/10.1080/00313831.2020.1755360

Habók, A., Magyar, A., Németh, M. B. and Csapó, B. (2020) 'Motivation and self-related beliefs as predictors of academic achievement in reading and mathematics: Structural equation models of longitudinal data', *International Journal of Educational Research*, Vol. 103, p. 101634. https://doi.org/10.1016/j.ijer.2020.101634

Jamieson, S. (2004) 'Likert scales: How to (ab)use them', *Medical Education*, Vol. 38, No. 12, pp. 1217–1218. https://doi.org/10.1111/j.1365-2929.2004.02012.x

Jurečka, V., Hon, M., Jánošíková, I., Kolcunová, E. and Spáčilová, L. (2013) *Mikroekonomie*. (Microeconomics), Prague: Grada Publishing.

Krpálek, P. and Krpálková Krelová, K. (2016) 'Possibilities for Developing Business Potential in Economic Education. Examples of Implementation in Slovakia and the Czech Republic', *Economics & Sociology,* Vol. 9, No. 4, pp. 119–133. https://doi.org/10.14254/2071-789X.2016/9-4/7

Langfeldt, L. and Kyvik, S. (2011) 'Researchers as evaluators: tasks, tensions and politics', *Higher Education*, Vol. 62, No. 2, pp. 199–212. https://doi.org/10.1007/s10734-010-9382-y

Lansford, J. E., Skinner, A. T., Sorbring, E., Giunta, L. D., Deater-Deckard, K. and Dodge, K. A., et al. (2012) 'Boys' and Girls' relational and physical aggression in nine Countries', *Aggressive Behavior*, Vol. 38, No. 4, pp. 298–308. https://doi.org/10.1002/ab.21433

Lebid, A. E. and Shevchenko, N. A. (2020) 'Gender Differentiation in Perceiving and Evaluating the Interactive Methods of Education by the Engineering Students', *European Journal of Contemporary Education*, Vol. 9, No. 4, pp. 819–826. https://doi.org/10.13187/ejced.2020.4.819

Lindberg, S. M., Hyde, J. S., Petersen, J. L. and Linn, M. C. (2010) 'New trends in gender and mathematics performance: A meta-analysis', *Psychological Bulletin*, Vol. 136, No. 6, pp. 1123–1135. https://doi.org/10.1037/a0021276

Marmeleira, J., Folgado, H., Guardado, I. M. and Batalha, N. (2020) 'Grading in Portuguese secondary school physical education: assessment parameters, gender differences and associations with academic achievement', *Physical Education and Sport Pedagogy*, Vol. 25, No. 2, pp. 119–136. https://doi.org/10.1080/17408989.2019.1692807

Norman, G. (2010) 'Likert scales, levels of measurement and the "laws" of statistics', *Advances in Health Sciences Education*, Vol. 15, No. 5, pp. 625–632. https://doi.org/10.1007/s10459-010-9222-y

Núñez-Peña, M. I., Suárez-Pellicioni, M. and Bono, R. (2016) 'Gender differences in test anxiety and their impact on higher education students´ academic achievement', in *Procedia – Social and Behavioral Sciences,* Amsterdam: Elsevier Ltd. https://doi.org/10.1016/j.sbspro.2016.07.023

OECD (2009) *Equally prepared for life? How 15-year-old boys and girls perform in school*, Paris: OECD Publishing. Available: https://www.oecd.org/pisa/pisaproducts/42843625.pdf

OECD (2010) *Mathematics teaching and learning strategies in PISA*. Paris: OECD Publishing. Available: https://doi.org/10.1787/9789264039520-en

OECD (2014) *PISA 2012 Results: Students and Money. Financial Literacy Skills for the 21st Century (Volume VI),* PISA: OECD Publishing. Available: https://doi.org/10.1787/19963777

Owoseni, A., Ibem, E. and Opoko, A. (2020) 'Socio-Economic Profiling of Students and its Impact on Learning Outcomes', *International Journal of Emerging Technologies in Learning*, Vol. 15, No. 19, pp. 213–222. https://doi.org/10.3991/ijet.v15i19.14649

Perelygina, L., Rayevska, I., Chusova, O., Horbatiuk, O. and Tomenko, M. (2020) 'Tools and Mechanisms of Motivation Support in the Process of Training Students of Higher Educational Institutions', *Applied Linguistics Research Journal*, Vol. 4, No. 9, pp. 148–158. https://doi.org/10.14744/alrj.2020.14892

Pintrich, P. R. and de Groot, E. V. (1990) 'Motivational and self-regulated learning components of classroom academic performance', *Journal of Educational Psychology*, Vol. 82, No. 1, pp. 33–40. https://doi.org/10.1037/0022-0663.82.1.33

Rivkin, S., Hanushek, E. and Kain, J. (2005) 'Teachers, schools, and academic achievement', *Econometrica*, Vol. 73, No. 2, pp. 417–458. https://doi.org/10.1111/j.1468-0262.2005.00584.x

Rockoff, J. (2004) 'The impact of individual teachers on student achievement: Evidence from panel data', *The American Economic Review*, Vol. 94, No. 2, pp. 247–252. https://doi.org/10.1257/0002828041302244

Silva-Arias, A. C., Sarmiento-Espinel, J. A. and González-Cogollos, L. V. (2020) 'Gender Inequalities in Academic Achievement among Internally Displaced children in Colombia', *Revista Latinoamericana de Ciencias Sociales, Niñez y Juventud*, Vol. 18, No. 2, pp. 1–19. http://dx.doi.org/10.11600/1692715x.18213

Stoet, G. and Geary, D. C. (2013) 'Sex differences in mathematics and reading achievement are inversely related: within-and across-nation assessment of 10 years of PISA data', *PLoS ONE,* Vol. 8, No. 3, p. e57988. https://doi.org/10.1371/journal.pone.0057988

Valente, S., Veiga-Branco, A., Rebelo, H., Lourenço, A. A. and Cristovão, A. M. (2020) 'The relationship between emotional intelligence ability and teacher efficacy', *Universal Journal of Educational Research,* Vol. 8, No. 3, pp. 916–923. https://doi.org/10.13189/ujer.2020.080324

Voyer, D. and Voyer, S. D. (2014) 'Gender Differences in Scholastic Achievement: A Meta-Analysis', *Psychological Bulletin*, Vol. 140, No. 4, pp. 1174–1204. http://dx.doi.org/10.1037/a0036620

Watts, R. L. and Zimmerman, J. L. (1978) 'Towards a Positive Theory of the Determination of Accounting Standards', *The Accounting Review*, Vol. 53, No. 1, pp. 112–134.

Wiliam, D. (2011) *Embedded Formative Assessment.* USA: Solution Tree Press.

Zhang, D. and Wang. C. (2020) 'The relationship between mathematics interest and mathematics achievement: mediating roles of self-efficacy and mathematics anxiety'. *International Journal of Educational Research*, Vol. 104, p. 101648. https://doi.org/10.1016/j.ijer.2020.101648

APPENDIX

Motivational factors (Economics)

	If item dropped Cronbach's α
F1_ECO	0.477
F2_ECO	0.553
F3_ECO	0.580
F4_ECO	0.454
F5_ECO	0.453
F6_ECO	0.521
F7_ECO	0.443
F8_ECO	0.460
F9_ECO	0.480

Motivational factors (Accounting)

	If item dropped Cronbach's α
F1_ACC	0.568
F2_ACC	0.646
F3_ACC	0.675
F4_ACC	0.568
F5_ACC	0.568
F6_ACC	0.644
F7_ACC	0.564
F8_ACC	0.579
F9_ACC	0.582

Teacher competences (Economics)

	If item dropped Cronbach's α
C1_ECO	0.821
C2_ECO	0.798
C3_ECO	0.799
C4_ECO	0.819
C5_ECO	0.817

Teacher competences (Accounting)

	If item dropped Cronbach's α
C1_ACC	0.853
C2_ACC	0.838
C3_ACC	0.830
C4_ACC	0.838
C5_ACC	0.838

Preparation time for the subject

	If item dropped Cronbach's α
T_ECO	0.638
T_ACC	0.587
T_MAT	0.701
T_MOT	0.674

School success rate in the subjects

	If item dropped Cronbach's α
SA2_ECO	0.745
SA2_ACC	0.735
SA2_MAT	0.774
SA2_MOT	0.820

MOTIVATION AND BEHAVIORAL ENGAGEMENT: THE MEDIATING ROLE OF MATHEMATICS SELF-EFFICACY IN PRIMARY EDUCATION

Achmad Hidayatullah[1]✉
Ratno Abidin[1]
Abdul Muqit[2]

[1]Universitas Muhammadiyah Surabaya, Indonesia

[2]State Polytechnic of Malang, Indonesia

✉ achmadhidayatullah@um-surabaya.ac.id

ABSTRACT

The primary objective of this study is to explore the interrelation between motivation, self-efficacy, and behavioral engagement in primary school mathematics learning. This study also examines the mediating role of self-efficacy in the relationship between motivation and behavioral engagement. The research involved 660 fifth and sixth grade students in four schools in Surabaya, Indonesia. The data collected was analyzed using structural equation modeling. The study revealed that motivation is key to enhancing students' self-efficacy and behavioral engagement during mathematics learning. Additionally, self-efficacy was found to be linked with students' behavioral engagement. Furthermore, self-efficacy was identified as a mediator in the relationship between motivation toward mathematics and behavioral engagement during mathematics learning. The study provides valuable insights into these variables in the Indonesian context, particularly in elementary schools for mathematics learning. The research also discusses the implications of the study for teaching practices.

KEYWORDS

Self-efficacy, motivation, behavioral engagement, mathematics learning

HOW TO CITE

Hidayatullah A., Abidin R., Muqit A. (2024) 'Motivation and Behavioral Engagement: The Mediating Role of Mathematics Self-Efficacy in Primary Education', *Journal on Efficiency and Responsibility in Education and Science*, vol. 17, no. 3, pp. 237–246. http://dx.doi.org/10.7160/eriesj.2024.170306

Article history
Received
October 20, 2023
Received in revised form
June 10, 2024
Accepted
August 23, 2024
Available on-line
September 30, 2024

Highlights

- *Students' behavioral engagement was explored in this study.*
- *Motivation is positively associated with self-efficacy and behavioral engagement.*
- *Self-efficacy mediated the association between behavioral engagement and motivation.*
- *There is an indirect association between motivation and behavioral engagement.*

INTRODUCTION

Over the past decade, there has been a growing recognition of the importance of students' behavioral engagement in mathematics education. Behavioral engagement is a critical factor that has been shown to affect students' performance in this subject. It refers to students' active class participation and practice sessions (Orji and Ogbuanya, 2022). Research indicates that those who actively manage their learning and engage in class achieve better results (Chong et al., 2018). Self-determination theory (Ryan and Deci, 2000) suggests that students' motivation is crucial to their outcomes in mathematics education. In other words, when students become motivated, they actively participate in mathematics learning and are more likely to achieve their goals (Józsa et al., 2022; Welesilassie and Nikolov, 2022).

Numerous studies in the literature have explored how motivation and engagement are related. For example, Orji and Ogbuanya's (2022) study in Nigeria found that students with high motivation were more likely to be engaged in their tasks because positive emotions encouraged them during the activity. Similarly, Flunger et al. (2022) discovered that in Germany, students invested greater effort in mathematics lessons when their interest was stimulated due to their high motivation. Additionally, research has suggested that students' behavioral engagement in mathematics learning is closely linked with their self-efficacy. According to self-efficacy theory (Bandura, 1997), those with high confidence and self-judgment about their abilities tend to put greater effort into academic learning and achieve their goals. Conversely, students without self-efficacy have a higher risk of negatively perceiving mathematics (Damrongpanit, 2019). Durksen et al. (2016) recommended that self-efficacy or mathematics agency is crucial in promoting positive motivation and behavioral engagement during mathematics learning.

Previous studies (Durksen et al., 2016; Flunger et al., 2022; Orji and Ogbuanya, 2022) have shed light on the close relationship between students' behavioral engagement, motivation, and self-efficacy in the context of education. However, there is still a gap in understanding the stability of this relationship in the Indonesian context and the mediating role of self-efficacy in the association between motivation and engagement in mathematics learning. To address this gap, our study aims to provide insights into the interplay among students' motivation, self-efficacy, and behavioral engagement during mathematics learning, focusing on primary education. Additionally, we will consider relevant demographic factors, such as age, gender, and grade, to better understand their potential impact on students' outcomes.

THEORETICAL FRAMEWORK

Motivation and Behavioral Engagement

Intrinsic motivation is a psychological process that stimulates an individual's interest in solving a specific problem or engaging in a particular activity based on their autonomy and competency (Latorre-Cosculluela et al., 2022). It signifies that an activity's performance is driven by self-satisfaction rather than external factors (Guay et al., 2010). When students possess high levels of intrinsic motivation, they are likely to learn mathematics with more autonomy. Researchers have observed that students become intrinsically motivated when they develop a fondness or positive sentiments towards an object (Cho and Perry, 2012). On the other hand, engagement refers to the degree of participation and involvement of students in educational practices (Durksen et al., 2017; Chiu, 2022). Literature review showed different types of academic engagement, such as emotional engagement, behavioral engagement, and cognitive engagement (Fredricks, Blumenfeld and Paris, 2004). Using different points of view, Skinner, Kindermann, and Furrer (2009) argued that engagement may consist of emotional engagement, behavioral engagement, and disaffected behavioral and emotional engagement. In the present study, we emphasized behavioral engagement, which many researchers agree on. Behavioral engagement refers to students' participation and involvement during academic activities (Skilling, Bobis and Martin, 2021).

Behavioral engagement plays a crucial role in students' academic outcomes in mathematics classrooms (Chiu, 2022). However, the factors that drive behavioral engagement are still being studied. Previous research has shown that behavioral engagement is closely linked to a student's motivation. When students feel motivated, they are more likely to participate actively in mathematics learning, such as solving problems, sharing their ideas, and participating in group discussions. Chiu (2022) suggested that when students feel more competent and autonomous, they become more engaged in learning. Durksen et al. (2017) reported that a student's motivation drives their behavioral engagement during mathematics learning, leading to an increase in interactions between students and teachers.

Similarly, Xia et al. (2022) found that the level of students' motivation was positively correlated with their engagement in mathematics learning. However, there is a lack of empirical investigation into this relationship for primary education in Indonesia. There is a shortage of information on the extent to which students' motivation towards mathematics and their behavioral engagement are associated, especially in the Indonesian context. Therefore, our study aimed to investigate students' motivation and behavior during mathematics learning in primary education in Indonesia. Based on previous research, we hypothesized a positive association between motivation and engagement in mathematics learning among primary education students.

The Role of Self-efficacy in Behavioral Engagements

Most of the previous study-defined self-efficacy definitions refer to Bandura's (1997) work, which described self-efficacy as individual judgments' of their performance to organize and execute a course of action on academic tasks to achieve success (Chao, McInerney and Bai, 2019). Students' self-efficacy are the individuals' convictions and expectations of what they can accomplish in certain situations (Xie, Yang and Xiao, 2022). There are a variety of constructs to measure personal beliefs depending on the purpose of the research.

The role of self-efficacy cannot be overstated in academic settings. These beliefs determine students' approach, perseverance, and effort when encountering obstacles (F. Pajares, 2003; Xie et al., 2022). Studies reveal that students with strong self-efficacy perform better and are more motivated (Usher and Pajares, 2009; Chong et al., 2018; Gao, 2020; Trautner and Schwinger, 2020). This is because their personal convictions enable them to maintain their efforts and persist in challenging situations. Additionally, students' level of self-confidence can even fuel their interest in mathematics (Jiang et al., 2022). When students believe in themselves, they are more likely to engage actively in class and put in the necessary effort and energy.

Previous studies (Archambault, Janosz, and Chouinard, 2012; Kareem, Thomas and Nandini, 2022; Orji and Ogbuanya, 2022) established a correlation between mathematical beliefs and engagement behavior. Metallidou and Vlachou (2007) have found that individuals with strong self-efficacy are better at regulating their cognitive engagement during mathematics learning. Additionally, Skaalvik et al. (2015) observed that self-efficacy was linked to students' intrinsic motivation and effort, which led to help-seeking behavior when faced with difficulties. Kareem et al. (2022) have investigated the relationship between self-efficacy, teacher and student attitudes, and engagement in mathematics learning. Similarly, Orji and Ogbuanya (2022) discovered that students with high self-efficacy were more involved in learning activities. However, there is a lack of research on the role of self-efficacy in students' engagement, especially in primary education. Therefore, our research aims to explore the relationship between self-efficacy and students' engagement in mathematics learning.

Self-efficacy as a Mediator Between Motivation and Behavioral Engagement

According to previous research studies, students' level of self-efficacy has a direct relationship with their motivation

(Chang et al., 2014) and engagement (Metallidou and Vlachou, 2007; Archambault, Janosz and Chouinard, 2012; Orji and Ogbuanya, 2022) during math learning. When students believe they possess a high proficiency in mathematics, they are more engaged in regulating their learning strategies (Metallidou and Vlachou, 2007). Correspondingly, Walker et al. (2006) suggested that self-efficacy predicts students' engagement. Y. Jiang and Zhang (2023) also found that self-efficacy positively correlates with behavioral engagement in mathematics.

Moreover, students' beliefs concerning their ability in math, which include their beliefs about problem-solving in math, self-regulated learning, and ability during math learning, are closely linked to their motivation (Walker, Greene and Mansell, 2006; Voica, Singer and Stan, 2020; Orji and Ogbuanya, 2022). For instance, Skaalvik et al. (2015) reported that students' beliefs about their ability in math are closely related to their intrinsic desire for math. When students derive pleasure from working with mathematics, their beliefs about their ability to solve problems in math increase. Orji and Ogbuanya (2022) also suggested that individuals with good motivation hold high self-efficacy. On the other hand, low motivation levels have been associated with low judgment levels of individuals' math learning ability (Habók et al., 2020). However, prior studies did not identify the mediation of self-efficacy for the relationship between motivation and behavioral engagement in math learning, particularly in primary education. Given that self-efficacy enhances students' engagement in math learning and is correlated with motivation, this study assumes that self-efficacy could mediate the relationship between attitude and intrinsic motivation during math learning.

Gender and Grade

In the literature reviews, demographic data such as gender and grade level were analyzed in relation to students' academic performance outcomes. Some studies found that gender and grade level could influence these outcomes (Hidayatullah and Csíkos, 2023b, 2023a). For example, Skaalvik and Skaalvik (2013) discovered that motivation and effort were correlated differently depending on gender and grade level. Additionally, Li (2019) observed that female students displayed more strategic self-regulated learning compared to male students. Similarly, Hidayatullah and Csíkos (2023c) determined that the association between cognitive and non-cognitive factors, including motivation, beliefs, and achievement, varied according to grade level. The authors noted that the path coefficient between motivation and achievement was stronger in sixth and fifth grades. Later, Hidayatullah and Csíkos (2023b) found that students differ in mathematical beliefs according to their gender and level of study. Accordingly, this study aims to investigate the consistency of the structural relationship among motivation, self-efficacy, and behavioral engagement during mathematics learning.

Present Study

This research focuses on investigating the mediating role of self-efficacy in the association between motivation and engagement in mathematics learning. As the researcher discussed earlier, although several previous studies have investigated the link between motivation and engagement (Metallidou and Vlachou, 2007; Kareem, Thomas and Nandini, 2022; Orji and Ogbuanya, 2022), the empirical research on the relationship between the two in mathematics learning, especially for primary education are scarce. In reality, students from developing countries like Indonesia suffer from mathematics scores (Chen et al., 2018) and scary of mathematics learning. Surabaya is one of the urban areas in Indonesia, and most of the students in this area come from middle-class families. The schools mostly have been supported by the internet and computers for education.

Furthermore, some empirical investigations also showed the connection between motivation and self-efficacy (Chang et al., 2014; Jiang et al., 2014; Skaalvik, Federici, and Klassen, 2015; Jiang and Zhang, 2023); the intermediation relations among the variables mentioned above have hardly been studied. Whether or not self-efficacy mediated the relationship between motivation and engagement during mathematics learning is still unexplored. Therefore, our study proposed the model association among the variables above. Our hypotheses below guide our investigation. Figure 1 depicts our model hypothesis:

1. Motivation is expected to be positively associated with behavioral engagement and self-efficacy toward mathematics.
2. Self-efficacy toward mathematics is expected to mediate the relationship between motivation and behavioral engagement.
3. The association between motivation, behavioral engagement, and self-efficacy differs according to grade and gender preferences.

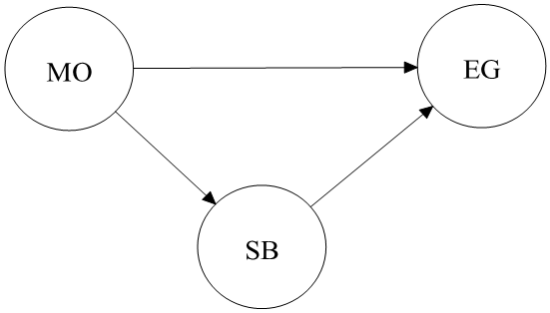


Figure 1: Hypothesis structural model of the relationship between attitude (ATM), self-efficacy (SFM), and intrinsic motivation (IMT) toward mathematics

METHODS

Participants

This cross-sectional study occurred in Surabaya, the capital city of East Java Province, Indonesia, in the first semester of 2022. Surabaya is an urban city in Indonesia, where students in the schools are from different socio-economic statuses. Six hundred sixty students (336 = boys, 324 = girls) from four schools participated in the present study. Three hundred and fifty students were from fifth grade, and three hundred and ten were from sixth grade. Students in these schools are of a variety of social and economic backgrounds. Participants ranged from 9-12 years old, with a mean age of 10.73 (SD = 0.70). Table 1 summarizes our demography's participants.

Characteristic	Full sample	Percentage
Gender		
Boys	336	50.9%
Girls	324	49.1%
Grade		
Fifth	350	53%
Sixth	310	47%
Age		
9 years	9	1.4%
10 years	244	37 %
11 years	317	48 %
12 years	90	13.6 %

Table 1: the sample structure

Instruments

In the present study, our study asked students to complete 28-item questionnaires for attitude, intrinsic motivation, self-efficacy about mathematics, and three items about their background (e.g., gender, grade, and age). All of these questionnaires were administered using the paper pencil-based test.

Motivation toward mathematics. Five items were adapted from the Academic Self-Regulation Questionnaire. The English version was adapted by Rubach and Bonanati (2021). This questionnaire related to the questions about students' enjoyment of mathematics lessons. For instance, "I learn mathematics because mathematics is fun for me" and "I learn mathematics because I want to know new things." The items of this questionnaire were rated using a five-point Likert scale (1 = strongly disagree to 5 = Strongly agree).

Self-efficacy. In the present study, we adapted six-item self-efficacy in mathematics from the academic efficacy scale (Dorman and Adams, 2004) teacher support, investigation, task orientation, cooperation, equity, involvement, personal relevance, shared control, student negotiation. This scale has been used to measure the extent to which students' beliefs about their capability in mathematics learning. For example: "I'm certain that I can master the skills taught in math this year " and "Even if the math is hard, I can learn it." These questionnaires were rated using a five-point Likert scale of 1-5 (1 = strongly disagree to 5 = strongly agree).

Behavioral Engagement. In the present study, five items of behavioral engagement were adapted from the students' engagement questionnaire (Kong, Wong and Lam, 2003). For instance, "I always take part in mathematics discussion in class" and "I concentrate when my teacher explains mathematics concepts." These items were rated using a five-point Likert scale (1 = strongly disagree to 5 = strongly agree). The validity of the behavioral engagement instrument was confirmed by performing a confirmatory factor analysis.

Data Analysis

In the present study, SPSS and Mplus 8 versions were performed to analyze the data. There are three steps to analyze the data. Two of them involved structural equation modeling analysis. In the first step, confirmatory factor analysis (CFA) was performed to confirm the validity of each questionnaire. Following this step, Cronbach's alpha was used to confirm the reliability of the questionnaire. In the second step, descriptive statistical analysis was used to explain the mean result of each variable as well as their correlations. In the third step, full structural equation modeling (SEM) analysis was performed to identify the association among variables and the mediation of self-efficacy for the relationship between motivation and engagement. Following this step, structural equation modeling was also performed to analyze the mediation of each type of self-efficacy about math (mathematics self-efficacy, problem-solving self-efficacy, and self-efficacy for self-regulated learning) for the relationship between motivation and engagement. During the model analysis, students' gender, age, and grade were included as variable controls.

Several parameters were used to gain the fit model. Maximum likelihood was used as a parameter estimate, and an absolute value loading factor of .40. Five good fit indices were used to measure the quality of the model: Tucker-Lewis's index (TLI) and comparative fit index (CFI), Chi-square, the Root Mean Square Error of Approximation (RMSEA), and Standardized Root Mean Squared Residual (SRMR). According to Hu & Bentler (1999), the value of CFI and TLI should be close to or greater than .90, SRMR value less than .08 (Hu and Bentler, 1999), and RMSEA value less than or equal to .10 (MacCallum, Browne and Sugawara, 1996; Hooper, Coughlan and Mullen, 2008). Chi-square statistics are reported (Hooper, Coughlan and Mullen, 2008) to assess a model fit. We evaluated the convergent data by calculating the average variance extracted (AVE). The coefficient value for AVE should be > 0.5. The discriminant validity was evaluated using the Forner Larcker criterion, where the root of

AVE should be greater than the correlation among latent variables constructs (Hair et al., 2019b).

Procedure

In the first step, the instruments were translated into Indonesian. Three experts and two mathematics teachers reviewed the items questionnaire before the researcher administered it. Each school principal was contacted, and a letter of permission was sent to the schools about the research. Then, 28 classes were randomly selected from six primary schools in Surabaya. Finally, 851 students participated in the present study. Data were gathered using paper-pencil tests.

RESULTS

Descriptive Statistic

Table 2 describes the descriptive statistics, internal reliability, and correlation of the latent variables. Most latent variables have good reliability according to the coefficient value of Cronbach alpha (range from 0.43-0.81) and internal reliability (range from.70-.84). Table 2 indicated that students have high motivation, high behavioral engagements, and moderate self-efficacy in mathematics learning according to the main result, ranging from 3.61-3.97 on a 5-point Likert scale. The skewness ranges from -.34 to -.65, and kurtosis ranges from.21 to.73, indicating the data was distributed normally (Kline, 2005).

Variables	M	SD	alpha	Skewness	Kurtosis
1. Motivation	3.96	.71	.84	-.58	.50
2. Self-efficacy	3.61	.74	.81	-.34	.21
3. Behavioral engagement	3.95	.67	.70	-.65	.73

Table 2: Descriptive statistics and correlation of each variable

Convergent Validity and Discriminant Validity

Confirmatory factor analysis was performed to confirm the construct validity of the questionnaire. With maximum likelihood parameter estimate, our instruments gained the good fit model, χ^2 (df = 100) = 385.29, p < .001, CFI = .93, TLI = .91, RMSEA = .06, SRMR = .04. Table 3 shows the convergent validity of the questionnaire. For the indicators reliability, loading factors > .70 are suggested and loading factors > .40 are acceptable (H_o, 2006; Kwong-Kay Wong, 2013). The result showed that the factors loading ranges from .35 to .81. The composite reliability (CR) for motivation, behavioral engagement, and self-efficacy was good, .84, .70, and .81, respectively. As suggested by Hair et al. (2019a), the composite reliability (CR) cut-off value should be higher than .70. Although the result shows that only the motivation construct obtained score of AVE was higher than .50, the score value of CR for self-efficacy and behavioral engagement was around .70, indicated that the indicators consistently measuring the latent variables. In other words, our items were internally consistent or convergent (Hair et al., 2019b).

Latent variables	Items	Loading factor	AVE	CR
Motivation			.52	.84
	M1	0.72		
	M2	0.78		
	M3	0.81		
	M4	0.69		
	M5	0.59		
Behavioral engagement			.33	.70
	EG1	0.68		
	EG2	0.68		
	EG3	0.43		
	EG4	0.51		
	EG5	0.52		
Self-Efficacy			.43	.81
	SE1	0.70		
	SE2	0.68		
	SE3	0.75		
	SE4	0.67		
	SE5	0.57		
	SE6	0.51		

Note. CR = composite reliability, AVE = average variance extracted

Table 3: Convergent validity and reliability of the construct

Furthermore, we evaluated the discriminant validity of the latent variables.by comparing the AVE score and the correlation between factors. Fornell and Larcker (1981) suggested that the root of AVE should be higher than the correlation between factors, indicating discriminant validity. The result showed that the correlation among the latent variables was weaker than the root AVE (see Table 4). Therefore, the structural equation modelling evaluation can be employed to evaluate the structural relationship among the aforementioned variables.

Variables	1	2	3
4. Motivation	(.72)		
5. Self-efficacy	.59**	(.66)	
6. Behavioral engagement	.70**	.53**	(.57)

*Note: *significant at the level.05 (p <.05), ** significant at the level 001 (p <.001). Diagonal data are the root of the average variance extracted.*

Table 4: Discriminant validity

SEM Evaluation

Our hypothesis proposed self-efficacy to mediate the relationship between attitude and intrinsic motivation. However, in the first step, we evaluated the normality data. We accept the criteria skewness +/- 3, and kurtosis +/ 8 can be considered as normal distribution (Kline, 2005). The Skewness and Kurtosis data indicated that our data was normal. Then, we performed structural equation modeling. Our model was not ideal, χ^2 (df = 101) = 436.53, p < .001, CFI = .91, TLI = .89, RMSEA = .07, SRMR = .05. We modified it to find the fit model (See figure 2) χ^2 (df = 100) = 385.29, p < .001, CFI = .93, TLI = .91, RMSEA = .06, SRMR = .04. The model suggested that motivation was positively associated with behavioral engagement (β = .79, p < .001) and self-efficacy about mathematics (β = .70, p < .001). Self-efficacy was positively related to behavioral engagement (β = .19, p < .001). In comparison, the indirect effect of motivation on behavioral engagement through self-efficacy was weaker than the direct effect (β = .13, p < .001).

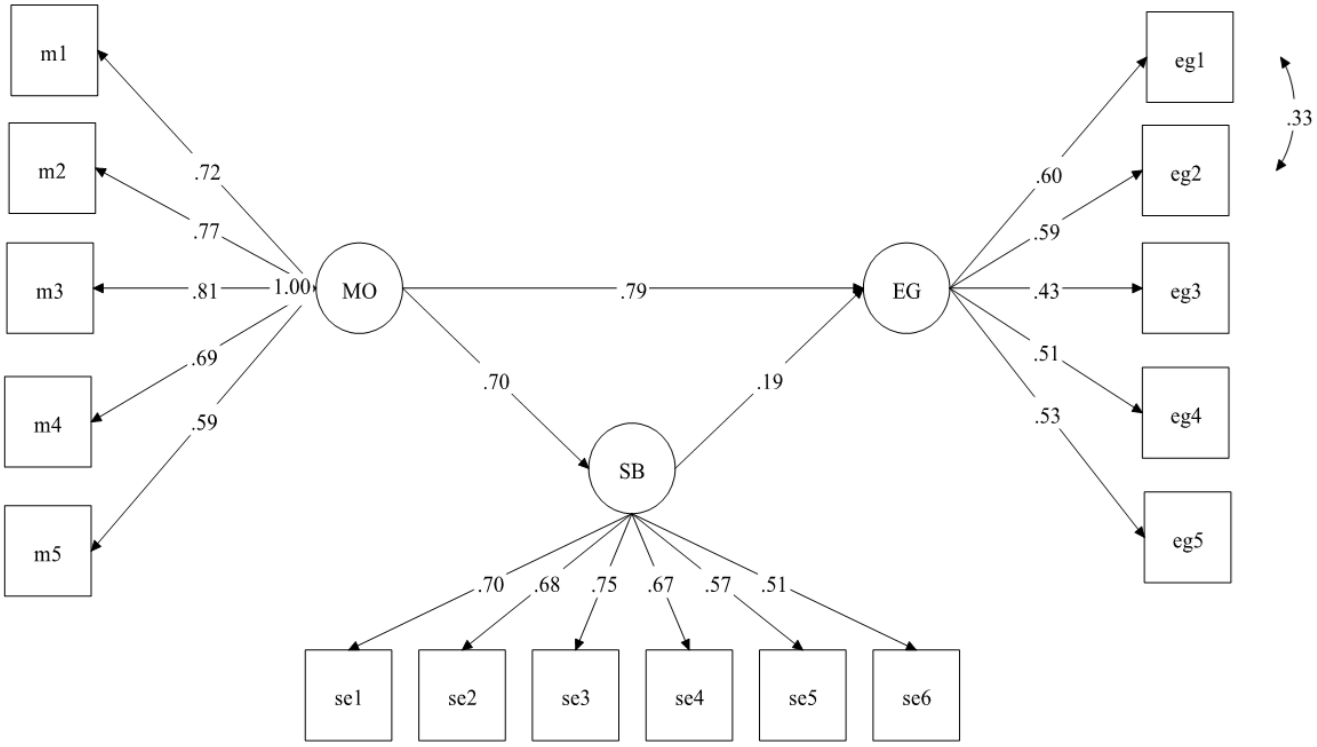


Figure 2: the mediation of self-efficacy (SB) with standardized coefficient

We further examined the association and the intermediation between latent variables by performing a bootstrapping approach. We calculated the direct association among these variables with a 95% confidence interval through bootstrapping 5,000 samples. The association between motivation and behavioral engagement was .79 (95% CI = [.66, .94]). Motivation was also positively associated with self-efficacy .75 (95% CI = [.54, .88]). Self-efficacy was directly associated with behavioral engagement .16 (95% CI = [.03, .28]). Self-efficacy positively mediated the relationship between motivation and behavioral engagement .13 (95% CI = [.03, .23]). Table 5 summarizes the bootstrapping result for the relationship among these variables.

Path	Standardized Estimate	Bootstrapping 95% CI	
		Lower Limit	Upper limit
Direct effect	.79	.66	.94
Motivation → Behavioral engagement	.70	.54	.88
Motivation → Self-efficacy	.75	.64	.87
Self-efficacy → Behavioral engagement	.16	.03	.28
Indirect effect			
Motivation → Self-efficacy → Behavioral engagement	.13	.03	.23

Table 5: Total, direct, and indirect effects

Furthermore, we performed the multigroup analysis to evaluate whether or not the association and the intermediation between these variables are stable across gender and grade-level studies. We found the fit model for the association among these variables in different in terms of grade level, χ^2 (df = 224) = 570.71, p < .001, CFI = .91, TLI = .91, RMSEA = .06, SRMR = .05. Motivation positively associated with behavioral engagement in grade 5 and grade 6, the path regression was (β = .74, p < .001) and (β = .64, p < .001), respectively. Self-efficacy is positively associated with behavioral engagement in grade 5 (β = .14, p = .02) and grade 6 (β = .17, p = .03). Self-efficacy has positively mediated the relationship between motivation and behavioral engagement in both grade 5 and grade 6, (β = .12, p = .02) and (β = .17, p = .03), respectively.

With respect to the association between motivation, behavioral engagement, and self-efficacy based on gender, the structural equation modeling consistently fit across gender χ^2 (df = 224) = 246.78, p < .001, CFI = .91, TLI = .91, RMSEA = .06, SRMR = .06. The differences between boys and girls in term of association between variables mentioned above have been identified. The direct association between motivation associated with behavioral engagement was stronger for boys (β = .87, p < .001) than for girls' students (β = .69, p < .001). In comparison, the direct association between self-efficacy and behavioral engagement was significant for girls (β = .27, p = .01) but not for boys (β = .12, p = .11). At the same time, the mediating role of self-efficacy in the relationship between motivation and behavioral engagement was only significant for girls β = .20, p = .01) but not for boys (β = .09, p = .10).

DISCUSSION

The main focus of this study was to investigate the structural model that describes the association between motivation, self-efficacy, and behavioral engagement in mathematics learning. Overall, the association and the mediation among the variables mentioned above were significant. This study contributed to the unpacking relationship among motivation, self-efficacy, and behavioral engagements for primary education in the Indonesian context.

Our study found that motivation was directly associated with self-efficacy and behavioral engagements. This finding supports hypothesis 1. What we found in the present study is consistent with the prior research (Chiu, 2022; Durksen et al., 2017), which suggested that the level of students' motivation would produce their involvement and active participation during mathematics learning. According to self-determination theory, students' motivation is the result of autonomy, relatedness, and competence support, which the teachers provide (Deci and Ryan, 2008; Chiu, 2022). When students became motivated, they were more willing to participate actively in mathematics learning. Our finding also showed that motivation was associated with self-efficacy in mathematics learning. This finding aligns with Skaalvik et al. (2015) and Hidayatullah and Csíkos (2023), who found a positive association between the two in mathematics learning. Orji and Ogbuanya (2022) argued that those with good motivation tend to have more confidence in judging their ability during mathematics learning. On the contrary, when students have less motivation, their beliefs about their capabilities decrease (Habók et al., 2020).

Our study also revealed that self-efficacy in mathematics learning positively mediated the association between motivation toward mathematics and behavioral engagement during mathematics learning. This finding supported the second hypothesis. Our finding is also consistent with Orji and Ogbuanya (2022) and Skaalvik et al. (2015), who suggested that beliefs in mathematics learning have been found to be the result of motivation. When students have good motivation, their beliefs also increase (Usher and Pajares, 2009; Chong et al., 2018; Gao, 2020; Trautner and Schwinger, 2020) and, in turn, promotes students' behavioral engagement during mathematics learning (Metallidou and Vlachou, 2007; Archambault, Janosz and Chouinard, 2012). According to social cognitive theory (Bandura, 2001), when students hold strong beliefs about their capability, they put much effort into being active and involved in academic learning. Therefore, prompt students' active participation during mathematics learning cannot be separated from motivation and self-efficacy.

Surprisingly, our study also revealed that the model relationship between motivation, self-efficacy, and behavioral engagement during mathematics learning was stable based on grade differences. This finding is in line with the previous studies (Hidayatullah & Csíkos, 2023; Skaalvik and Skaalvik, 2013), which find the weight of association among latent variables tends to vary based on the gender and grade level study. This study revealed the same result as the study by Suherman and Vidákovich (2024), which found a variety of weight associations between attitude and mathematics achievement in the SEM evaluation. In this study, the direct association of motivation on self-efficacy and behavioral engagement in mathematics learning was positive in grades 5 and grade 6. The direct association between motivation and behavioral engagement was stronger in grade five. At the same time, the mediation of self-efficacy for the relations between motivation and behavioral engagement was stronger in grade six. Our interpretation for this stage is that students' motivation towards mathematics in grade six decreased. Therefore, to promote their behavioral engagement, they also need the motivation to increase their self-efficacy in mathematics learning. However, a further empirical study is necessary to clarify this speculation for future research.

The data also showed that the model was stable for boys' and girls' students. The differences in the weight association between motivation, self-efficacy, and behavioral engagement have been identified. This study revealed the same result as Skaalvik and Skaalvik (2013), who found significant differences in the association between the effect and motivational aspects in primary education. In the current study, the direct association between self-efficacy and behavioral engagement was significant only for girls, not for boys. Consequently, the mediating role of self-efficacy for the relationship between motivation and engagement was significant only for girls. This finding contradicts Oppermann, Brunner and Anders (2019), who suggested that the association between engagement and self-efficacy was stronger for male students. It can be interpreted that girls might be more inclined to internalize their desire for mathematics and develop a stronger self-judgment about their

ability, generating a stronger association between motivation, self-efficacy, and behavioral engagement than boys' students. However, an additional study is needed to confirm this finding. Our study revealed that motivation and self-efficacy were directly associated with students' behavioral engagement during mathematics learning. However, the effect of motivation was stronger than self-efficacy on behavioral engagements. Self-efficacy also serves as a mediator for the relationship between motivation and behavioral engagements. In other words, when students become motivated to study mathematics, their confidence in their capability to overcome any obstacle in mathematics learning also increases, in turn promoting their involvement during mathematics learning.

LIMITATIONS AND FUTURE RESEARCH

Although this study provided a wealth of data and information, several limitations should be noted. First, the research was a cross-sectional study with the survey, which cannot be stated as a causal relationship between motivation, self-efficacy, and behavioral engagement. Second, this study only measures the mediating role of self-efficacy, and there may be a bi-direction of the relationship among the variables mentioned above. Therefore, a longitudinal study is needed for future research to confirm the causal relationship among these variables. Third, this study emphasized the structural model of non-cognitive factors without investigating the implication for students' achievement and cognitive engagement during mathematics learning. Fourth, in the present study, the data was gathered using self-report, which leaves much space to be improved. In future research, a deep interview is important to strengthen the results of the self-report. This study also investigated the stability of the model based on grade and gender differences. However, there was no conclusion to confirm the differences between the groups. Future research is necessary to clarify why there are differences in the structural relationship among variables mentioned above based on gender and grade differences in the form of a longitudinal study.

CONCLUSION AND IMPLICATION

To summarize, this study's finding showed that motivation and self-efficacy in mathematics learning positively predict students' behavioral engagement. Also, this study revealed that self-efficacy in mathematics learning mediates the relationship between motivation toward mathematics and behavioral engagement in mathematics learning. The findings of this study provide theoretical contributions. This study provided empirical evidence of the relationship between motivation, self-efficacy, and behavioral engagement. The more motivated students to study mathematics, their self-efficacy to study math increased, and they were more likely to participate actively in mathematics, such as always taking part in mathematics group discussions. However, the differences in the association between self-efficacy, motivation, and behavioral engagement by gender in mathematics lessons may encourage the discussion of the relevance of motivation and self-efficacy theory.

These findings have implications for teaching practices. Since the finding of this study told us that motivation and self-efficacy are key factors in promoting students' behavioral engagements, this finding enlightened us that mathematics educators should set the math class to shape students' positive motivation because the positive motivation will elevate self-judgment about the capability to discuss in group work, to solve the mathematical task, and to regulate mathematics learning, then facilitate students' behavioral engagements. Grouping students in mathematics group work may also promote their motivation and beliefs about their mathematics learning ability. Providing autonomy support and encouraging students to be more connected with others, such as their peers in the classroom, will strengthen their motivation toward mathematics. Persuading and appreciating students' work also increases students' motivation and beliefs to do well in mathematics learning.

DECLARATION OF COMPETING INTEREST

No conflict of interest exists.

REFERENCES

Archambault, I., Janosz, M. and Chouinard, R. (2012) 'Teacher beliefs as predictors of adolescents' cognitive engagement and achievement in mathematics', *Journal of Educational Research*, Vol. 105, No. 5, pp. 319–328. https://doi.org/10.1080/00220671.2011.629694

Bandura, A. (1997) *Self-Efficacy: The exercise of control*, New York:W. H. Freeman and Company.

Bandura, A. (2001) 'Social cognitive theory: An agentic perspective', *Annual Reviews Psychology*, Vol. 52, pp. 1–26. https://doi.org/10.1146/annurev.psych.52.1.1

Chang, C.-S., Liu, E. Z.-F., Sung, H.-Y., Lin, C.-H., Chen, N.-S. and Cheng, S.-S. (2014) 'Effects of online college student's Internet self-efficacy on learning motivation and performance', *Innovations in Education and Teaching International*, Vol. 51, No. 4, pp. 366–377. https://doi.org/10.1080/14703297.2013.771429

Chao, C. N. G., McInerney, D. M. and Bai, B. (2019) 'Self-efficacy and self-concept as predictors of language learning achievements in an Asian bilingual context', *Asia-Pacific Education Researcher*, Vol. 28, No. 2, pp. 139–147. https://doi.org/10.1007/s40299-018-0420-3

Chen, L.-K., Dorn, E., Krawitz, M., SH Lim, Ch. and Mourshed, M. (2018) *Drivers of Student Performance: Insights from Asia*. McKinsey & Company. Available: https://www.mckinsey.com/~/media/McKinsey/Industries/Social Sector/Our Insights/What drives student performance in latin america/Drivers-of-Student-Performance.ashx

Chiu, T. K. F. (2022) 'Applying the self-determination theory (SDT) to explain student engagement in online learning during the COVID-19 pandemic', *Journal of Research on Technology in Education*, Vol. 54 No.1, pp. 14–30. https://doi.org/10.1080/15391523.2021.1891998

Cho, Y. J. and Perry, J. L. (2012) 'Intrinsic motivation and employee attitudes: Role of managerial trustworthiness, goal-directedness, and extrinsic reward expectancy', *Review of Public Personnel Administration*, Vol. 32, No. 4, pp. 382–406. https://doi.org/10.1177/0734371X 11421495

Chong, W. H., Liem, G. A. D., Huan, V. S., Kit, P. L. and Ang, R. P. (2018) 'Student perceptions of self-efficacy and teacher support for learning in fostering youth competencies: Roles of affective and cognitive engagement', *Journal of Adolescence*, Vol. 68, No. 1, pp. 1–11. https://doi.org/10.1016/j.adolescence.2018.07.002

Damrongpanit, S. (2019) 'From modern teaching to mathematics achievement: The mediating role of mathematics attitude, achievement motivation, and self-efficacy', *European Journal of Educational Research*, Vol. 8, No. 3, pp. 713–727. https://doi.org/10.12973/eu-jer.8.3.713

Deci, E. L., & Ryan, R. M. (2008). *Self-determination theory: A macrotheory of human motivation, development, and health*. Canadian Psychology, 49(3), pp. 182–185. https://doi.org/10.1037/a0012801

Dorman, J. and Adams, J. (2004) 'Associations between students' perceptions of classroom environment and academic efficacy in Australian and British secondary schools', *Westminster Studies in Education*, Vol. 27, No. 1, pp. 69–85. https://doi.org/10.1080/0140672040270106

Durksen, T. L., Chu, M.-W., Ahmad, Z. F., Radil, A. I. and Daniels, L. M. (2016) 'Motivation in a MOOC: a probabilistic analysis of online learners' basic psychological needs', *Social Psychology of Education*, Vol. 19, No. 2, pp. 241–260. https://doi.org/ 10.1007/s11218-015-9331-9

Durksen, T. L., Way, J., Bobis, J., Anderson, J., Skilling, K. and Martin, A. J. (2017) 'Motivation and engagement in mathematics: a qualitative framework for teacher-student interactions', *Mathematics Education Research Journal*, Vol. 29, No. 2, pp. 163–181. https://doi.org/10.1007/s13394-017-0199-1

Flunger, B., Hollmann, L., Hornstra, L. and Murayama, K. (2022) 'It's more about a lesson than a domain: Lesson-specific autonomy support, motivation, and engagement in math and a second language', *Learning and Instruction*, Vol. 77, No. 1, p. 101500. https://doi.org/10.1016/j.learninstruc.2021.101500

Fornell, C. and Larcker, D. F. (1981) 'Evaluating structural equation models with unobservable variables and measurement error', *Journal of Marketing Research*, Vol. 18, No. 1, p. 39. https://doi.org/10.2307/3151312

Fredricks, J. A., Blumenfeld, P. C. and Paris, A. H. (2004) 'School engagement: Potential of the concept, state of the evidence', *Review of Educational Research*, Vol. 74. No. 1. pp. 59–109. https://doi.org/10.3102/00346543074001059

Gao, J. (2020) 'Sources of mathematics self-efficacy in Chinese students: A mixed-method study with Q-Sorting procedure', *International Journal of Science and Mathematics Education*, Vol. 18, No. 4, pp. 713–732. https://doi.org/10.1007/s10763-019-09984-1

Guay, F., Ratelle, C. F., Roy, A. and Litalien, D. (2010) 'Academic self-concept, autonomous academic motivation, and academic achievement: Mediating and additive effects', *Learning and Individual Differences*, Vol. 20, No. 6, pp. 644–653. https://doi.org/10.1016/j.lindif.2010.08.001

Habók, A., Magyar, A., Németh, M. B. and Csapó, B. (2020) 'Motivation and self-related beliefs as predictors of academic achievement in reading and mathematics: Structural equation models of longitudinal data', *International Journal of Educational Research*, Vol. 103, No. 2, pp. 1–12. https://doi.org/10.1016/j.ijer.2020.101634

Hair, J. F., Babin B., Anderson, R. E. and Black W. (2019a) *Multivariate data analysis, multivariate data analysis*. Cheriton House: UK: Cengage Learning, EMEA.

Hair, J. F., Risher, J. J., Sarstedt M. and Ringle C. M. (2019b) 'When to use and how to report the results of PLS-SEM abstract', *European Business Review*, Vol. 3, No. 1, pp. 3–24. https://doi.org/10.1108/EBR-11-2018-0203

Hidayatullah, A. and Csíkos, C. (2023a) 'Exploring students ' mathematical beliefs : gender, grade, and culture differences', Journal on Efficiency and Responsibility in Education and Science, Vol. 16, No. 3, pp. 186–195. https://doi.org/10.7160/eriesj.2023.160303

Hidayatullah, A. and Csíkos, C. (2023b) 'Students' responses to the realistic word problems and their mathematics-related beliefs in primary education', Pedagogika, Vol. 150, No. 2, pp. 21–37. https://doi.org/10.15823/p.2023.150.2

Hidayatullah, A. and Csíkos, C. (2023c) 'The role of students' beliefs, parents' educational level, and the mediating role of attitude and motivation in students' mathematics achievement', The Asia-Pacific Education Researcher. https://doi.org/10.1007/s40299-023-00724-2

Ho, R. (2006) *Handbook of univariate and multivariate data analysis and interpretation with SPSS*, Abingdon: Taylor & Francis.

Hooper, D., Coughlan, J. and Mullen, M. R. (2008) 'Structural equation modelling: Guidelines for determining model fit', *Electronic Journal of Business Research Methods*, Vol. 6, No. 1, pp. 53–60.

Hu, L. T. and Bentler, P. M. (1999) 'Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives', *Structural Equation Modeling*, Vol. 6, No. 1, pp. 1–55. https://doi.org/10.1080/10705519909540118

Jiang, H., Turnbull, D., Wang, X., Chugh, R., Dou, Y. and Chen, S. (2022) 'How do mathematics interest and self-efficacy influence coding interest and self-efficacy? A structural equation modeling analysis', *International Journal of Educational Research*, Vol. 115, No. 5, pp. 1–10. https://doi.org/10.1016/j.ijer.2022.102058

Jiang, Y., Song, J., Lee, M. and Bong, M. (2014) 'Self-efficacy and achievement goals as motivational links between perceived contexts and achievement', *Educational Psychology*, Vol. 34, No. 1, pp. 92–117. https://doi.org/10.1080/01443410.2013.863831

Jiang, Y. and Zhang, L. (2023) 'High school students' expectancy, value, and cost profiles and their relations with engagement and achievement in Math and English', *Learning and Individual Differences*, Vol. 101, No. 1, pp. 1–14. https://doi.org/10.1016/j.lindif.2022.102252

Józsa, G., Oo, T. Z., Amukune, S. and Józsa, K. (2022) 'Predictors of the intention of learning in higher education: Motivation, self-handicapping, executive function, parents' education, and school achievement', *Education Sciences*, Vol. 12, No. 12, p. 906. https://doi.org/10.3390/educsci12120906

Kareem, J., Thomas, R. S. and Nandini, V. S. (2022) 'A conceptual model of teaching efficacy and beliefs, teaching outcome expectancy, student technology use, student engagement, and 21st-century learning attitudes: A STEM education study', *Interdisciplinary Journal of Environmental and Science Education*, Vol. 18, No. 4, pp. 1–11. https://doi.org/10.21601/ijese/ 12025

Kline, R. B. (2005) *Principles and practice of structural equation modeling*. New York and London: The Guilford Press : New York and London.

Kong, Q. P., Wong, N. Y. and Lam, C. C. (2003) 'Student engagement in mathematics: Development of instrument and validation of construct', *Mathematics Education Research Journal*, Vol. 15, No. 1, pp. 4–21. https://doi.org/10.1007/BF03217366

Kwong-Kay Wong, K. (2013) 'Partial least squares structural equation modeling (PLS-SEM) techniques using SmartPLS', *Marketing Bulletin*, Vol. 24, No. 1, pp. 1–32.

Latorre-Cosculluela, C., Sierra-Sánchez, V., Rivera-Torres, P. and Liesa-Orús, M. (2022) 'Emotional well-being and social reinforcement as predictors of motivation and academic expectations', *International Journal of Educational Research*, Vol. 115, No. 5, pp. 1–12. https://doi.org/10.1016/j.ijer.2022.102043

Li, K. (2019) 'MOOC learners' demographics, self-regulated learning strategy, perceived learning and satisfaction: A structural equation modeling approach', *Computers and Education*, Vol. 132, No. 5, pp. 16–30. https://doi.org/10.1016/j.compedu.2019.01.003

MacCallum, R. C., Browne, M. W. and Sugawara, H. M. (1996) 'Power analysis and determination of sample size for covariance structure modeling', *Psychological Methods*, Vol. 1, No. 2, pp. 130–149. https://doi.org/10.1037/1082-989X.1.2.130

Metallidou, P. and Vlachou, A. (2007) 'Motivational beliefs, cognitive engagement, and achievement in language and mathematics in elementary school children', *International Journal of Psychology*, Vol. 42, No. 1, pp. 2–15. https://doi.org/10.1080/00207590500411179

Oppermann, E., Brunner, M. and Anders, Y. (2019) 'The interplay between preschool teachers' science self-efficacy beliefs, their teaching practices, and girls' and boys' early science motivation', *Learning and Individual Differences*, Vol. 70, No.1, pp. 86–99. https://doi.org/10.1016/j.lindif.2019.01.006

Orji, C. T. and Ogbuanya, T. C. (2022) 'Mediating roles of ability beliefs and intrinsic motivation in PBL and engagement in practical skills relations among electrical/electronic education undergraduate', *Innovations in Education and Teaching International*, Vol. 59, No. 3, pp. 326–336. https://doi.org/10.1080/14703297.2020.1813188

Pajares, F. (2003) 'Self-efficacy, motivation, and achievement in writing: A review of the literature', *Reading and Writing Quarterly*, Vol. 19, No. 2, pp. 139–158. https://doi.org/10.1080/10573560308222

Rubach, C. and Bonanati, S. (2021) 'The impact of parents' home- and school-based involvement on adolescents' intrinsic motivation and anxiety in math', *Psychology in the Schools*, Vol. 5, No. 6, pp. 1–21. https://doi.org/10.1002/pits.22577

Ryan, R. M. and Deci, E. L. (2000) 'Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being', *American Psychologist*, Vol. 55, No. 1, pp. 68–78. https://doi.org/10.1037/0003-066X.55.1.68

Skaalvik, E. M., Federici, R. A. and Klassen, R. M. (2015) 'Mathematics achievement and self-efficacy: Relations with motivation for mathematics', *International Journal of Educational Research*, Vol. 72, No. 4, pp. 129–136. https://doi.org/10.1016/j.ijer.2015.06.008

Skaalvik, E. M. and Skaalvik, S. (2013) 'School goal structure: Associations with students' perceptions of their teachers as emotionally supportive, academic self-concept, intrinsic motivation, effort, and help seeking behavioral', *International Journal of Educational Research*, Vol. 61, No, 5, pp. 5–14. https://doi.org/10.1016/j.ijer.2013.03.007

Skilling, K., Bobis, J. and Martin, A. J. (2021) 'The "ins and outs" of student engagement in mathematics: shifts in engagement factors among high and low achievers', *Mathematics Education Research Journal*, Vol. 33, No. 3, pp. 469–493. https://doi.org/10.1007/s13394-020-00313-2

Skinner, E. A., Kindermann, T. A. and Furrer, C. J. (2009) 'A motivational perspective on engagement and disaffection', *Educational and Psychological Measurement*, Vol. 69, No. 3, pp. 493–525. https://doi.org/10.1177/0013164408323233

Suherman, S. and Vidákovich, T. (2024) 'Mathematical Creative Thinking-Ethnomathematics based Test: Role of Attitude toward Mathematics, Creative Style, Ethnic Identity, and Parents' Educational Level', *Revista de Educación a Distancia*, Vol. 24, No.77, pp. 1–22. https://doi.org/10.6018/red.581221

Trautner, M. and Schwinger, M. (2020) 'Integrating the concepts self-efficacy and motivation regulation: How do self-efficacy for motivation regulation influence self-regulatory success?', *Learning and Individual Differences*, Vol. 80, No. 4. https://doi.org/ 10.1016/j.lindif.2020.101890

Usher, E. L. and Pajares, F. (2009) 'Sources of self-efficacy in mathematics: A validation study', *Contemporary Educational Psychology*, Vol. 34, No. 1, pp. 89–101. https://doi.org/ 10.1016/j.cedpsych.2008.09.002

Voica, C., Singer, F. M. and Stan, E. (2020) 'How are motivation and self-efficacy interacting in problem-solving and problem-posing?', *Educational Studies in Mathematics*, Vol. 105, No. 3, pp. 487–517. https://doi.org/10.1007/s10649-020-10005-0

Walker, C. O., Greene, B. A. and Mansell, R. A. (2006) 'Identification with academics, intrinsic/extrinsic motivation, and self-efficacy as predictors of cognitive engagement', *Learning and Individual Differences*, Vol. 16, No. 1, pp. 1–12. https://doi.org/ 10.1016/j.lindif.2005.06.004

Welesilassie, M. W. and Nikolov, M. (2022) 'Relationships between motivation and anxiety in adult EFL learners at an Ethiopian university', *Ampersand*, Vol. 9. No. 1, pp. 1–10. https://doi.org/10.1016/j.amper.2022.100089

Xia, Q., Yin, H., Hu, R., Li, X. and Shang, J. (2022) 'Motivation, engagement, and mathematics achievement: An exploratory study among Chinese primary students', *SAGE Open*, Vol. 12, No. 4, pp. 1–13. https://doi.org/ 10.1177/21582440221134609

Xie, F., Yang, Y. and Xiao, C. (2022) 'Gender-math stereotypes and mathematical performance: the role of attitude toward mathematics and math self-concept', *European Journal of Psychology of Education*, Vol. 39, pp. 695–708. https://doi.org/10.1007/s10212-022-00631-y

Full research paper

NAVIGATING THE COLLEGE STUDENTS' ADVERSITIES: THE ROLE OF ACADEMIC BUOYANCY AND MOTIVATION ON LEARNING ACHIEVEMENT

Mohamad Arief Rafsanjani[1]✉
Handri Dian Wahyudi[2]
Retno Mustika Dewi[1]
Putri Ulfa Kamalia[1]

[1]Universitas Negeri Surabaya, Indonesia

[2]Universitas Negeri Malang, Indonesia

✉ mohamadrafsanjani@unesa.ac.id

ABSTRACT

Understanding the predictor of learning achievement among college students is crucial to adopting the appropriate learning strategy. Academic buoyancy is one of the predictors of learning achievement, playing a vital role in helping students navigate academic setbacks and adversities. However, the previous studies failed to reveal a robust link between the two variables. Therefore, this study aims to explain the nexus between academic buoyancy and students' learning achievement more clearly by introducing the mediating variable, motivational constructs, which include self-efficacy, persistence, and anxiety. This study involved 493 college students in Indonesia. Structural equation modeling (SEM) was utilized to examine the research hypotheses. The results show that academic buoyancy directly affects learning achievement. Furthermore, motivational constructs (self-efficacy, persistence, and anxiety) significantly mediate the relationship between academic buoyancy and learning achievement. This study contributes to the literature by explaining how academic buoyancy affects learning achievement through motivational constructs as a mediating variable. Furthermore, the university must promote students' academic buoyancy and motivational constructs by providing counseling services and encouraging students to recognize and address the adversities during lecture activities.

KEYWORDS

Academic buoyancy, learning achievement, self-efficacy, persistence, anxiety

HOW TO CITE

Rafsanjani M. A., Wahyudi H. D., Dewi R. M., Kamalia P. U. (2024) 'Navigating the College Students' Adversities: The Role of Academic Buoyancy and Motivation on Learning Achievement', *Journal on Efficiency and Responsibility in Education and Science*, vol. 17, no. 3, pp. 247–256. http://dx.doi.org/10.7160/eriesj.2024.170307

Article history
Received
January 9, 2024
Received in revised form
August 10, 2024
Accepted
September 2, 2024
Available on-line
September 30, 2024

Highlights

- *Academic buoyancy plays a pivotal role in navigating setbacks and adversities among college students.*
- *Academic buoyancy positively predicts students' learning achievement.*
- *Motivational constructs (self-efficacy, persistence, and anxiety) comprehensively elucidate the role of academic buoyancy on learning achievement.*
- *The university must take into account the promotion of students' academic buoyancy and motivational constructs by providing counseling services.*

INTRODUCTION

Learning achievement is seen as an important indicator of success in a wide range of areas, such as job satisfaction after employment (Choi, 2018), job performance (Dogaru and Popescu, 2021; Dyer, 1987; Wise, 1975), and career maturity (Bae, 2017; Bae, 2022). Accordingly, most scholars exploring the key predictor of learning achievement, e.g., the teaching and learning strategy (Grønlien et al., 2021; Tong et al., 2022; Yiin and Chern, 2023), teaching quality (Alp Christ et al., 2022; Sanfo and Malgoubri, 2023), cognitive ability (Cadoret et al., 2018; Demetriou et al., 2019; Pluck et al., 2020), prior knowledge (Bosch et al., 2021; Köller, 2012; Schneider and Preckel, 2017), and socio-economic status (Hopfenbeck et al., 2018; Marks and O'Connell, 2023). Besides, social-emotional development is another factor beyond the facia prima that affects learning achievement (Corcoran et al., 2018; Durlak et al., 2011; Liu et al., 2022).

Social-emotional development is becoming popular due to its role in helping students navigate challenges during daily school activities (Durlak et al., 2011; Martin and Marsh, 2020).

Academic buoyancy is part of social-emotional development. Academic buoyancy refers to the student's ability to successfully deal with minor adversities and setbacks during everyday school activities, such as receiving poor grades, negative feedback, exam pressure, competing deadlines, and declining motivation (Martin and Marsh, 2020; Putwain and Wood, 2023). Academic buoyancy helps students as the front-line protection against minor academic adversities. Furthermore, academic buoyancy prevents minor adversities from becoming major, such as school refusal and chronic underachievement (Putwain et al., 2020; Putwain and Wood, 2023). Therefore, academic buoyancy is an important factor affecting students learning achievement.

The previous studies revealed that students' ability to be buoyant in facing academic adversities and setbacks affects learning achievement directly (Granziera et al., 2022; Yu et al., 2019). Although the relationship between academic buoyancy and students' learning achievement is significant, the effect sizes were relatively small (rs = 0.07 to 0.19) (Datu and Yang, 2021; Fong and Kim, 2021; Lei et al., 2022) and failed to reveal a robust link between the two variables. At the same time, academic buoyancy is closely related to motivational and emotional outcomes (Yu et al., 2019). Accordingly, the current study includes the motivational and emotional aspects as a mediating variable to determine the possible linking mechanism between academic buoyancy and learning achievement. More precisely, we examined the relationship between academic buoyancy and learning achievement via an indirect mechanism. Therefore, the current study tries to comprehensively advance the understanding of how academic buoyancy affects students' learning achievement.

This study provides three main contributions. First, this study tries to explain more comprehensively how academic buoyancy affects students' learning achievement by including motivational and emotional aspects as mediating variables in the research model. Second, the study of academic buoyancy has been highlighted in various countries such as Australia (Bostwick et al., 2022), the United Kingdom (Putwain and Wood, 2023), Finland (af Ursin et al., 2021; Hirvonen et al., 2020), USA (Fong and Kim, 2021), and China (Yu et al., 2019). However, the study of academic buoyancy among Indonesian college students has received less attention from scholars. Therefore, the current study provides Indonesian university administrators and lecturers a basic understanding of the other predictors of learning achievement beyond the popular predictor (e.g., teaching and learning strategy, cognitive ability, and prior knowledge). Last, the current study will contribute to the existing body of literature, especially on how academic buoyancy affects learning achievement among college students.

THEORETICAL FRAMEWORK

Academic Buoyancy and Learning Achievement

Academic buoyancy is conceptualized as students' ability to successfully deal with minor adversities and setbacks during everyday school activities, such as receiving poor grades or negative feedback, exam pressure, competing deadlines, and declining motivation (Martin and Marsh, 2020; Putwain and Wood, 2023). Academic buoyancy also refers to the student's ability to show adaptive responses when facing challenges and adversity posed by routine school activities (Martin and Marsh, 2020). Students with adaptive responses to academic challenges and adversities are expected to achieve high academic achievement. Furthermore, academic buoyancy also plays a buffer role against the negative impact of adversities and setbacks during everyday school activities (Martin and Marsh, 2020; Putwain et al., 2020). Therefore, students with high academic buoyancy would be expected to mitigate the adverse effects of academic adversities on learning achievement.

The empirical study demonstrates a clear link between academic buoyancy and learning achievement. Students with higher academic buoyancy tend to achieve better results in examinations, particularly in subjects like English, science, math, numeracy, and literacy tests (Putwain et al., 2020; Putwain and Wood, 2023). Another study also shows that higher academic buoyancy is related to higher learning achievement (Datu and Yang, 2021; Yun et al., 2018).

Furthermore, the previous study also explains that academic buoyancy involves coping strategies that help students manage academic stress, lower their anxiety levels, and foster academic performance (Hirvonen et al., 2020). The study by Putwain et al. (2020) indicates that students with high academic buoyancy are adept at planning and prioritizing tasks and frequently engage in reflective practices by evaluating their learning process. These practices help them stay organized and focused, allowing them to identify their strengths and weaknesses, contributing to better academic performance.

The Mediating Role of Motivational Aspects

Although empirical studies show the direct effect of academic buoyancy on learning achievement, the effect sizes were relatively small (rs = 0.07 to 0.19) (Putwain and Wood, 2023). Furthermore, the other study shows no relationship after controlling the beliefs construct (Collie et al., 2015; Putwain and Aveyard, 2018). This finding indicates the presence of other factors that mediate the link between academic buoyancy and learning achievement.

A previous study found that academic buoyancy is associated with psychological aspects, such as motivation, anxiety, boredom, hopelessness, and shame (Datu and Yang, 2021; Hirvonen et al., 2020). Furthermore, academic buoyancy positively affects motivational and emotional factors (Collie et al., 2015). Academic buoyancy is also related to the process linked with students' learning achievement, such as motivational constructs in the form of persistence and self-efficacy (Collie et al., 2015). Additionally, anxiety was identified as another factor that explains the link between academic buoyancy and learning achievement (Collie et al., 2015). Accordingly, we propose that motivational aspects, such as self-efficacy and persistence, mediate the relationship between academic buoyancy and learning achievement.

The motivational construct is conceptualized as what initiates, sustains, and helps individuals complete tasks (Yu et al., 2019). Furthermore, motivational constructs are identified as having three core components: expectancy, value, and affective. Expectancy components revolve around students' belief in their ability to accomplish their goals, called self-efficacy. As Bandura (1997) explains, self-efficacy refers to an individual's confidence in their capability to perform a particular behavior in a given situation. The value components include how individuals value a task and within that, whereby persistence can be considered.

Last, affective components deal with the feelings or emotions that individuals experience before and during a particular task. Anxiety becomes the most extensive and significant focus in the theoretical and empirical analysis. Therefore, considering the three core components of the motivational construct, we use self-efficacy, persistence, and anxiety as mediating variables that reflect the motivational set in the current research model.

Objectives and Hypothesis

This study examines the mediating role of motivation in the link between academic buoyancy and learning achievement to comprehensively advance the understanding of how academic buoyancy affects students' learning achievement. Accordingly, this study examined the relationship between academic buoyancy and learning achievement via an indirect mechanism through motivational constructs. Furthermore, based on the literature review, motivational constructs consist of three core components: expectancy, value, and affective. Self-efficacy, persistence, and anxiety represent each core of motivational constructs.

According to the theoretical framework in the previous section, we draw the conceptual research model (figure 1) and present the research hypotheses as follows:

H1. Academic buoyancy affects learning achievement.

H2. Self-efficacy mediates the link between academic buoyancy and learning achievement.

H3. Persistence mediates the link between academic buoyancy and learning achievement.

H4. Anxiety mediates the link between academic buoyancy and learning achievement.

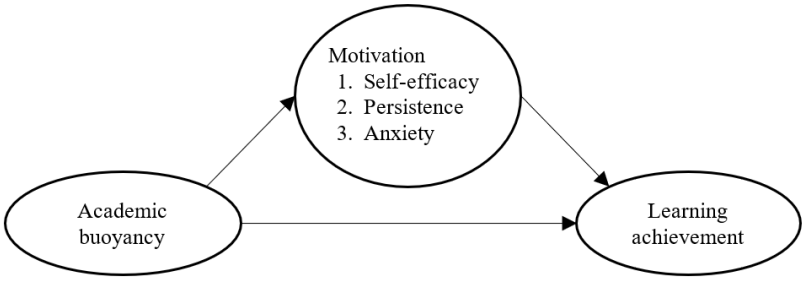


Figure 1: Conceptual research model

MATERIALS AND METHODS

Procedure and Participants

This study was conducted on university students in Indonesia from July to September 2023. To select research participants, we employed the cluster random sampling method. We used the five main islands in Indonesia as the clusters, namely Java, Sulawesi, Kalimantan, Sumatra, and Papua. Students were invited to participate via email and through student association networks. We sent invitations to 1,100 college students across 40 universities within these five clusters. The email contains the research objective, significance, and questionnaire link. Five hundred thirty-two (532) college students participated in the survey (response rate was 48.36%). After cleaning the data, we dropped 39 samples due to missing values in the data set. Therefore, this study used 493 participants as research participants. The detailed characteristics of the research participants are presented in Table 1.

Characteristics		Public University		Private University		Total Participants
		Participants	GPA's on average	Participants	GPA's on average	
Gender	Female	128	3.32	173	3.31	301
	Male	76	3.29	116	3.27	192
Region of the university	Java	82	3.48	113	3.42	195
	Sulawesi	35	3.16	47	3.23	82
	Kalimantan	31	3.04	42	3.15	73
	Sumatera	49	3.31	68	3.27	117
	Papua	11	2.99	15	3.02	26
Discipline	Economics, management & accounting	63	3.39	85	3.44	148
	Engineering	27	3.07	35	3.14	62
	Education	34	3.48	45	3.42	79
	Arts and humanities	32	3.36	43	3.29	75
	Law	29	3.31	26	3.25	55
	Others	40	3.27	34	3.32	74
Academic standing	1st-year student	71	3.06	97	3.14	168
	2nd-year student	65	3.51	88	3.48	153
	3rd-year student	53	3.32	72	3.28	125
	4th-year student	20	3.29	27	3.24	47

Table 1: Participants' characteristics (*N = 493)*

Instruments

We adopt the previous instruments to measure the research variable. We also adjusted the item to fit the research and participant context: college students. We measured academic buoyancy through four items from the Academic Buoyancy Scale (Martin and Marsh, 2020). For motivational constructs (self-efficacy, persistence, anxiety), we adopt the instruments from the Motivational and Engagement Scale: University/College (Martin, 2008). Each construct of motivation consists of four items. Last, we use the grade point average (GPA) to measure learning achievement.

Data Analysis

We used PLS-SEM to examine the research hypotheses using the SmartPLS 3.0 software package. We follow the multistage process by Hair et al. (2014). This process includes model specification, outer model evaluation, and inner model evaluation. In more detail, we propose the conceptual research model (Figure 1) into model specification based on the theoretical framework. We evaluate the outer model through confirmatory factor analysis (CFA), which involves the validity and reliability of each construct in the research model. Last, we evaluate the inner model through the coefficient of determination (R^2), cross-validated redundancy (Q^2), and path coefficients.

RESULTS

Outer Model Evaluation

In this section, we conduct the outer model evaluation through convergent validity, discriminant validity, and composite reliability. Convergent validity is established when the loading factor score of the measurement item is higher than 0.70 and the average variance extracted (*AVE*) score is higher than 0.50 (Hair et al., 2020). The result in Table 2 shows the loading factor of all measurement items involved in the current study had scores ranging from 0.804 to 0.962, which exceeds the threshold of 0.70. Furthermore, the *AVE* score of all constructs was higher than the threshold of 0.50, which confirms the convergent validity.

For the discriminant validity, we used the Fornell-Larcker criterion, which proposed that the squared root of the AVE score of each construct should be greater than that of the other constructs (Fornell and Larcker, 1981). The result in Table 3 indicates that each construct had a higher square root of AVE score than the others, which means the discriminant validity of the current model was established.

Later, we estimated the composite reliability using the composite reliability (*CR*). The result in Table 2 shows that the Cronbach's alpha value of all constructs ranged from 0.911 to 0.946. In contrast, the composite reliability value ranged from 0.938 to 0.961, which means the outer model has internal consistency and reliability (Hair et al., 2020).

Constructs	Item	Loading factor	AVE	Cronbach's alpha	Composite reliability
Academic buoyancy	AB1	0.919	0.845	0.939	0.956
	AB2	0.883			
	AB3	0.941			
	AB4	0.932			
Self-efficacy	SE1	0.962	0.815	0.924	0.946
	SE2	0.878			
	SE3	0.938			
	SE4	0.827			
Persistence	Pers1	0.903	0.791	0.911	0.938
	Pers2	0.934			
	Pers3	0.911			
	Pers4	0.804			
Anxiety	Anxy1	0.953	0.861	0.946	0.961
	Anxy2	0.910			
	Anxy3	0.925			
	Anxy4	0.923			
Learning achievement	GPA	1.000	1.000	1.000	1.000

Table 2: Loading factor, AVE, and composite reliability

	Academic buoyancy	Self-efficacy	Persistence	Anxiety	Learning achievement
Academic buoyancy	0.919				
Self-efficacy	0.782	0.903			
Persistence	0.734	0.766	0.889		
Anxiety	-0.654	-0.607	-0.542	0.928	
Learning achievement	0.555	0.535	0.546	-0.471	1.000

Table 3: Discriminant validity (Fornell and Larcker, 1981)

Inner Model Evaluation

The second step of multistage analysis is inner model evaluation. We evaluate the inner model through coefficient determination (R^2), cross-validated redundancy (Q^2), and path coefficients. The coefficient determination (R^2) was to estimate the model's predictive accuracy. The result in Table 4 shows the R^2 values of self-efficacy (SE), persistence (Pers), and anxiety (Anxy) were 0.628, 0.545, and 0.606, respectively. This result indicates that more than 50 percent variance of the three motivational constructs (SE, Pers, and Anxy) can be explained by academic buoyancy (AB). Furthermore, the R^2 value of learning achievement (LA) was 0.501, which means 50.1 percent of the LA variance can be explained by AB, SE, Pers, and Anxy variables at a robust level.

Furthermore, we also performed Stone-Geisser's Q^2 (cross-validated redundancy) to examine the model's predictive accuracy. Table 4 shows all the consequent variables (SE, Pers, Anxy, and LA) have a value of Q^2 greater than zero, indicating the current research model has a predictive relevance (Hair et al., 2020; Sarstedt et al., 2019).

Relationship	Adjusted R^2	Q^2
AB → SE	0.628	0.481
AB → Pers	0.545	0.243
AB → Anxy	0.606	0.473
AB, SE, Pers, and Anxy → LA	0.501	0.307

Note: AB = Academic buoyancy; SE = Self-efficacy; Pers = Persistence; Anxy = Anxiety; LA = Learning achievement

Table 4: Coefficient determination (R^2) and cross-validated redundancy (Q^2)

The path coefficients were examined to evaluate the research hypotheses through the significance of the structural relationship among the variables. The evaluation performed a bootstrap resampling method with 5000 iterations at a *p-value* of 0.05 significant level. The results are presented in Table 5 and Figure 2. The first result confirms the first hypothesis that academic buoyancy significantly affects learning achievement (β = 0.106 and *p-value* = 0.024). The result indicates that academic buoyancy is crucial in promoting students' learning achievement.

The next result shows that self-efficacy significantly mediates the relationship between academic buoyancy and learning achievement (β = 0.175 and *p-value* = <0.001), indicating self-efficacy explains how academic buoyancy affects learning achievement. Academic buoyancy will promote students' self-efficacy, which in turn enhances their academic performance (learning achievement). Next, the output shows persistence also plays a significant mediator in the relationship between academic buoyancy and learning achievement (β = 0.245 and *p-value* = <0.001), remarking that persistence also explains how academic buoyancy influences learning achievement. Good academic buoyancy will shape strong persistence, leading to better academic performance. The last finding indicates that anxiety also significantly mediates the effect of academic buoyancy on learning achievement (β = 0.315, *p-value* < 0.001). This suggests that academic buoyancy can reduce anxiety, which, in turn, can enhance learning achievement.

Hypotheses	Relationship	β-value	SE	*t*-value	*p*-value	Remarks
H1	AB → LA	0.106	0.026	2.682	0.024	Supported
H2	AB → SE → LA	0.175	0.019	5.086	<0.001	Supported
H3	AB → Pers → LA	0.245	0.013	5.924	<0.001	Supported
H4	AB → Anxy → LA	0.315	0.028	5.902	<0.001	Supported

Note: AB = Academic buoyancy; SE = Self-efficacy; Pers = Persistence; Anxy = Anxiety; LA = Learning achievement

Table 5: Summary of the tested hypotheses

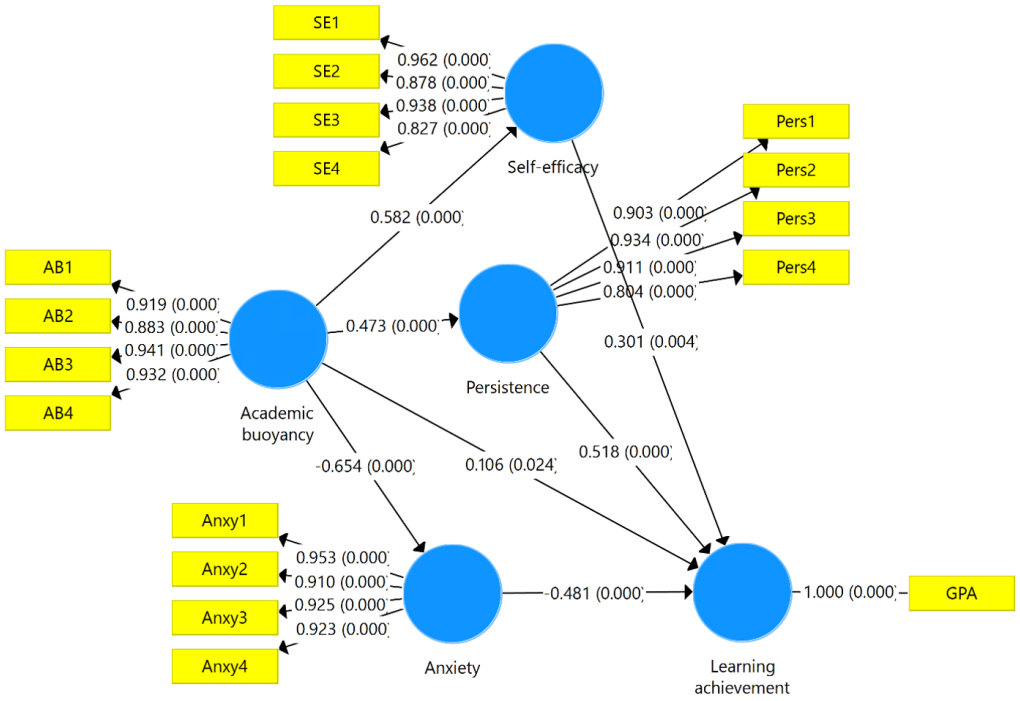


Figure 2: Result of structural model analysis

DISCUSSION

Academic buoyancy and learning achievement

The result of examining H1 shows that academic buoyancy has a significant positive effect on learning achievement. This finding indicates that students who successfully navigated academic adversities and setbacks during daily school had more opportunities for higher learning. The link between academic buoyancy and learning achievement can be explained for the following reasons. Students with high academic buoyancy show higher adaptive capability in response to minor challenges and adversities during daily lectures than those with low academic buoyancy. Students with higher adaptive ability have a higher chance of mitigating the adverse effects of academic adversities. Therefore, students with high academic buoyancy are more likely to achieve higher learning outcomes because they are better equipped to handle and overcome academic challenges.

Furthermore, students with good academic buoyancy can adjust their behavior, emotions, expectations, and assumptions when facing unexpected circumstances and situations, such as receiving negative feedback or a lower grade on a piece of work, exam pressures, and competing deadlines. These abilities make students more successful in dealing with the challenges and adversities during daily lectures, which then leads to positive learning achievement.

This study confirmed that academic buoyancy buffers against the negative impact of adversities and setbacks during everyday school activities (Martin and Marsh, 2020; Putwain et al., 2020). It also confirms previous findings that students with higher academic buoyancy tend to achieve higher learning achievement (Collie et al., 2015; Datu and Yang, 2021; Granziera et al., 2022; Lei et al., 2022; Yun et al., 2018).

The mediating role of motivational constructs

Although we confirmed the significant effect of academic buoyancy on learning achievement, the path coefficient is small (β-*value* = 0.106). We consider this result as a finding to strengthen the proposed hypothesis, the presence of other factors that mediate the link between academic buoyancy and learning achievement. As explained in the theoretical framework, we present motivational constructs (self-efficacy, persistence, and anxiety) as mediating variables in the research model to explain how academic buoyancy affects learning achievement (H2 – H4). The result of examining the mediating variable (H2 – H4) shows that all motivational constructs (self-efficacy, persistence, and anxiety) significantly mediate the relationship between academic buoyancy and learning achievement (*p-value* < 0.05). The detailed mechanism of how each construct plays a mediating variable will be explained as follows.

The first mediating variable is self-efficacy. The current study shows that self-efficacy significantly mediates academic buoyancy and learning achievement. This study shows that students with high academic buoyancy are confident in managing and dealing with academic adversities. The students who reported high academic buoyancy and self-efficacy were less likely to experience college work difficulties, leading to high learning achievement. This situation is different if compared to the students with low academic buoyancy. They demonstrated lower self-efficacy in dealing with daily lecture adversities, such as exam pressures or tight deadlines, and were likelier to have lower learning achievement. Furthermore, students with high self-efficacy respond to academic adversities in a positive and proactive manner. Therefore, we argue that self-efficacy is a factor that can explain how the mechanism of academic buoyancy affects learning achievement. This finding is in line with the previous study that proposed self-efficacy as

a factor that linked academic buoyancy and students' learning achievement (Caprara et al., 2011; Collie et al., 2015; Datu and Yang, 2021; Haidari et al., 2023). Additionally, the results are consistent with earlier research indicating that students with high academic buoyancy strongly believe in their ability to manage academic setbacks, which positively influences their learning performance (Yun et al., 2018).

The second mediating variable is persistence. This study evidences that persistence is another variable that significantly mediates the relationship between academic buoyancy and learning achievement. This study found that more buoyant students demonstrate better persistence and effort in overcoming daily lecture challenges. Furthermore, students are more likely to achieve better learning achievement with better persistence and effort. In contrast, students with lower academic buoyancy exhibit less persistence and effort in overcoming academic challenges, which can lead to lower academic achievement compared to their more buoyant peers. This finding aligns with the previous study that revealed persistence is a variable that links academic buoyancy and learning achievement (Collie et al., 2015; Granziera et al., 2022; Senler, 2022). Therefore, we can conclude that persistence is a factor that can explain how academic buoyancy affects learning achievement.

The last mediating variable is anxiety. As shown in the summary of the tested hypotheses (table 5), anxiety significantly mediates the relationship between academic buoyancy and learning achievement. This study revealed that students with high buoyancy have positive control of emotions over learning difficulties and adversities. Furthermore, due to the ability to control emotions positively, they are more likely to focus on the probability of success instead of failure. This ability will reduce anxiety when facing academic adversities, such as tight deadlines, negative feedback, and exam pressure. Because of the positive emotion approach instead of negative in facing adversities, students with higher buoyancy are more likely to experience lower anxiety, thus leading to better learning achievement. This study is in line with the previous research that revealed buoyancy contributes to helping students to control negative emotions (e.g., anxiety) (af Ursin et al., 2021; Collie et al., 2015; Hirvonen et al., 2020) and promote positive emotions (e.g., enjoyment) (Datu and Yang, 2021; Jia and Cheng, 2022; Wang and Hui, 2024) in facing adversities during daily schoolwork. The findings also align with prior research demonstrating that students with greater academic buoyancy exhibit better emotional responses to learning situations, which can enhance learning achievement (Collie et al., 2015; Granziera et al., 2022). Accordingly, we conclude that anxiety is another variable that explains the mechanism of academic buoyancy affecting learning achievement.

According to the result, this study provides theoretical and practical implications. Theoretically, this study contributes to the body of literature by explaining more clearly how academic buoyancy affects learning achievement through mediating variables (motivational constructs). The previous studies failed to evidence a robust link between the two variables due to small effect sizes. This study succeeds in revealing the link through the mediation model. Therefore, this study provides evidence that motivational constructs, including self-efficacy, persistence, and anxiety, are variables that can more clearly explain the relationship between academic buoyancy and learning achievement.

This study provides several practical points. First, lecturers are encouraged to help students recognize the adversities and challenges associated with routine lecture activities. By recognizing the adversities and challenges during lecture activities, students can be better prepared and develop effective strategies to successfully deal with academic adversities, such as pressure situations (tests and tight deadlines). Second, it is beneficial for students to learn adaptive responses to academic adversities. Furthermore, developing confidence and positive emotional control can help students manage difficulties more effectively. Students can reduce anxiety by developing confidence and positive emotional control abilities, such as fear of failure in tests or examinations. Lastly, universities are advised to consider providing college counselors to support students in managing their emotional and motivational needs, particularly in dealing with academic adversities and difficulties during routine lecture activities. Counselors can help students develop confidence to maximize success opportunities. Furthermore, counselors can assist students by illustrating that mistakes can be stepping stones toward success and do not define a person's worth. They can also reframe the concept of success, emphasizing personal progress and improvement rather than simply outperforming others.

Limitation

Although this study explained the relationship between academic buoyancy, motivational constructs, and learning achievement, several limitations should be acknowledged. First, the research was conducted on college students as participants, so the findings cannot be generalized to other educational levels, such as elementary and middle school. College students often face greater academic pressures, such as higher academic demands, more challenging coursework, and increased competition, than elementary or middle school students. Second, this study excluded demographical factors, such as gender and discipline background, from the analysis. While our research provides valuable insights into relationships within the variables (academic buoyancy, motivation, and learning achievement), the need to consider these demographic variables may overlook significant differences that could influence the results. Including gender, discipline background, and other demographic factors in future studies could offer a more nuanced understanding of how these variables impact the observed phenomena. We acknowledge this limitation and suggest that subsequent research incorporate these factors to enhance the robustness and generalizability of the findings. Last, this study used a cross-sectional design to examine the research hypotheses. This type of design captures data at a single point in time, which restricts our ability to infer causality or observe changes and developments over time. Consequently, the relationships observed in this study may only partially reflect the dynamic nature of the variables involved. We recommend that future research employ longitudinal designs to better capture the complexities and evolution of the phenomena under investigation.

CONCLUSION

This study aims to explain how academic buoyancy affects learning achievement by presenting motivational and emotional constructs as mediating variables in the research model. The findings indicate that academic buoyancy positively predicts learning achievement. Furthermore, the motivational constructs (self-efficacy, persistence, and anxiety) significantly mediate the link between academic buoyancy and learning achievement. The motivational constructs clearly explain the mechanism of academic buoyancy affecting learning achievement.

Students with high academic buoyancy often possess strong self-efficacy. This belief in their ability to overcome challenges empowers them to engage actively in learning and seek help when needed. In Addition, academic buoyancy also fosters persistence by equipping students with the resilience to bounce back from setbacks. They are more likely to persevere in their studies even when faced with challenges. These situations will help students achieve better academic performance. Lastly, academic buoyancy enables students to face adversities and setbacks with a positive outlook, fostering emotional control. This ability to manage emotions helps reduce anxiety when confronted with academic adversities, ultimately leading to better learning achievement.

REFERENCES

af Ursin, P., Järvinen, T. and Pihlaja, P. (2021) 'The Role of Academic Buoyancy and Social Support in Mediating Associations Between Academic Stress and School Engagement in Finnish Primary School Children', *Scandinavian Journal of Educational Research*, Vol. 65, No. 4, pp. 661–675. https://doi.org/10.1080/00313831.2020.1739135

Alp Christ, A., Capon-Sieber, V., Grob, U. and Praetorius, A.-K. (2022) 'Learning processes and their mediating role between teaching quality and student achievement: A systematic review', *Studies in Educational Evaluation*, Vol. 75, p. 101209. https://doi.org/10.1016/j.stueduc.2022.101209

Bae, S.-M. (2017) 'An analysis of career maturity among Korean youths using latent growth modeling', *School Psychology International*, Vol. 38, No. 4, pp. 434–449. https://doi.org/10.1177/0143034317709527

Bae, S. M. (2022) 'The Analysis of a Causal Relationship between Career Maturity and Academic Achievement on Korean Adolescents Using Autoregressive Cross-Lagged Modeling', *International Journal of Environmental Research and Public Health*, Vol. 19, No. 9. https://doi.org/10.3390/ijerph19095572

Bandura, A. (1997) *Self-efficacy: The exercise of control, New York: Henry Holt & Co.*

Bosch, E., Seifried, E. and Spinath, B. (2021) 'What successful students do: Evidence-based learning activities matter for students' performance in higher education beyond prior knowledge, motivation, and prior achievement', Learning and Individual Differences, Vol. 91, p. 102056. https://doi.org/10.1016/j.lindif.2021.102056

Bostwick, K. C. P., Collie, R. J., Burns, E. C., Hare, N., Cox, S., Flesken, A. and McCarthy, I. (2022) 'Academic buoyancy in high school: A cross-lagged multilevel modeling approach exploring reciprocal effects with perceived school support, motivation, and engagement', *Journal of Educational Psychology*, Vol. 114, pp. 1931–1949. https://doi.org/10.1037/edu0000753

Cadoret, G., Bigras, N., Duval, S., Lemay, L., Tremblay, T. and Lemire, J. (2018) 'The mediating role of cognitive ability on the relationship between motor proficiency and early academic achievement in children', *Human Movement Science*, Vol. 57, pp. 149–157. https://doi.org/10.1016/j.humov.2017.12.002

Caprara, G. V., Vecchione, M., Alessandri, G., Gerbino, M. and Barbaranelli, C. (2011) 'The contribution of personality traits and self-efficacy beliefs to academic achievement: A longitudinal study', *British Journal of Educational Psychology*, Vol. 81, No. 1, pp. 78–96. https://doi.org/10.1348/2044-8279.002004

Choi, S. (2018) 'Analysis of Relationship between Academic Achievement and Job Satisfaction of Secondary NCS Vocational Education', *Journal of Service Research and Studies*, Vol. 8, No. 1, pp. 41–57. https://doi.org/10.18807/JSRS.2018.8.1.041

Collie, R. J., Martin, A. J., Malmberg, L.-E., Hall, J. and Ginns, P. (2015) 'Academic buoyancy, student's achievement, and the linking role of control: A cross-lagged analysis of high school students', *British Journal of Educational Psychology*, Vol. 85, No. 1, pp. 113–130. https://doi.org/10.1111/bjep.12066

Corcoran, R. P., Cheung, A. C. K., Kim, E. and Xie, C. (2018) 'Effective universal school-based social and emotional learning programs for improving academic achievement: A systematic review and meta-analysis of 50 years of research', *Educational Research Review*, Vol. 25, pp. 56–72. https://doi.org/10.1016/j.edurev.2017.12.001

Datu, J. A. D. and Yang, W. (2021) 'Academic buoyancy, academic motivation, and academic achievement among filipino high school students', *Current Psychology*, Vol. 40, No. 8, pp. 3958–3965. https://doi.org/10.1007/s12144-019-00358-y

Demetriou, A., Kazi, S., Spanoudis, G. and Makris, N. (2019) 'Predicting school performance from cognitive ability, self-representation, and personality from primary school to senior high school', *Intelligence*, Vol. 76, p. 101381. https://doi.org/10.1016/j.intell.2019.101381

Dogaru, M. M. and Popescu, M. (2021) 'The Relevance of Academic Performance in Career Development According to the Big Five Model', *International Journal of Innovation in the Digital Economy (IJIDE)*, Vol. 12, No. 4, pp. 64–75. https://doi.org/10.4018/IJIDE.2021100105

Durlak, J. A., Weissberg, R. P., Dymnicki, A. B., Taylor, R. D. and Schellinger, K. B. (2011) 'The Impact of Enhancing Students' Social and Emotional Learning: A Meta-Analysis of School-Based Universal Interventions', *Child Development,* Vol. 82, No. 1, pp. 405–432. https://doi.org/10.1111/j.1467-8624.2010.01564.x

Dyer, E. D. (1987) 'Can University Success and First-Year Job Performance Be Predicted from Academic Achievement, Vocational Interest, Personality and Biographical Measures?', *Psychological Reports*, Vol. 61, No. 2, pp. 655–671. https://doi.org/10.2466/pr0.1987.61.2.655

Fong, C. J. and Kim, Y. W. (2021) 'A clash of constructs? Re-examining grit in light of academic buoyancy and future time perspective', *Current Psychology*, Vol. 40, No. 4, pp. 1824–1837. https://doi.org/10.1007/s12144-018-0120-4

Fornell, C. and Larcker, D. F. (1981) 'Structural Equation Models with Unobservable Variables and Measurement Error: Algebra and Statistics', *Journal of Marketing Research*, Vol. 18, No. 3, pp. 382–388. https://doi.org/10.2307/3150980

Granziera, H., Liem, G. A. D., Chong, W. H., Collie, R. J., Bishop, M. and Tynan, L. (2022) 'The role of teachers' instrumental and emotional support in students' academic buoyancy, engagement, and academic skills: A study of high school and elementary school students in different national contexts', *Learning and Instruction*, Vol. 80, p. 101619. https://doi.org/10.1016/j.learninstruc.2022.101619

Grønlien, H. K., Christoffersen, T. E., Ringstad, Ø., Andreassen, M. and Lugo, R. G. (2021) 'A blended learning teaching strategy strengthens the nursing students' performance and self-reported learning outcome achievement in an anatomy, physiology and biochemistry course – A quasi-experimental study', *Nurse Education in Practice*, Vol. 52, p. 103046. https://doi.org/10.1016/j.nepr.2021.103046

Haidari, S. M., Koçoğlu, A. and Kanadlı, S. (2023) 'Contribution of Locus of Control, Self-Efficacy, and Motivation to Student Achievement: A Meta-Analytic Structural Equation Modelling', *Journal on Efficiency and Responsibility in Education and Science*, Vol. 16, No. 3, pp. 245–261. https://doi.org/10.7160/eriesj.2023.160308

Hair, J. F., Howard, M. C. and Nitzl, C. (2020) 'Assessing measurement model quality in PLS-SEM using confirmatory composite analysis', *Journal of Business Research*, Vol. 109, pp. 101–110. https://doi.org/10.1016/j.jbusres.2019.11.069

Hair, J. F., Sarstedt, M., Hopkins, L. and G. Kuppelwieser, V. (2014) 'Partial least squares structural equation modeling (PLS-SEM)', *European Business Review*, Vol. 26, No. 2, pp. 106–121. https://doi.org/10.1108/EBR-10-2013-0128

Hirvonen, R., Putwain, D. W., Määttä, S., Ahonen, T. and Kiuru, N. (2020) 'The role of academic buoyancy and emotions in students' learning-related expectations and behaviours in primary school', *British Journal of Educational Psychology*, Vol. 90, No. 4, pp. 948–963. https://doi.org/10.1111/bjep.12336

Hopfenbeck, T. N., Lenkeit, J., El Masri, Y., Cantrell, K., Ryan, J. and Baird, J.-A. (2018) 'Lessons Learned from PISA: A Systematic Review of Peer-Reviewed Articles on the Programme for International Student Assessment', *Scandinavian Journal of Educational Research*, Vol. 62, No. 3, pp. 333–353. https://doi.org/10.1080/00313831.2016.1258726

Jia, Y. and Cheng, L. (2022) 'The Role of Academic Buoyancy and Social Support on English as a Foreign Language Learners' Motivation in Higher Education', *Frontiers in Psychology*, Vol. 13, pp. 1–7. https://doi.org/10.3389/fpsyg.2022.892603

Köller, O. (2012) 'What works best in school? Hatties Befunde zu Effekten von Schul- und Unterrichtsvariablen auf Schulleistungen [What works best in school? Hattie's findings on effects of school and teaching variables on academic achievement]', *Psychologie in Erziehung und Unterricht*, Vol. 59, pp. 71–78. https://doi.org/10.2378/peu2012.art06d

Lei, W., Wang, X., Dai, D. Y., Guo, X., Xiang, S. and Hu, W. (2022) 'Academic self-efficacy and academic performance among high school students: A moderated mediation model of academic buoyancy and social support', *Psychology in the Schools*, Vol. 59, No. 5, pp. 885–899. https://doi.org/10.1002/pits.22653

Liu, S., Liu, S., Liu, Z., Peng, X. and Yang, Z. (2022) 'Automated detection of emotional and cognitive engagement in MOOC discussions to predict learning achievement', *Computers & Education*, Vol. 181, p. 104461. https://doi.org/10.1016/j.compedu.2022.104461

Marks, G. N. and O'Connell, M. (2023) 'The importance of parental ability for cognitive ability and student achievement: Implications for social stratification theory and practice', *Research in Social Stratification and Mobility*, Vol. 83, p. 100762. https://doi.org/10.1016/j.rssm.2023.100762

Martin, A. J. (2008) *The Motivation and Engagement Scale: University/College*, [Online], Available: https://lifelongachievement.com/products/motivation-and-engagement-scale-university-college [February 17, 2023]

Martin, A. J. and Marsh, H. W. (2020) 'Investigating the reciprocal relations between academic buoyancy and academic adversity: Evidence for the protective role of academic buoyancy in reducing academic adversity over time', *International Journal of Behavioral Development*, Vol. 44, No. 4, pp. 301–312. https://doi.org/10.1177/0165025419885027

Pluck, G., Bravo Mancero, P., Ortíz Encalada, P. A., Urquizo Alcívar, A. M., Maldonado Gavilanez, C. E. and Chacon, P. (2020) 'Differential associations of neurobehavioral traits and cognitive ability to academic achievement in higher education', *Trends in Neuroscience and Education*, Vol. 18, p. 100124. https://doi.org/10.1016/j.tine.2019.100124

Putwain, D. W. and Aveyard, B. (2018) 'Is perceived control a critical factor in understanding the negative relationship between cognitive test anxiety and examination performance?', *School Psychology Quarterly*, Vol. 33, pp. 65–74. https://doi.org/10.1037/spq0000183

Putwain, D. W., Gallard, D. and Beaumont, J. (2020) 'Academic buoyancy protects achievement against minor academic adversities', *Learning and Individual Differences*, Vol. 83–84, p. 101936. https://doi.org/10.1016/j.lindif.2020.101936

Putwain, D. W. and Wood, P. (2023) 'Riding the bumps in mathematics learning: Relations between academic buoyancy, engagement, and achievement', *Learning and Instruction*, Vol. 83, p. 101691. https://doi.org/10.1016/j.learninstruc.2022.101691

Sanfo, J.-B. M. B. and Malgoubri, I. (2023) 'Teaching quality and student learning achievements in Ethiopian primary education: How effective is instructional quality in closing socioeconomic learning achievement inequalities?', *International Journal of Educational Development*, Vol. 99, p. 102759. https://doi.org/10.1016/j.ijedudev.2023.102759

Sarstedt, M., Hair, J. F., Cheah, J.-H., Becker, J.-M. and Ringle, C. M. (2019) 'How to specify, estimate, and validate higher-order constructs in PLS-SEM', *Australasian Marketing Journal (AMJ)*, Vol. 27, No. 3, pp. 197–211. https://doi.org/10.1016/j.ausmj.2019.05.003

Schneider, M. and Preckel, F. (2017) 'Variables associated with achievement in higher education: A systematic review of meta-analyses', *Psychological Bulletin*, Vol. 143, pp. 565–600. https://doi.org/10.1037/bul0000098

Senler, B. (2022) 'Relationship Between Student Perceptions of a Constructivist Learning Environment, and their Motivational Beliefs and Self-Regulation of Effort', *Journal on Efficiency and Responsibility in Education and Science*, Vol. 15, No. 2, pp. 72–81. https://doi.org/10.7160/eriesj.2022.150202

Tong, D. H., Uyen, B. P. and Ngan, L. K. (2022) 'The effectiveness of blended learning on students' academic achievement, self-study skills and learning attitudes: A quasi-experiment study in teaching the conventions for coordinates in the plane', *Heliyon*, Vol. 8, No. 12, p. e12657. https://doi.org/10.1016/j.heliyon.2022.e12657

Wang, X. and Hui, L. (2024) 'Buoyancy and engagement in online English learning: The mediating roles and complex interactions of anxiety, enjoyment, and boredom', *System*, Vol. 125, p. 103418. https://doi.org/10.1016/j.system.2024.103418

Wise, D. A. (1975) 'Academic Achievement and Job Performance', *The American Economic Review*, Vol. 65, No. 3, pp. 350–366. https://www.jstor.org/stable/1804839

Yiin, S.-J. and Chern, C.-L. (2023) 'The effects of an active learning mechanism on cognitive load and learning achievement: A new approach for pharmacology teaching to Taiwanese nursing students', *Nurse Education Today*, Vol. 124, p. 105756. https://doi.org/10.1016/j.nedt.2023.105756

Yu, K., Martin, A. J., Hou, Y., Osborn, J. and Zhan, X. (2019) 'Motivation, Engagement, Academic Buoyancy, and Adaptability: The Roles of Socio-Demographics among Middle School Students in China', *Measurement: Interdisciplinary Research and Perspectives*, Vol. 17, No. 3, pp. 119–132. https://doi.org/10.1080/15366367.2019.1584518

Yun, S., Hiver, P. and Al-Hoorie, A. H. (2018) 'Academic buoyancy: Exploring learners' everyday resilience in the language classroom', *Studies in Second Language Acquisition*, Vol. 40, No. 4, pp. 805–830. https://doi.org/10.1017/S0272263118000037

EFFECT OF 2-PL AND 3-PL MODELS ON THE ABILITY ESTIMATE IN MATHEMATICS BINARY ITEMS

Rukayat Oyebola Iwintolu[1]
Oluwaseyi Aina Gbolade Opesemowo[2]✉
Phebean Oluwaseyi Adetutu[3]

[1]Osun State University, Nigeria

[2]University of Johannesburg, South Africa

[3]Obafemi Awolowo University, Nigeria

✉ oopesemowo@uj.ac.za

Article history
Received
April 13, 2024
Received in revised form
July 10, 2024
Accepted
September 3, 2024
Available on-line
September 30, 2024

ABSTRACT

The investigation delves into examining the influence of 2-parameter logistic (PL) and 3-parameter logistic models on the ability estimates of students in mathematical binary items. It ascertained the parameters of the items in the 2-PL and 3-PL models. We employed Item Response Theory (IRT) in the design of this research survey, with a sample comprising 1015 senior secondary (SS) students in SS III classes who were analyzed using both models in the investigation. The Mathematics Achievement Test instrument was adapted from the General Mathematics Paper 1 of the Senior School Certificate Examination administered by the West Africa Examinations Council (WAEC). Results indicated that the 2-PL model shows lower difficulty levels but higher discriminatory indices. Statistical analysis revealed a significant ($F = 19.52$, $p < 0.05$ and $F = 18.52$, $p < 0.05$) effect of both models, respectively, on ability estimates in mathematics binary items among Nigerian secondary school students. We established that item parameters in the 2-PL and 3-PL models significantly affected the ability estimate of Nigeria secondary school students in binary mathematics items, while the 3-PL model provided a better ability estimate than the 2-PL model.

KEYWORDS

Items, binary items, models, item response theory, item parameters

HOW TO CITE

Iwintolu R. O., Opesemowo O. A. G., Adetutu P. O. (2024) 'Effect of 2-PL and 3-PL Models on the Ability Estimate in Mathematics Binary Items', *Journal on Efficiency and Responsibility in Education and Science*, vol. 17, no. 3, pp. 257–272. http://dx.doi.org/10.7160/eriesj.2024.170308

Highlights

- *The efficacy of 2-PL and 3-PL models were evaluated in estimating students' ability in Mathematics binary items.*
- *The models had a significant effect in estimating students' ability*
- *The 3-PL model estimated ability better than the 2-PL model.*
- *Continued research into the effects of item parameters on ability estimation across different subjects and grade levels will be crucial for advancing assessment practices and promoting academic success.*

INTRODUCTION

Tests are standardized instruments used to obtain a sample of an examinee's best attempt at aptitude/achievement test, which gives an estimate of their performance/ability (Adetutu and Iwintolu, 2017; Breuer et al., 2023; Gates, 2023; Opesemowo et al., 2018) or a representation of an individual's standard performance on surveys or assessments where they reveal their typical emotions, beliefs, preferences, or responses to situations (O'Connor et al., 2019; Powers, 2019). Different peculiarities, strengths, and weaknesses characterize the aptitude/achievement tests, including the essay and objective tests. There are various objective tests: the short-answered test, the completion test, multiple choice, matching, cloze tests, and binary choice tests. The multiple-choice tests (binary scored) have gained significant acceptance among item-generation experts, even in Nigeria's standardized tests (Opesemowo et al., 2023). Among the objective test types, the multiple-choice test is generally known as the most commonly relevant, valuable, and used (Danh et al., 2020). It is fit for measuring complex outcomes in knowledge, understanding, application, and problem-solving skills. In Nigeria, like other countries, multiple-choice items are popular test types among examination organizations, as Douglas et al. (2023), Kalhori and Abbasi (2017), and Rios and Soland (2022) alluded. The organizations include the West Africa Examinations Council (WAEC), National Examinations Council (NECO), National Teachers Institution (NTI) Examination, and Joint Admission and Matriculation Board (JAMB) are some of the organizations involved.

Essentially, users of multiple-choice tests typically use binary scoring (i.e., assigning a value of one for a correct response and zero for an incorrect response), commonly analyzed using Classical Test Theory (CTT) techniques due to its ease of interpretation. Using CTT, examinees' raw scores are summed;

therefore, all tests and examinees are considered together in this model. Despite its widespread use, CTT has been critiqued based on its inability to capture the essence of a test taker's ability, as the actual score is not an inherent trait. Moreover, the difficulty of individual test items may fluctuate based on the composition of the test-takers, making comparison of results difficult across different tests. In contrast, the Item Response Theory (IRT) technique is garnering recognition in the fields of psychological and educational testing owing to its provision of more flexible and efficient approaches to test development, evaluation, and scoring compared to those stemming from CTT (Adetutu and Iwintolu, 2017; Olagunju and Iwintolu, 2023). In IRT, individual items and individual test takers are the objects of analysis (Awopeju and Afolabi, 2016; Setiawati et al., 2023). As a result, IRT is contingent upon the individual items within a test instead of a collective measure of item responses like test scores for its basic concepts (Alordiah, 2015; Baker, 2001).

Furthermore, among the ultimate assumptions of IRT is that the respective test examinee responding to a test item has some level of the underlying ability that the item is intended to measure. This underlying ability is called a latent trait, and IRT models seek to estimate the latent based on the examinee's responses to various test items. By accurately estimating an examinee's latent trait, IRT can provide more precise and reliable measures of ability than traditional test-scoring methods. However, in practical terms, we cannot directly measure the value of the examinee's ability parameter; thus, the best approach is to estimate it (Ayanwale, 2023; Bichi and Talib, 2018). A numerical score on the ability scale can represent each examinee. At different ability levels (θ), there is a probability that an examinee will answer an item correctly regardless of their ability level. This probability, denoted as P(θ), is low for examinees with lower abilities and high for those with higher abilities. When an examinee faces a set of test items during an examination, they bring their inherent ability (θ) or trait into the testing environment (Rafi et al., 2023; Zanon et al., 2016). Tests are designed to evaluate an examinee's position on the ability scale, enabling a standardized comparison of examinees to ascertain their relative placements (Rudner, 2019; Scheibling-Sève et al., 2020). Obtaining ability measures for everyone taking the test can help achieve two critical objectives. Firstly, it allows for appraising the examinee's underlying ability level. Secondly, it enables comparisons among examinees to determine rates, assign grades, award scholarships, and more.

IRT models, such as the 2 Parameter Logistic and 3 Parameter Logistic (PL) models, have become instrumental in educational assessment, particularly in measuring the student's abilities in mathematics. These models provide an advanced framework for analyzing binary items, where responses are either correct or incorrect, and have been extensively utilized in various educational settings (Jimoh et al., 2022). The influential nature of these models in estimating mathematics abilities is a subject of significant interest and research due to its implications for curriculum design, instructional strategies, and student evaluation. IRT presupposes an examinee can steadily provide correct responses to test items, contingent upon possessing the requisite abilities as demanded by the items. The interaction between the individual's trait and the parameters of the items determines the probability of answering a test item correctly. One of the main objectives of IRT is to establish a relationship between latent variables, such as the examinee's ability, and the likelihood of providing correct responses to test items. The primary models currently utilized are the 1-PL, 2-PL, and 3-PL.

THEORETICAL STRUCTURE

1-Parameter Logistic Model: This model is regarded as the most foundational IRT model. It is presumed that only one item parameter underpins the item response procedure. IRT literature often refers to this item parameter as difficulty, symbolized by *b* in the 1-PL model (Yustiandi and Saepuzaman, 2021). The *b*-parameter, representing a test item, typically aligns with θ, indicating a trait of an individual under consideration. Within this framework, all test items display an identical Item Characteristic Curve (ICC), differing solely in their positioning along the horizontal axis (θ). The b-parameter represents the item's or task's cognitive resistance in each cognitive task. The formula is presented below:

$$P_i\left(\theta_j\right)=\frac{1}{1+\exp\left[-\left(\theta_j-b_i\right)\right]} \quad (1)$$

where $P_i(\theta_j)$ = probability of examinee with ability θ_j answering item *i* correctly
exp = exponential is constant (2.718)
θ_j = ability estimates
b_i = item *i* difficulty parameter

2-Parameter Logistic Model: the parameter uncovers possible flaws in the 1-PL model if all test items have identical shapes in the ICC. In response, the 2-PL model introduces a parameter called discrimination, expressed by *a*, which permits the ICC for diverse items to have distinct slopes (Perez and Loken, 2023). The discrimination parameter will enable us to model items with more significant (or weaker) relationships to the assessed construct (θ) than others; a high discrimination index indicates stronger ties between the item and the construct, while a low discrimination index indicates a weaker relationship. The a-parameter is significant in IRT since it directly influences an item's information. This model assumes that the examinee's competence and difficulty level of the question ascertain the chance of responding correctly to an item. The level of simplicity and practicality of the 2PL model has given it a wide application in educational testing. It provided valuable insights into students' abilities based on their responses to binary items. Mathematically, it is expressed as below:

$$P_i\left(\theta_j\right)=\frac{1}{1+\exp\left[-a_i\left(\theta_j-b_i\right)\right]} \quad (2)$$

where $P_i(\theta_j)$ = probability of examinee with ability θ_j answering item *i* correctly
θ_j = ability estimates
a_i = item *i* discrimination parameter
b_i = item *i* difficulty parameter

3-Parameter Logistic Model: the 2-PL model aims to tackle a specific critique of the Rasch model, which assumes that all test items exhibit uniform discriminating capability. Still, it has been found deficient in addressing another possibly essential factor that may vary among items: the lower asymptote of the ICC refers to the anticipated proportion of correct or key responses exhibited by participants with exceedingly low θ scores. Including the *c*-parameter in the 3-PL model accounts for the likelihood of correctly predicting the item, causing the lower asymptote of the ICC to be potentially non-zero. This contrasts the 1-PL and 2-PL models, where the ICC's lower asymptote is permanently set to zero (Paek et al., 2023). This parameter accounts for the likelihood of guessing correct responses to items despite examinees' lack of the necessary knowledge or skill to be successful at items. In mathematics assessment, the 3-PL model provides a refined method for estimating abilities, wildly when guessing influences IRT, such as in multiple-choice tests that allow partial credit. In assessments comprising multiple-choice questions, examinees lacking the requisite knowledge of the accurate solution are likely to resort to guessing, leading to the need for a non-zero lower asymptote (and on occasion, they may select the correct option). The formula is presented below:

$$P_i\left(\theta_j\right) = c_i + \frac{1 - c_i}{1 + \exp^{-1.7a_i\left(\theta_j - b_i\right)}} \quad (3)$$

where $P_i(\theta_j)$ = probability of participant with ability θ_j answering item *i* correctly
θ_j = ability estimates
a_i = item *i* discrimination parameter
b_i = item *i* difficulty parameter
c_i = item *i* guessing parameter

Statement of the Problem

The objective of testing is to attain an accurate measure of what we need to assess, and the accuracy of test measurements is decided mainly by the test data at each examinee's ability level. The fundamental concept of IRT originates from the principles of the item response model, which entails a mathematical function elucidating the likelihood of particular responses to an item based on various quantitative attributes of the respondents (Frick et al., 2024; Jimoh et al., 2022; von Davier, 2019). Three IRT models, namely the 1-PL, 2-PL, and 3-PL models, have been developed and implemented in various studies for item calibration. This process involves determining the properties of items and estimating the examinees' abilities. As a result, the difficulty in calibrating objects stems from deciding which models to use. However, the Rasch model claims to be a reliable measuring criterion. It contends that factors other than the difficulty of the items are likely to influence examinees' responses (Stemler and Naples, 2021). Nevertheless, other researchers have dismissed it as experimentally meaningless because it does not account for changes in discrimination and guessing (i.e., chance) factors. The Rasch model implies that guessing is irrelevant and that all objects have the same discrimination value. This study did not examine the Rasch model. Conversely, this study aims to ascertain the effect of the 2-PL and 3-PL on examinees' ability estimates in binary mathematics items. Specifically, the study also aims to establish difficulty and discrimination indices in the 2-PL and 3-PL models of the examinees' responses, determine the effect of the 2-PL model on ability estimates in mathematics binary items, and assess the effect of the 3-PL model on ability estimates in mathematics binary items. By analyzing the binary data collected from administering the instrument, we hope to understand how each model affects the accuracy of the ability estimates. This study will provide valuable insight for educators and test developers looking to elevate the consistency and authenticity of their assessment tools in mathematics education.

Research Questions

1. What are the items' difficulty and discrimination indices in the 2-PL?
2. What are the items' difficulty, discrimination, and guessing indices in the 3-PL?

Research Hypotheses

1. The effect of the 2-PL model on the ability estimates in mathematics binary items is deemed significant.
2. The effect of the 3-PL model on the ability estimates in mathematics binary items is deemed significant.

MATERIALS AND METHODS

A descriptive survey design was utilized in this study. This type of research design focuses on describing data without manipulating variables. There were 1015 participants in the study, all from senior secondary school three (SS3). The study sample consisted of 522 (51.4%) male students and 493 (48.6%) female students. Approximately 42% of the participants were enrolled in private educational institutions, whereas 58% were in public institutions, including federal and state-owned schools. We adapted the Mathematics Achievement Test (MAT) instrument from the WAEC General Mathematics Paper 1 of the June/July (2006-2014) SSCE. WAEC has been responsible for conducting standardized examinations across the West African region (Kennedy and Ebuwa, 2022). This exam is crucial in determining students' academic performance and progress in their respective countries. The WAEC ensures that the exams are fair and transparent, allowing all students to showcase their knowledge and ability. The instrument (MAT) consists of a 20-item mathematics multiple-choice test, and we scored the response binarily. Using a stratified random sampling method, the participants were chosen to guarantee inclusivity from private and public educational institutions. The MAT was administered to the students under standard examination conditions to measure their mathematical achievement. The data were analyzed using IRTPRO. The IRTPRO is a statistical analysis software tool used for IRT analysis for binary and polytomous datasets. It can also perform unidimensional and multidimensional IRT analysis and support models, including 1, 2, and 3-PL models.

Analysis

The data went through a preliminary analysis. Descriptive analysis was employed to ascertain the occurrence rate

of each item's response, alongside calculating the mean, maximum, minimum, and standard deviation for each item. The test unidimensionality, which corresponds with the IRT assumptions, was established. According to Choi et al. (2023), Kim (2017), and Opesemowo et al. (2023), the assumption of unidimensionality implies that the item examines a single ability and that the response satisfies the local independence principle, which states that item responses depend on a particular ability level independently. Nevertheless, an exploratory factor analysis (EFA) can verify unidimensionality if one of the two conditions is met. First, the unrotated factor matrix should show that the first component explains at least 20% of the variance based on the inter-item correlation matrix. Second, the eigenvalues of the first component must be greater than those of the second factor. In this study, we assessed unidimensionality using exploratory factor analysis. Furthermore, a scree plot was created to see if unidimensionality could be inferred. A scree plot is a valuable diagram for visualizing a principal component analysis (PCA) leading factor. In a scree plot, a dominating factor stands out over the ICC's elbow break.

The item difficulty parameter estimations were analyzed when responding to the research questions. Items with high *b*-values are often tricky (difficult) items under the IRT model; these are the questions that low-ability examinees are unlikely to answer accurately. Items with low *b*-values, on the other hand, are classed as easy (simple) items; these are questions that most examinees, including those with little aptitude, will have at least a moderate chance of answering correctly. Consequently, when interpreting the difficulty values, the following criteria are used: Difficulty values (*b*) that ranged between $-3.00 \leq -2.00$ is classified as very easy; *b*-values that ranged between $-2.00 \leq -1.00$ is classified as easy; $-1.00 \leq 1.00$ is classified as moderately difficult; $1.00 \leq 2.00$ is classified as difficult while ≥ 2.00 is categorized as very difficult (Bichi and Talib, 2018). In addition, the discriminating value reveals how well an item distinguishes between examinees of varied abilities. Discrimination indices for good items typically range from 0.5 to 2.0. In the 3-PL model, item discrimination is proportional to the slope of the item response function at the inflection point (0.25). The c-parameter has a theoretical range of $0 \leq C \leq 1.0$, although values higher than 0.35 are unacceptable (Adedoyin and Adedoyin, 2013; Baker, 2001).

To conduct the first hypothesis test, the effects of the 2-PL model (item difficulty and discrimination) on ability estimates were analyzed using an ANOVA. An ANOVA of the 3-PL model's (item difficulty, discrimination, and guessing) effects on ability estimates was also used to test the second hypothesis.

Ethical Consideration

Before the data collection, the ethical consideration was approved, and the participants were informed about the need to complete the MAT instrument responsibly and honestly, and their participation was completely voluntary. We established a confidentiality agreement to ensure that we would keep the collected data anonymous and use it solely for the research project. All ethical rules and procedures were strictly followed throughout the data collection process to preserve the participants' rights. We also notified participants that they could opt out of the study without repercussions. This made participants feel comfortable and confident that their privacy was respected throughout the research process. The ethical considerations the researchers took were vital in upholding the integrity of the study and respecting the individuals who had chosen to participate. Ultimately, these measures helped establish trust between the researchers and participants, creating a safe and respectful data collection environment.

RESULTS

Table 1 exhibits MAT items' mean, standard deviation, minimum and maximum scores, and response frequency.

Table 1 displays the frequencies of each item answer option and MAT's mean, maximum, minimum, and standard deviation.

Table 2 presents the eigenvalues and total variance explained with evidence that the test is unidimensional.

Table 2 showcases the EFA conducted on the 20-MAT. It produced six eigenvalues that are significantly more than one. The initial eigenvalue, 4.282, was more significant than the successive five eigenvalues (1.279, 1.131, 1.085, 1.010, and 1.004). The initial factor accounted for 21.41% of the total variation in the sample. The subsequent component accounted for 6.397% of the residual variance. However, the other 18 factors accounted for the rest of the variance. A scree plot, demonstrated in Figure 1, further validated the data's unidimensionality.

Research Question 1: What are the items' difficulty and discrimination indices in the 2-PL?

The items were subjected to a 2-PL model in the IRTPRO. Table 4 displays the item parameters, including the difficulty and discrimination indices.

Table 3 indicates that none of the items in the 2-PL were rated tricky. In contrast, only two items (1 and 15) exhibited poor discrimination since their discrimination indices fell below the 0.5 threshold.

Figure 2 illustrates the Total Information Curve (TIC), which compares the test data against the theta (ability) levels with their standard measurement error. The TIC enables researchers to visually analyze the relationship between test data and ability levels, providing valuable insight into the accuracy and precision of the test measurements. By examining the curve, researchers can assess how well the test differentiates between individuals with different ability levels and identify where measurement error is most likely to occur. This information can be used to make informed decisions about test design and interpretation, improving assessments' overall quality and reliability.

The curve shows a normal ability distribution, indicating the highly discriminatory test. The test characteristics curve (TCC) shown in Figure 3 further verified this conclusion. To display the psychometric structure, we also presented each item using an item category curve (ICC) (see Appendix A). Appendix A contains graphical representations of the components, known as ICC. They can show items that discriminate effectively and items that do not distinguish individuals at different levels of the items.

	Statistical Properties				Frequency of Response Options	
Item	M	SD	Min	Max	0	1
1	0.15	0.36	0	1	863	152
2	0.61	0.49	0	1	393	622
3	0.31	0.46	0	1	697	318
4	0.31	0.46	0	1	697	318
5	0.57	0.50	0	1	467	548
6	0.82	0.38	0	1	181	834
7	0.74	0.44	0	1	264	751
8	0.33	0.47	0	1	676	339
9	0.55	0.50	0	1	452	564
10	0.47	0.50	0	1	535	480
11	0.62	0.49	0	1	390	625
12	0.54	0.50	0	1	465	550
13	0.39	0.49	0	1	615	400
14	0.72	0.45	0	1	285	730
15	0.50	0.50	0	1	506	509
16	0.64	0.48	0	1	362	653
17	0.13	0.34	0	1	882	133
18	0.30	0.46	0	1	712	303
19	0.33	0.47	0	1	681	334
20	0.53	0.5	0	1	479	536

Note: Response option frequencies for each item total 1015 responses.

Table 1: Descriptive Statistics of MAT

Factor	Initial Eigenvalues		
	Total	% of Variance	Cumulative %
1	4.282	21.411	21.411
2	1.279	6.397	27.809
3	1.131	5.656	33.465
4	1.085	5.427	38.892
5	1.010	5.051	43.943
6	1.004	5.018	48.961

Extraction method: PCA

Table 2: Eigenvalues and Total Variance Explained

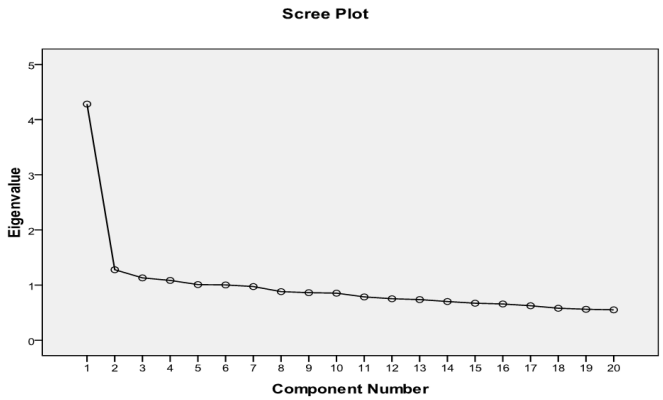


Figure 1: Scree Plot of MAT

Item	*b* (difficulty)	*a* (discrimination)
1	-3.53	-0.52
2	-0.28	1.21
3	-0.47	1.18
4	-0.47	1.03
5	-0.21	0.91
6	-1.71	1.09
7	-1.10	0.79
8	-0.17	1.12
9	-0.21	1.79
10	0.18	0.63
11	-0.36	0.89
12	-0.21	1.02
13	-0.18	1.31
14	-0.77	2.14
15	-0.02	0.46
16	-0.53	1.78
17	-1.62	0.94
18	-0.44	1.97
19	-0.05	0.82
20	-0.12	1.35

Table 3: Item Parameters of 2-PL Model

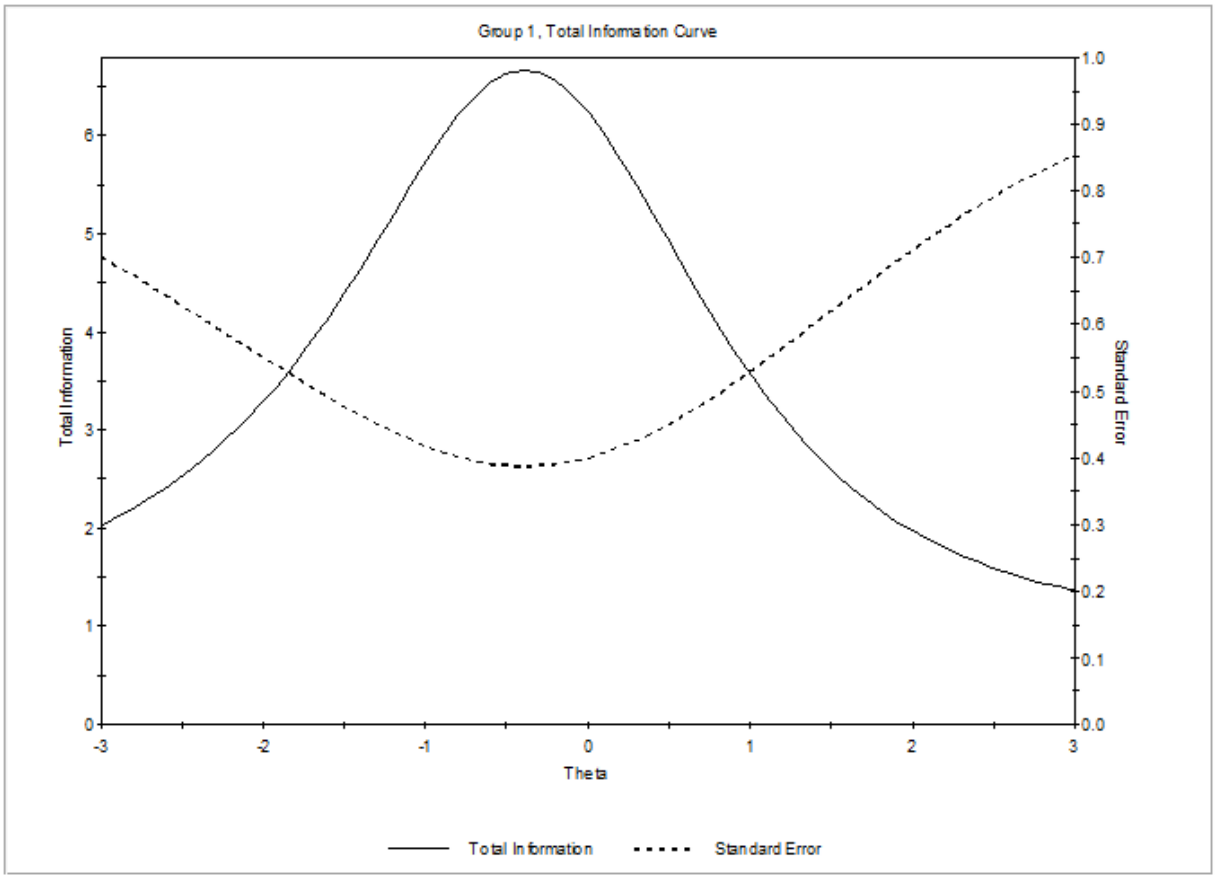


Figure 2: Test Information Curve of MAT

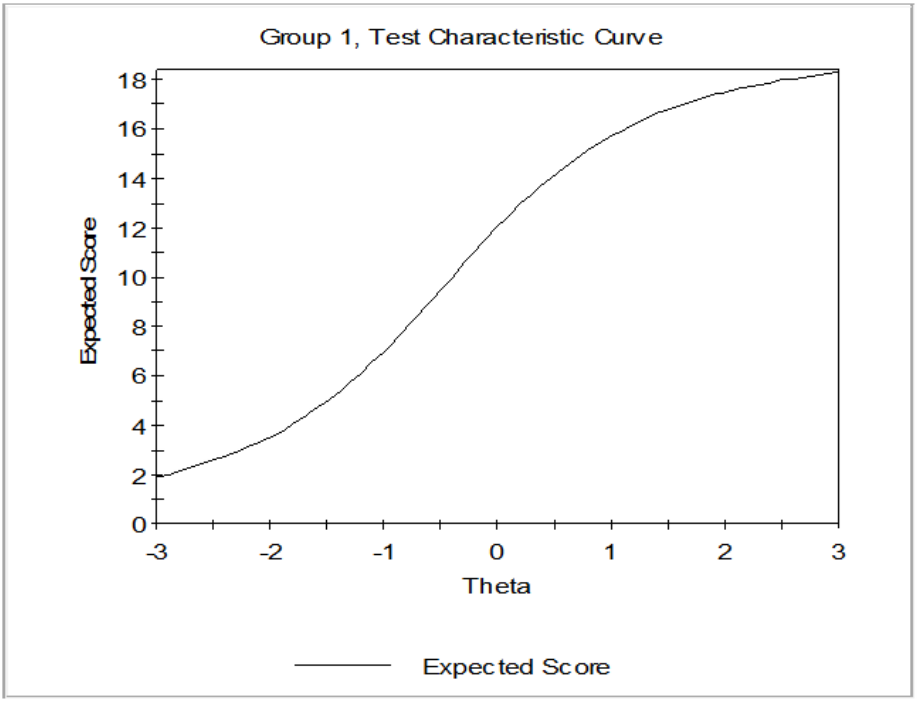


Figure 3: Test Characteristic Curve of MAT

Research Questions 2: What are the items' difficulty, discrimination, and guessing indices in the 3-PL?

The items underwent a 3-PL model within the IRTPRO software. Table 4 displays the item parameters, encompassing difficulty, discrimination, and guessing indices.

Table 4 exhibits the item parameters of the 3-PL model. Only item 1 was adjudged very difficult among the item difficulty parameter estimates; 11 items are moderately difficult, while eight items are easy. Based on the discrimination parameter estimations presented in Table 5, it was observed that only one item from a total of twenty failed to differentiate among the examinees. It is further obverse that the guessing parameter (c) highlights the exclusion of six items per the predefined criteria.

Item	*b* (difficulty)	*a* (discrimination)	*c* (guessing)
1	467.72	0.37	-171.99
2	0.12	1.61	-0.19
3	-0.02	1.55	0.03
4	-0.01	1.32	0.01
5	0.17	1.09	-0.19
6	-1.44	1.11	1.60
7	-0.68	0.85	0.58
8	0.36	1.91	-0.69
9	0.10	2.66	-0.27
10	0.93	1.15	-1.07
11	0.01	1.00	-0.01
12	0.20	1.30	-0.26
13	0.13	1.73	-0.23
14	-0.48	2.68	1.30
15	0.79	0.60	-0.47
16	-0.28	2.12	0.59
17	-1.23	0.98	1.21
18	-0.28	2.16	0.60
19	0.55	1.26	-0.69
20	0.14	1.67	-0.24

Table 4: Item Parameters of 3-PL model

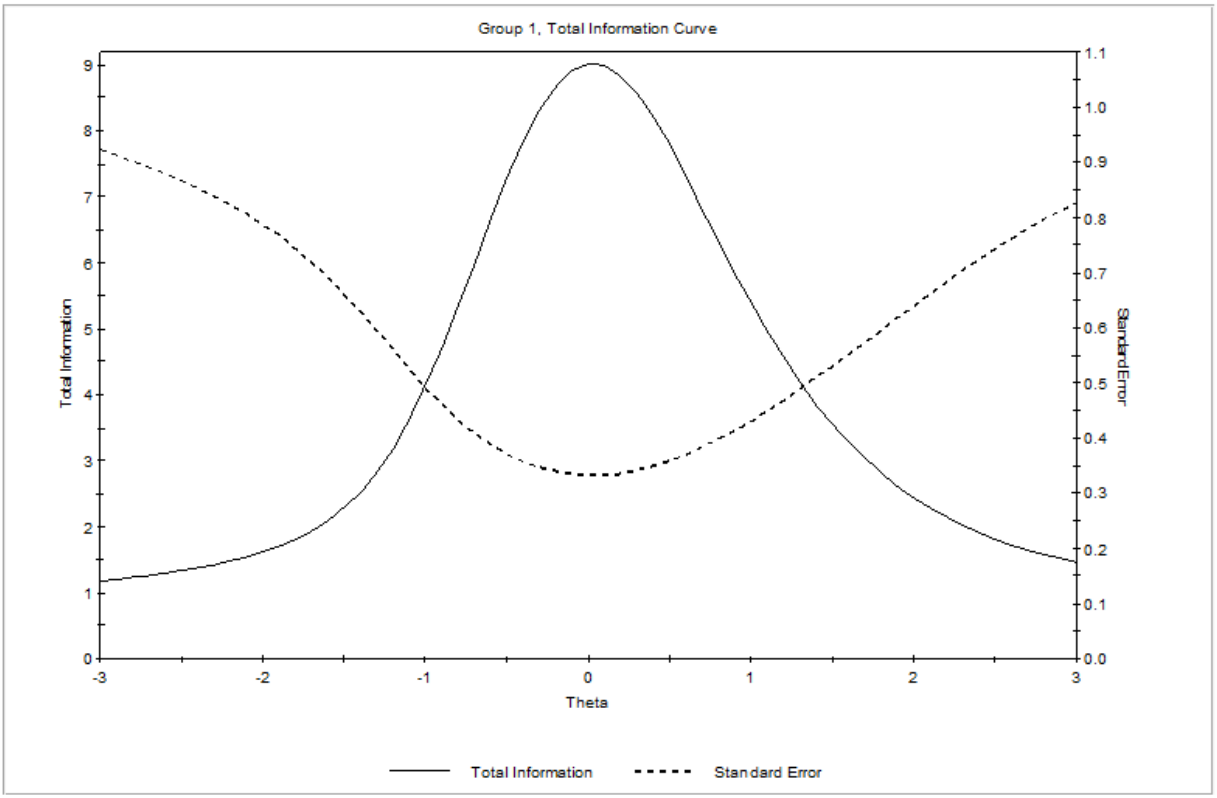


Figure 4: Total Information Curve of MAT

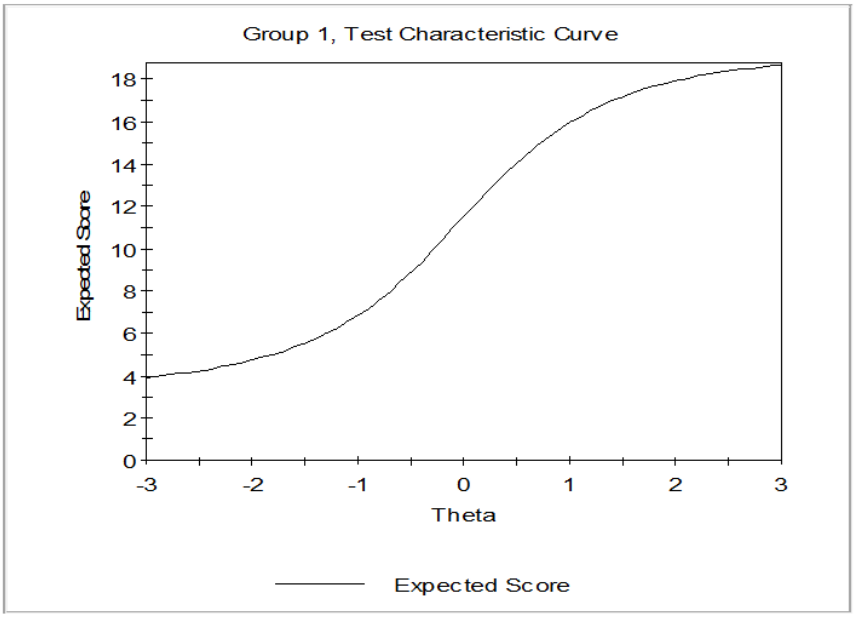


Figure 5: Test Characteristics Curve of MAT

The diagram (Figure 4) illustrates a typical distribution of abilities, demonstrating the high level of discrimination in the test. This allows researchers to identify potential areas for improvement in the test to ensure that it accurately reflects individuals' true abilities. Test developers can also make adjustments to minimize measurement error and increase the test's reliability by analyzing the curve. Figure 5 provided additional evidence supporting the Test Characteristic Curve (TCC). Each item's psychometric structure was characterized using an item category curve (see Appendix B). The item category curves presented in Appendix B serve as visual representations of the items, allowing for the identification of items that effectively discriminate and those that do not differentiate between individuals with different ability levels. Hypothesis one: the effect of the 2-PL model on the ability estimates in mathematics binary items is deemed significant.

	Sum of Squares	*df*	Mean Square	*F*	Sig.
Between groups	31.44	2	15.72		
Within groups	847.32	1052	0.80	19.52	0.000
Total	878.76	1054			

Table 5: ANOVA of the Effect of 2-PL parameters on Ability Estimates

The ANOVA results for the effect of the 2-PL model on ability estimates indicate a significant difference among the groups, as evidenced by a large F-value of 19.52 ($p < 0.05$). The between-groups variance (15.72) is substantially higher than the within-groups variance (0.80), suggesting that the variation in mathematics ability estimates of examinees because of item parameters (difficulty and discrimination) is explained mainly by differences between the groups rather than within them. This implies that the 2-PL model notably impacts ability estimates, underscoring the importance of considering these parameters in psychometric modelling.

Hypothesis Two: the effect of the 3-PL model on ability estimates in mathematics binary items is deemed significant.

	Sum of Squares	df	Mean Square	F	Sig.
Between groups	12290.52	3	4096.84		
Within groups	236861.19	1071	221.16	18.52	0.000
Total	249151.71	1074			

Table 6: ANOVA of the Effect of the 3-PL Model on Ability Estimates

The result unveils that the effect of the 3-PL (difficult, discrimination, and guessing) model on ability estimates in mathematics binary items is deemed significant with $F = 18.52$; $p < 0.05$ while the between-groups variance (4096.84) is higher than the within-groups variance (221.16). Therefore, the alternative hypothesis was supported.

Comparing the two ANOVA results in the 2-PL and 3-PL models, it is explicitly uncovered that the effect of the model on ability estimates is higher in the 3-PL model than in the 2-PL model. However, both have a significant effect on ability estimates. The 3-PL model can capture more nuanced variations in ability levels compared to the 2-PL model. This suggests that incorporating the additional parameter (i.e., guessing parameter) in the 3-PL model allows for a more accurate and precise estimation of individuals' abilities. These findings highlight the importance of selecting the appropriate item response theory model to ensure a valid and reliable measurement of abilities in educational and psychological assessments.

DISCUSSION

The 2-PL and 3-PL models utilized in the estimation of binary data raised apprehension due to the inclusion of parameters representing the item's difficulty, discrimination, and lower asymptote. These parameters were crucial in precisely evaluating the correlation between an individual's proficiency level and their reactions to particular items. The 2-PL model was beneficial for measuring discrimination between individuals with different levels of ability, while the 3-PL model also accounted for the guessing behavior of participants. Overall, these models provided a comprehensive framework for understanding and interpreting binary data in assessments and measurements of abilities. The assumption in the mathematics items holds reasonably with the factor analysis results.

The parameters derived from the 2-PL model exhibit low difficulty levels but with high discriminating indices (i.e., only two items exhibit low discrimination values). These findings suggest that the 2-PL model may not be the most suitable for accurately measuring examinees' abilities, as it tends to underestimate the discrimination of items. Additional research is needed to explore alternative models that may provide more precise estimations of item parameters and better reflect the true abilities of individuals. Additionally, considering the potential effect of item discrimination on test validity and reliability, it is crucial for researchers and practitioners to carefully evaluate the appropriateness of the chosen IRT model for their specific assessment needs. This was in tandem with the study of Perez and Loken (2023), who affirmed that item parameters in the 2-PL IRT model demonstrated well-estimated difficulties but noticeably underestimated discriminations, indicating low discrimination values rather than high. The result further aligned with the findings of Setiawati et al. (2023), presenting that the 2-PL model in the study showed low item difficulties but high discrimination indices, with low discrimination values indicating unique characteristics of the items.

The result of the 3-PL model introduced a c-parameter that affected the relationship between item difficulty and discrimination indices. Under the 3-PL model, item difficulties tend to be mostly average, leading to a higher discriminating index than in the 2-PL model (Perez and Loken, 2023). This validates the findings of Sweeney et al. (2022), who found that item difficulty and discrimination are mostly positively connected in the 3-PL model, as opposed to the negative correlation reported in the 2-PL. As a result, the 3-PL model's incorporation of the c-parameter alters these relationships, which may contribute to the model's higher discriminating indices (Ferreira-Junior et al., 2023). As a result, integrating the c-parameter in the 3-PL model significantly affects the item's difficulty and discrimination characteristics, distinguishing it from the 2-PL model. This suggests that most examinees, including those of low and medium aptitude, will have a reasonable probability of answering correctly.

Notably, item 1 in the 3-PL had the highest difficulty and the lowest discrimination indices. We traced this to uncertainty in the item's development, which resulted in examinees misinterpreting it.

The findings revealed that both models accounted for the effect of item parameters (item difficulty, discrimination, and guessing) on examinees' ability assessments in mathematics binary items. The 3-PL model introduces a c-parameter, affecting the relationship between item difficulty and discrimination indices. Under the 3-PL model, item difficulties tend to be mostly average, leading to a higher discriminating index than in the 2-PL model (Perez and Loken, 2023). The result supports the findings of Sweeney et al. (2022), revealing that item difficulty and discrimination are mostly positively correlated in the 3-PL model, in contrast to the negative correlation observed in the 2-PL model. Additionally, the 3-PL model's incorporation of the c-parameter influences these relationships, potentially contributing to the higher discriminating indices observed in this model (Ferreira-Junior et al., 2023). As a result, integrating the c-parameter in the 3-PL model significantly impacts the item's difficulty and discrimination indices, separating it from the 2-PL model. This implies that examinee estimates were dependent on item parameters. Although the models' item parameters differed in the index, both impacted ability estimates. The findings of this study correspond with the findings of Setiawati et al. (2023), which suggest that few and possibly non-significant differences exist in the assessment of item parameters in the 1-PL, 2-PL, and 3-PL.

In estimating a person's abilities, some empirical studies have explored the efficacy of IRT compared to CTT. A more accurate estimation of abilities is possible when using IRT due to its sensitivity to item characteristics (Suparman and Juandi, 2022). Some studies have shown slight differences between CTT and IRT estimates of abilities (Mutiawani et al., 2022). These studies emphasize the importance of accurate ability estimation in educational assessments, especially in scenarios like Computer Adaptive Testing (CAT), where precise estimations are crucial (Oladele et al., 2022; Opesemowo and Ndlovu, 2023). Additionally, IRT models, such as many-facet Rasch models (MFRMs), have been purported to enhance accuracy in measuring higher-order abilities, considering factors like rater severity and task difficulty (Sideridis and Alahmadi, 2022). Overall, the research underscores the significance of employing advanced statistical models like IRT for a more precise and reliable estimation of a person's abilities in various assessment contexts.

Despite the valuable results of this study, we should identify some limitations. First, the participants in the research were restricted to senior secondary school students in Nigeria. Hence, it is worth noting that other studies should consider junior and senior secondary school students from different countries to have a broader perspective of the student's ability estimate. Second, the subject of focus was mathematics assessment in Nigeria, which may have restricted the generalizability of the findings. However, it is essential to conduct further research across different disciplines and other countries to understand the ability to estimate better. Third, the data used in this study was binary, which may have impeded the findings.

Further studies can incorporate polytomous data to conduct a broader analysis. Lastly, the study concentrated on quantitative data, neglecting qualitative data that might provide impactful information. Integrating quantitative data into future research could provide more insight into the effects of 2-PL and 3-PL models on ability estimates in mathematics binary items.

CONCLUSION

Based on the study results, we concluded that the item parameters of the 2-PL and 3-PL models affected the ability estimates of examinees in Nigerian secondary schools for mathematics binary items. Furthermore, the study found that the 3-PL model bestowed more precise estimates of examinees' abilities than the 2-PL model. This suggests that using the 3-PL model for mathematics assessments in Nigerian secondary schools may create more accurate and reliable results. In addition, the study recommended further investigation into the effect of item parameters on ability estimation in other subject areas and grade levels to improve assessment practices in the country. In conclusion, the study highlighted the importance of utilizing advanced measurement models, such as the 3-PL model, to enhance the accuracy of ability estimates in mathematics assessments. By implementing this model in secondary school, educators and policymakers can make more informed decisions about students' academic performance and tailor instructional strategies to meet their needs better. Moving forward, continued research into the effects of item parameters on ability estimation across different subjects and grade levels will be crucial for advancing assessment practices and promoting academic success in Nigerian schools.

REFERENCES

Adedoyin, O. O. and Adedoyin, J. (2013) 'Assessing the Comparability between Classical Test Theory (CTT) and Item Response Theory (IRT) Models in Estimating Test Item Parameters', *Herald Journal of Education and General Studies*, Vol. 2, No. 3, pp. 107–114. http://dx.doi.org/10.12691/education-6-3-11

Adetutu, P. O. and Iwintolu, R. O. (2017) 'An Item Response Theory Analysis of the Academic Amotivation Inventory for Secondary School Students in Southwestern Nigeria', *Journal of Research and Method in Education (IOSR-JRME),* Vol. 7, No. 4, pp. 22–31. https://dx.doi.org/10.9790/7388-0704032231

Alordiah, C. (2015) '*Comparison of Index of Differential Item Functioning under the Methods of Item Response Theory and Classical Test Theory in Mathematics'*, An unpublished Ph. D thesis of Delta State University, Abraka, Delta State, Nigeria.

Awopeju, O. and Afolabi, E. (2016) 'Comparative Analysis of Classical Test Theory and Item Response Theory based Item Parameter Estimates of Senior School Certificate Mathematics Examination'*, European Scientific Journal*, Vol. 12, No. 28, pp. 263–284. http://dx.doi.org/10.19044/esj.2016.v12n28p263

Ayanwale, M. A. (2023) 'Test Score Equating of Multiple-Choice Mathematics Items: Techniques from Characteristic Curve of Modern Psychometric Theory', *Discover Education*, Vol. 2, No. 1, p. 30. https://doi.org/10.1007/s44217-023-00052-z

Baker, F. B. (2001) *The basics of item response theory*, 2nd Edition, Available: https://files.eric.ed.gov/fulltext/ED458219.pdf [10 January 2024].

Bichi, A. A. and Talib, R. (2018) 'Item Response Theory: An Introduction to Latent Trait Models to Test and Item Development', *International Journal of Evaluation and Research in Education*, Vol. 7, No. 2, pp. 142–151. https://doi.org/10.11591/ijere.v7i2.12900

Breuer, S., Scherndl, T. and Ortner, T. M. (2023) "Effects of Response Format on Achievement and Aptitude Assessment Results: Multi-Level Random Effects Meta-Analyses', *Royal Society Open Science*, Vol. 10, No. 5, p. 220456. https://doi.org/doi:10.1098/rsos.220456

Choi, S., Jang, Y. and Kim, H. (2023) 'Influence of Pedagogical Beliefs and Perceived Trust on Teachers' Acceptance of Educational Artificial Intelligence Tools', *International Journal of Human–Computer Interaction*, Vol. 39, No. 4, pp. 910–922. https://doi.org/10.1080/10447318.2022.20491455

Danh, T., Desiderio, T., Herrmann, V., Lyons, H. M., Patrick, F., Wantuch, G. A. and Dell, K. A. (2020) 'Evaluating the Equality of Multiple-Choice Questions in a NAPLEX Preparation Book', *Currents in Pharmacy Teaching and Learning*, Vol. 12, No. 10, pp. 1188–1193. https://doi.org/10.1016/j.cptl.2020.05.006

Douglas, K. A., Neumann, K. and Oliveri, M. E. (2023) 'Contemporary Approaches to Assessment of Engineering Competencies for Diverse Learners', in Johri, A. (ed.) *International handbook of Engineering Education Research*, New York: Routledge. https://doi.org/10.4324/9781003287483-37

Ferreira-Junior, M., Reinaldo, J. T., Neto, E. A. L. and Prudencio, R. B. (2023) 'β4-IRT: A new β3-IRT with Enhanced Discrimination Estimation', ArXiv Preprint arXiv:2303.17731. https://doi.org/10.48550/arXiv.2303.17731

Frick, S., Krivosija, A. and Munteanu, A., (2024) 'Scalable Learning of Item Response Theory Models', *International Conference on Artificial Intelligence and Statistics (AISTATS),* London, pp. 1234–1242.

Gates, J. A. (2023) '*School Board Member Perceptions of the Race Based Academic Achievement Disparity* [Ed.D., Concordia University Chicago]', ProQuest Dissertations & Theses Global. United States -- Illinois. https://www.proquest.com/dissertations-theses/school-board-member-perceptions-race-based/docview/2910119963/se-2?accountid=13425

Jimoh, K., Opesemowo, O. A. and Faremi, Y. A. (2022) 'Psychometric Analysis of Senior Secondary School Certificate Examination (SSCE) 2017 Neco English Language Multiple Choice Test Items in Kwara State Using Item Response Theory', *Journal of Applied Research and Multidisciplinary Studies*, Vol. 3, No. 2, pp. 1–19. https://doi.org/10.32350/jarms.32.01

Kalhori, R. P. and Abbasi, M. (2017) 'Are Faculty Members of Paramedics Able to Designed Accurate Multiple Choice Questions', *Global Journal of Health Science*, Vol. 9, No. 1, pp. 211–216. http://dx.doi.org/10.5539/gjhs.v9n1p211

Kennedy, I. and Ebuwa, S. O. (2022) 'Assessing Score Dependability of West Africa Examinations Council (WAEC) 2019 Mathematics Objective Test using Generalisability Theory', *British Journal of Contemporary Education*, Vol. 2, No. 1, pp. 64–73. https://doi.org/10.52589/BJCE-OCA9OZJT

Kim, K. Y. (2017) '*IRT Linking Methods for the Bifactor Model: A Special Case of the Two-Tier Item Factor Analysis Model* [Ph.D., The University of Iowa]', ProQuest Dissertations & Theses Global. United States -- Iowa. https://www.proquest.com/dissertations-theses/irt-linking-methods-bifactor-model-special-case/docview/1964934014/se-2?accountid=13425

Mutiawani, V., Athaya, A. M., Saputra, K. and Subianto, M. (2022) 'Implementing Item Response Theory (IRT) Method in Quiz Assessment System', *TEM Journal*, Vol. 11, No. 1, pp. 210–218. https://doi.org/10.18421/TEM111-26

O'Connor, P. J., Hill, A. and Martin, B. (2019) 'The Measurement of Emotional Intelligence: A Critical Review of the Literature and Recommendations for Researchers and Practitioners', *Frontiers in Psychology*, Vol. 10, No. 1116, pp. 1–19. https://doi.org/10.3389/fpsyg.2019.01116

Oladele, J. I., Ndlovu, M. and Spangenberg, E. D. (2022) 'Simulated Computer Adaptive Testing Method Choices for Ability Estimation with Empirical Evidence', *International Journal of Evaluation and Research in Education (IJERE)* Vol. 3, pp. 1392–1399. https://doi.org/10.11591/ijere.v11i3.21986

Olagunju, B. A. and Iwintolu, R. O. (2023) 'Development of Pedagogical Competence Scale for Lecturers in Universities Using Item Response Theory', *Mimbar Sekolah Dasar*, Vol. 10, No.1, pp. 135–148. https://doi.org/10.53400/mimbar-sd.v10i1.51422

Opesemowo, O., Afolabi, E. and Oluwatimilehin, T. (2018) 'Development of a Scale for Measuring Students' Testwiseness in Senior Secondary School Examination in Nigerian', *International Journal of Research*, Vol. 5, No. 19, pp. 464–474.

Opesemowo, O. A. G., Ayanwale, M. A., Opesemowo, T. R. and Afolabi, E. R. I. (2023) 'Differential Bundle Functioning of National Examinations Council Mathematics Test Items: An Exploratory Structural Equation Modelling Approach', *Journal of Measurement and Evaluation in Education and Psychology*, Vol. 14, No. 1, pp. 1–18. https://doi.org/10.21031/epod.1142713

Opesemowo, O. A. G. and Ndlovu, M. (2023) 'Status and Experience of Mathematics Teachers' Perception of Integrating Computer Adaptive Testing into Unified Tertiary Matriculation Examination Mathematics', *Multicultural Education*, Vol. 9, No. 2, pp. 66–78.

Paek, I., Lin, Z., and Chalmers, R. P. (2023) 'Investigating Confidence Intervals of Item Parameters When Some Item Parameters Take Priors in the 2PL and 3PL Models', *Educational and Psychological Measurement*, Vol. 83, No. 2, pp. 375–400. https://doi.org/10.1177/00131644221096431

Perez, A. L. and Loken, E. (2023) 'Person Specific Parameter Heterogeneity in the 2PL IRT Model', *Multivariate Behavioral Research*, pp. 1–7. https://doi.org/10.1080/00273171.2023.2224312

Powers, K. (2019) '*Personality, Attitudes, and Behaviors*', in Workplace Psychology, Chemeketa Community College.

Rafi, I., Retnawati, H., Apino, E., Hadiana, D., Lydiati, I. and Rosyada, M. N. (2023) 'What Might Be Frequently Overlooked Is Actually Still Beneficial: Learning from Post National-Standardized School Examination', *Pedagogical Research*, Vol. 8, No. 1, pp. 1–15. https://doi.org/10.29333/pr/12657

Rios, J. A. and Soland, J. (2022) 'An Investigation of Item, Examinee, and Country Correlates of Rapid Guessing in PISA', *International Journal of Testing*, Vol. 22, No. 2, pp. 154–184. https://doi.org/10.1080/15305058.2022.2036161

Rudner, L. M. (2019) 'Scoring and Classifying Examinees using Measurement Decision Theory', *Practical Assessment, Research, and Evaluation*, Vol. 14, pp. 1–12. https://doi.org/10.7275/vksg-rh07

Scheibling-Sève, C., Pasquinelli, E. and Sander, E. (2020) 'Assessing Conceptual Knowledge through Solving Arithmetic Word Problems', *Educational Studies in Mathematics*, Vol. 10*3,* No. 3, 293–311. https://doi.org/10.1007/s10649-020-09938-3

Setiawati, F. A., Amelia, R. N., Sumintono, B. and Purwanta, E. (2023) 'Study Item Parameters of Classical and Modern Theory of Differential Aptitude Test: is it Comparable?', *European Journal of Educational Research*, Vol. 12, No. 2, pp. 1097–1107. https://doi.org/10.12973/eu-jer.12.2.1097

Sideridis, G. and Alahmadi, M. (2022) 'Estimation of Person Ability under Rapid and Effortful Responding', *Journal of Intelligence*, Vol. 10, No. 3, p. 67. https://www.mdpi.com/2079-3200/10/3/67

Stemler, S. E. and Naples, A. (2021) 'Rasch Measurement v. Item Response Theory: knowing when to cross the line', *Practical Assessment, Research and Evaluation*, Vol. 26, No. 11, pp. 1–16. https://doi.org/10.7275/v2gd-4441

Suparman, S. and Juandi, D. (2022) 'Self-efficacy and Mathematical Ability: A Meta-Analysis of Studies conducted in Indonesia', *Pedagogika/Pedagogy*, Vol. 147, No. 3, pp. 26–57. https://doi.org/10.15823/p.2022.147.2

Sweeney, S. M., Sinharay, S., Johnson, M. S. and Steinhauer, E. W. (2022) 'An Investigation of the Nature and Consequence of the Relationship between IRT Difficulty and Discrimination', *Educational Measurement: Issues and Practice*, Vol. 41, No. 4, pp. 50–67. https://doi.org/10.1111/emip.12522

von Davier, M. (2019) 'TIMSS 2019 Scaling Methodology: Item Response Theory, Population Models, and linking across Modes', *Methods and procedures: TIMSS*, Vol. 11, No. 11, pp. 11–25.

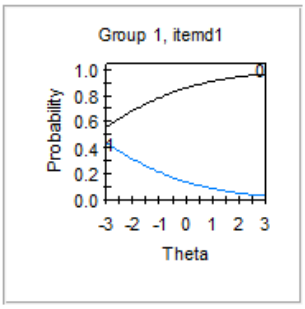
Yustiandi, Y. and Saepuzaman, D. (2021) 'Analysis of Model Fit and Item Parameter of Work and Energy Test using Item Response Theory', *Gravity: Jurnal Ilmiah Penelitian dan Pembelajaran Fisika*, Vol. 7, No, 2, pp. 21–28. https://doi.org/10.30870/gravity.v7i2.10563

Zanon, C., Hutz, C. S., Yoo, H. and Hambleton, R. K. (2016) 'An Application of Item Response Theory to Psychological Test Development', *Psicologia: Reflexão e Crítica*, Vol. 29, No. 18. https://doi.org/10.1186/s41155-016-0040-x

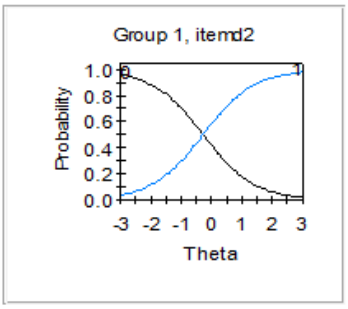
APPENDIX

APPENDIX A

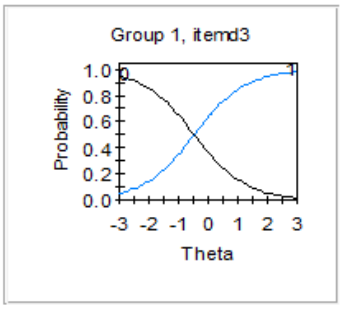
MAT Item Category Curves of 2-PL



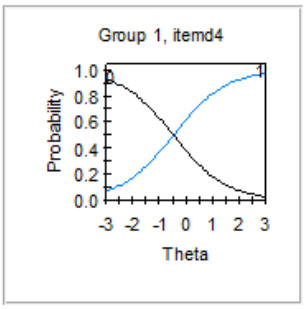
a = -0.52 b = -3.53 M = 0.15



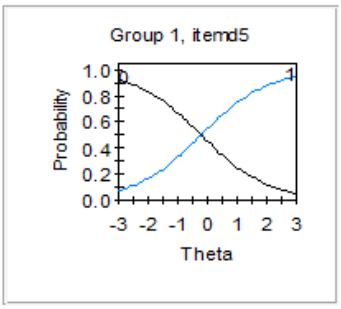
a = 1.21 b = -0.28 M = 0.61



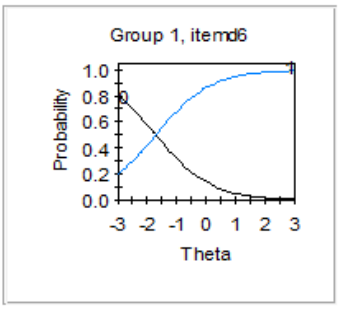
a= 1.18 b = -0.47 M = 0.31



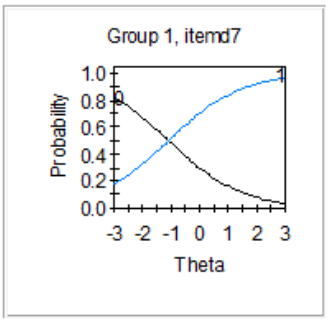
a = 1.03 b = -0.47 M = 0.31



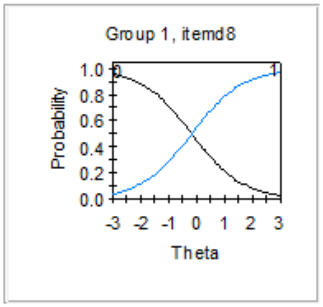
a = 0.91 b = -0.21 M = 0.57



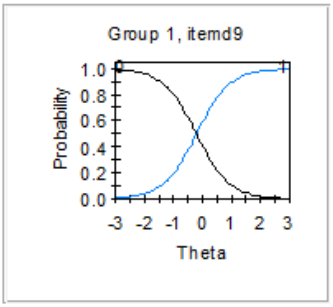
a = 1.09 b = -1.71 M = 0.82



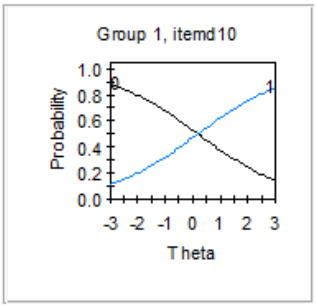
a = 0.79 b = -1.10 M = 0.74



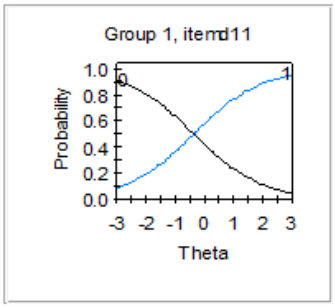
a = 1.12 b = -0.17 M = 0.33



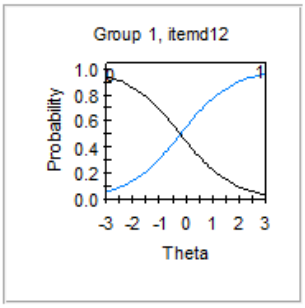
a = 1.79 b =-0.21 M=0.55



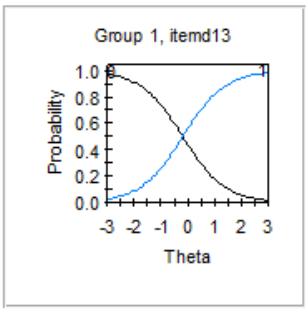
a = 0.63 b = 0.18 M = 0.47



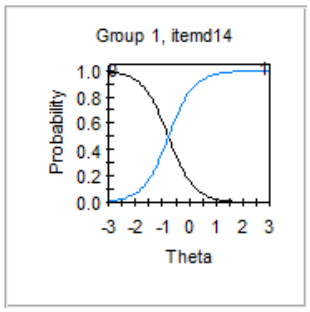
a = 0.89 b = -0.36 M = 0.62



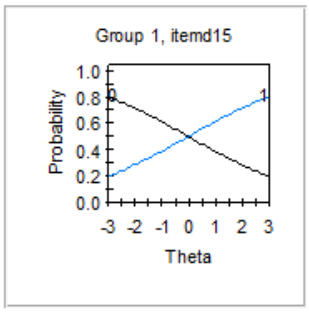
a = 1.02 b =-0.21 M = 0.54



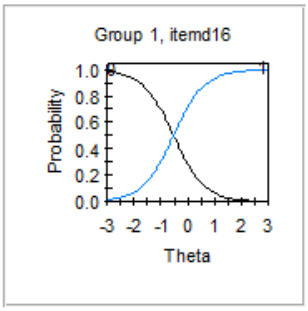
$a = 1.31$ $b = -18$ $M = 0.39$



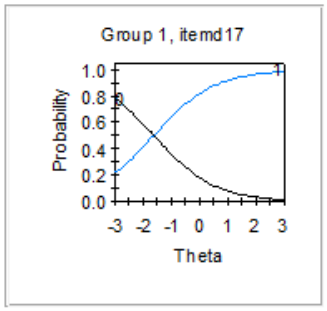
$a = 2.14$ $b = -0.77$ $M = 0.72$



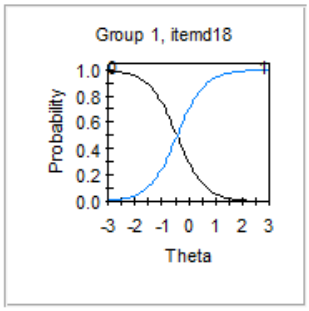
$a = 0.46$ $b = -0.02$ $M = 0.50$



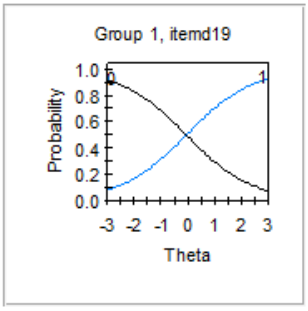
$a = 1.78$ $b = -0.53$ $M = 0.64$



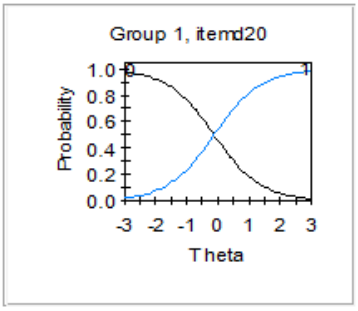
$a = 0.94$ $b = -1.62$ $M = 0.13$



$a = 1.97$ $b = -0.44$ $M = 0.30$



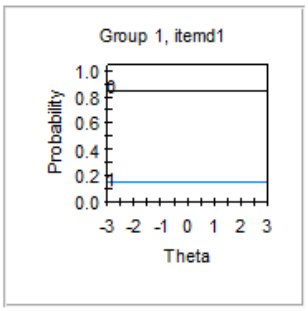
$a = 0.82$ $b = -0.05$ $M = 0.33$



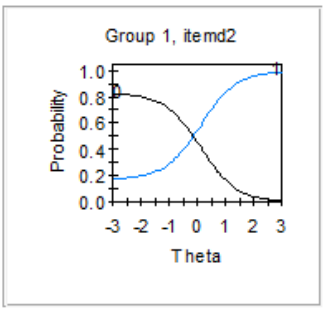
$a = 1.35$ $b = -0.12$ $M = 0.53$

APPENDIX B

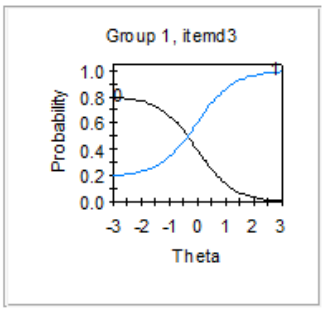
MAT Item Category Curves of 3-PL



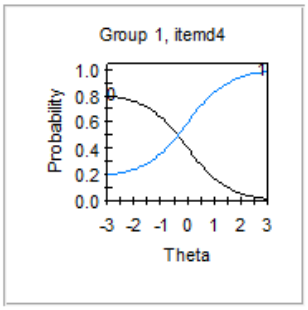
$a = 0.37$ $b = 467.72$ $c = -171.99$



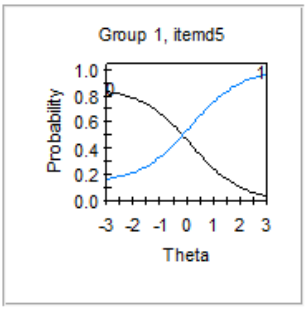
$a = 1.61$ $b = 0.12$ $c = -0.19$



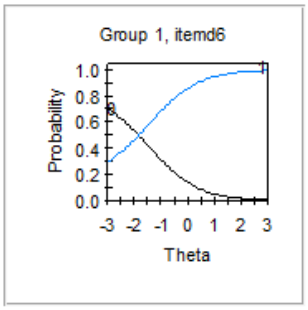
$a = 1.55$ $b = -0.02$ $c = 0.03$



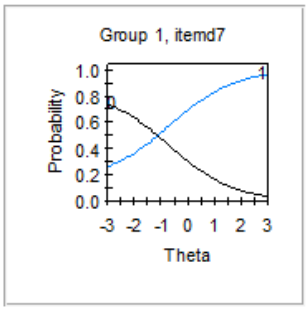
$a = 1.32$ $b = -0.01$ $c = 0.01$



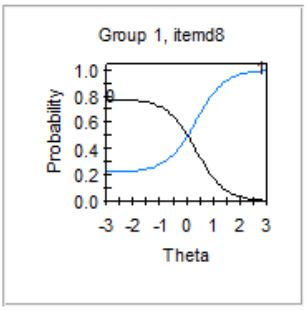
$a = 1.09$ $b = 0.17$ $c = -0.19$



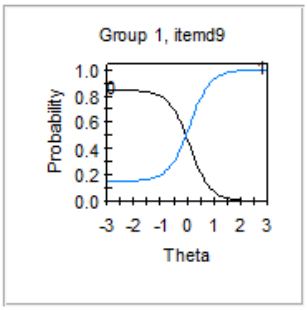
$a = 1.11$ $b = -1.44$ $c = 1.60$



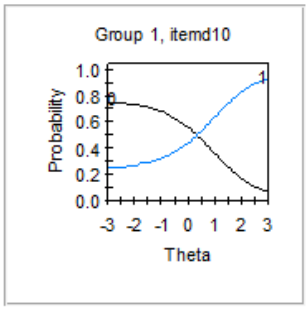
$a = 0.85$ $b = -0.68$ $c = 0.58$



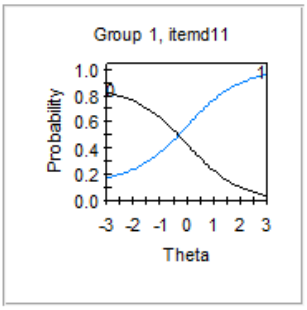
$a = 1.91$ $b = 0.36$ $c = -0.69$



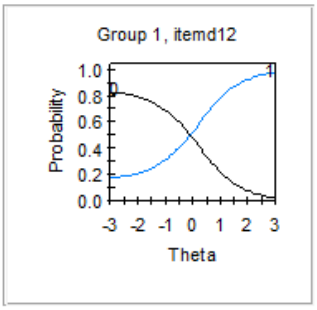
$a = 2.66$ $b = 0.10$ $c = -0.27$



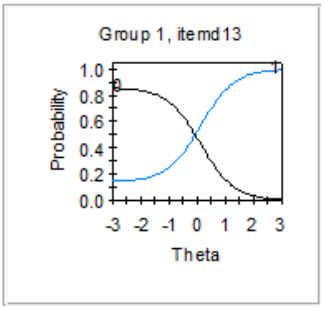
$a = 1.15$ $b = 0.93$ $c = -1.02$



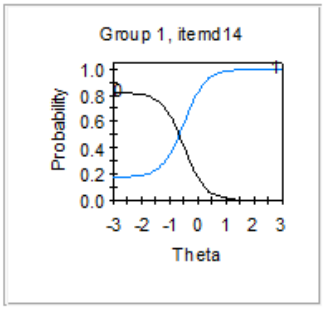
$a = 1.00$ $b = 0.01$ $c = -0.01$



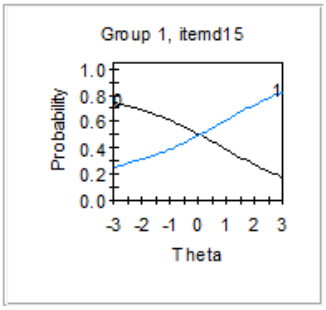
$a = 1.30$ $b = 0.20$ $c = -0.26$



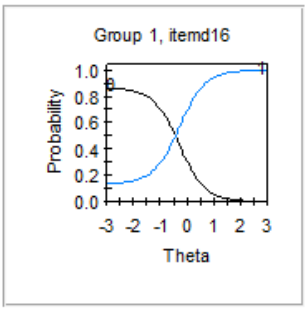
$a = 1.73$ $b = 0.13$ $c = -0.23$



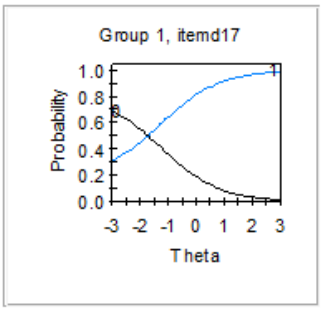
$a = 2.68$ $b = -0.48$ $c = 1.30$



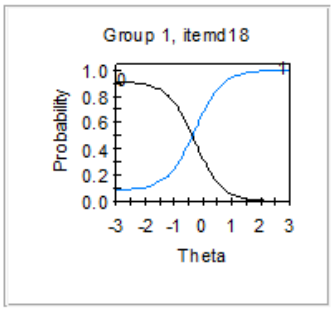
$a = 0.60$ $b = 0.79$ $c = -0.47$



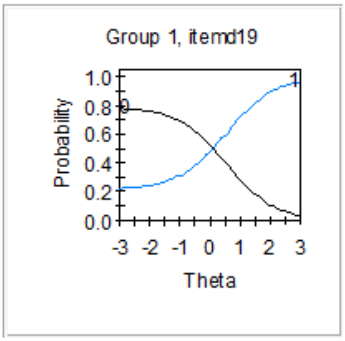
$a = 2.12$ $b = -0.28$ $c = 0.59$



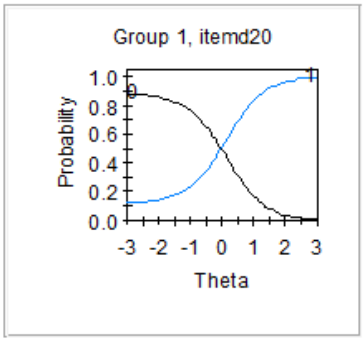
$a = 0.98$ $b = -1.23$ $c = 1.21$



$a = 2.16$ $b = -0.28$ $c = 0.60$



$a = 1.26$ $b = 0.55$ $c = -0.69$



$a = 1.67$ $b = 0.14$ $c = -0.24$