

Statistica Applicata

ISSN: 1125-1964

Volume 38
Number 1
January - April 2026

ITALIAN
JOURNAL
OF APPLIED
STATISTICS

FI
FIRENZE
UNIVERSITY
PRESS

Statistica Applicata – ITALIAN JOURNAL OF APPLIED STATISTICS

EDITORIAL TEAM

EDITORIAL TEAM

EDITOR IN CHIEF

Francesco Palumbo, Università di Napoli Federico II, Naples, Italy

<https://orcid.org/0000-0002-9027-5053>

fpalumbo@unina.it

CO-EDITORS ON A SPECIFIC SUBJECT

Co-editor Studies in Behavioural and Health Sciences

Luigi Fabbri, Università di Padova, Padua, Italy

<https://orcid.org/0000-0001-8657-8361>

luigi.fabbri@unipd.it

Co-editor Studies in Studies in Business, Industry and Economics

Paolo Mariani, Università di Milano Bicocca, Milan, Italy

<https://orcid.org/0000-0002-8848-8893>

paolo.mariani@unimib.it

SCIENTIFIC COMMITTEE

Irene D'Epifanio, Universitat Jaume I, Castelló de la Plana, Spain

<https://orcid.org/0000-0002-6973-311X>

epifanio@uji.es

Vincenzo Esposito Vinzi, ESSEC Paris, France

<https://orcid.org/0000-0001-8772-042X>

vinzi@essec.edu

Michael J. Greenacre, UPF, Barcelona, Spain

<https://orcid.org/0000-0002-0054-3131>

michael.greenacre@upf.edu

Salvatore Ingrassia, Università di Catania, Catania, Italy
<https://orcid.org/0000-0003-2052-4226>
salvatore.ingrassia@unict.it

Ron S. Kenett, KPA Ltd. and Samuel Neaman Institute, Technion, Haifa, Israel
<https://orcid.org/0000-0003-2315-0477>
ron@kpa-group.com

Eric Marlier, LISER, Luxembourg Institute of Socio-Economic Research
Université de Luxembourg, Luxembourg
<https://orcid.org/0000-0002-5559-3689>
eric.marlier@liser.lu

Stefania Mignani, Università di Bologna Alma Mater, Bologna, Italy
<https://orcid.org/0000-0003-4746-1130>
stefania.mignani@unibo.it

Tormod Naes, NOFIMA, Oslo, Norway
<https://orcid.org/0000-0001-5610-3955>
tormod.naes@nofima.no

Alessandra Petrucci, Università di Firenze, Florence, Italy
<https://orcid.org/0000-0001-9952-0396>
alessandra.petrucci@unifi.it

Monica Pratesi, Università di Pisa, Pisa, Italy
<https://orcid.org/0000-0002-9959-5483>
monica.pratesi@unipi.it

Maurizio Vichi, Sapienza Università di Roma, Rome, Italy
<https://orcid.org/0000-0002-3876-444X>
maurizio.Vichi@uniroma1.it

Giorgio Vittadini, Università di Milano Bicocca, Milan, Italy
<https://orcid.org/0000-0002-0375-9763>
giorgio.vittadini@unimib.it

Adalbert Wilhelm, Constructor University Bremen, Bremen, Germany
<https://orcid.org/0000-0001-6160-1596>
awilhelm@constructor.university

ASSOCIATE EDITORS

Ali Abuzaid, Department of Mathematics, Al Azhar University - Gaza, Gaza, Palestine
<https://orcid.org/0000-0002-6680-7371>
alizaid33@yahoo.com

Luca Bagnato, Università Cattolica del Sacro Cuore, Milan, Italy
<https://orcid.org/0000-0003-3129-5385>
luca.bagnato@unicatt.it

Francesca Bassi, Università di Padova, Padua, Italy
<https://orcid.org/0000-0002-3257-7029>
francesca.bassi@unipd.it

Rossella Berni, Università di Firenze, Florence, Italy
<https://orcid.org/0000-0002-7782-1777>
rossella.berni@unifi.it

Paolo Berta, Università di Milano Bicocca, Milan, Italy
<https://orcid.org/0000-0003-0984-4288>
paolo.bera@unimib.it

Bruno Bertaccini, Università di Firenze, Florence, Italy
<https://orcid.org/0000-0002-5816-2964>
bruno.bertaccini@unifi.it

Matilde Bini, Università Europea, Rome, Italy
<https://orcid.org/0000-0003-1989-8685>
matilde.bini@unier.it

Giovanna Boccuzzo, Università di di Padova, Padua, Italy
<https://orcid.org/0000-0003-2143-7730>
giovanna.boccuzzo@unipd.it

John Castura, Compusense Inc., Guelph, Canada
<https://orcid.org/0000-0002-1640-833X>
jcastura@compusense.com

Maurizio Carpita, Università di di Brescia, Brescia, Italy
<https://orcid.org/0000-0001-7998-5102>
maurizio.carpita@unibs.it

Carlo Cavicchia, Erasmus Rotterdam University, Netherlands
<https://orcid.org/0000-0003-1816-3521>
cavicchia@ese.eur.nl

Carolina Chaya, Universitat Politecnica de Madrid, Spain
<https://orcid.org/0000-0002-5518-886X>
carolina.chaya@upm.es

Alessandro Celegato, AICQ Centronord - Quality and Technology in production
alessandro.celegato@gmail.com

Giuliana Coccia, ASVIS, Rome, Italy
giuliana.coccia1@gmail.com

Fabio Crescenzi, Independent researcher, Italy
<https://orcid.org/0000-0002-2660-6209>
fabio7826@gmail.com

Cristina Davino, Università di Napoli Federico II, Naples, Italy
<https://orcid.org/0000-0003-1154-4209>
cristina.davino@unina.it

Tonio Di Battista, Università di Chieti-Pescara “Gabriele D’Annunzio”, Pescara, Italy
<https://orcid.org/0000-0003-2139-7273>
tonio.dibattista@unich.it

Francesca Di Iorio, Università di Napoli Federico II, Naples, Italy
<https://orcid.org/0000-0002-9586-3380>
francesca.diiorio@unina.it

Simone Di Zio, Università di Chieti-Pescara “Gabriele D’Annunzio”, Pescara, Italy
<https://orcid.org/0000-0002-9139-1451>
simone.dizio@unich.it

Filippo Domma, Università della Calabria, Rende, Italy
<https://orcid.org/0000-0002-1489-1065>
f.domma@unical.it

Angela M. D’Uggento, Università di Bari, Bari, Italy
<https://orcid.org/0000-0001-9768-651X>
angelamaria.duggento@uniba.it

Monica Ferraroni, Università di Milano, Milan, Italy
<https://orcid.org/0000-0002-4542-4996>
monica.ferraroni@unimi.it

Livio Finos, Università di Padova, Padova, Italy
<https://orcid.org/0000-0003-3181-8078>
livio.finos@unipd.it

Giuseppe Giordano, Università di Salerno, Salerno, Italy
<https://orcid.org/0000-0002-0582-109X>
ggiordano@unisa.it

Michela Gnaldi, Università di Perugia, Perugia, Italy
<https://orcid.org/0000-0002-2785-3279>
michela.gnaldi@unipg.it

Paolo Girardi, Università Ca' Foscari, Venezia, Italy
<https://orcid.org/0000-0001-8330-9414>
paolo.girardi@unive.it

Maria Gabriella Grassia, Università degli Studi di Napoli Federico II, Naples, Italy
<https://orcid.org/0000-0002-7128-7323>
mariagabriella.grassia@unina.it

Domenica Fioredistella Iezzi, Università di Roma Tor Vergata, Rome, Italy
<https://orcid.org/0000-0001-8041-2846>
stella.iezzi@uniroma2.it

Giuseppe Lamberti, Universitat Autònoma de Barcelona, Spain
<https://orcid.org/0000-0002-8666-796X>
giuseppe.lamberti@uab.cat

Filomena Maggino, Sapienza Università di Roma, Italy
<https://orcid.org/0000-0002-9071-9750>
filomena.maggino@uniroma1.it

Angelos Markos, Democritus University of Thrace, Greece
<https://orcid.org/0000-0002-4204-3573>
amarkos@eled.duth.gr

Maria Cristina Martini, Università di Modena e Reggio Emilia, Modena, Italy
<https://orcid.org/0000-0001-5622-9187>
mariacristiana.martini@unimore.it

Fulvia Mecatti, Università di Milano Bicocca, Milan, Italy
<https://orcid.org/0000-0002-0403-9231>
fulvia.mecatti@unimib.it

Isabella Morlini, Università di Modena e Reggio Emilia, Modena, Italy
<https://orcid.org/0000-0002-2900-2742>
isabella.morlini@unimore.it

Biagio Palumbo, Università di Napoli Federico II, Naples, Italy
<https://orcid.org/0000-0003-1036-8127>
bpalumbo@unina.it

Antonio Punzo, Università di Catania, Catania, Italy
<https://orcid.org/0000-0001-7742-1821>
antonio.punzo@unict.it

Rosaria Romano, Università di Napoli Federico II, Naples, Italy
<https://orcid.org/0000-0002-9708-1753>
rosaroma@unina.it

Rosaria Simone, Università di Napoli Federico II, Naples, Italy
<https://orcid.org/0000-0002-6844-6418>
rosaria.simone@unina.it

Silvia Salini Università di Milano, Milan, Italy
<https://orcid.org/0000-0001-6106-9835>
silvia.salini@unimi.it

Luigi Salmaso, Università di di Padova, Padua, Italy
<https://orcid.org/0000-0001-6501-1585>
luigi.salmaso@unipd.it

Germana Scepi, Università di Napoli Federico II, Naples, Italy
<https://orcid.org/0000-0001-8565-0085>
germana.scepi@unina.it

Giorgio Tassinari, Università di Bologna Alma Mater, Bologna, Italy
<https://orcid.org/0000-0002-5161-7989>
giorgio.tassinari@unibo.it

Laura Trinchera, Neoma Business School, Paris, France
<https://orcid.org/0000-0001-9679-0956>
laura.trinchera@neoma-bs.fr

Rosanna Verde, Università della Campania “Luigi Vanvitelli”, Caserta, Italy
<https://orcid.org/0000-0002-9959-4675>
rosanna.verde@unicampania.it

Maria Prosperina Vitale, Università di Salerno, Salerno, Italy
<https://orcid.org/0000-0003-2735-7029>
mvitale@unisa.it

Susanna Zaccarin, Università di Trieste, Trieste, Italy
<https://orcid.org/0000-0002-3595-0407>
susanna.zaccarin@units.it

Emma Zavarrone, IULM Milano, Milan, Italy
<https://orcid.org/0000-0001-9509-8773>
emma.zavarrone@iulm.it
Editorial staff

Rosa Fabbricatore, Università di Napoli Federico II, Naples, Italy
<https://orcid.org/0000-0002-4056-4375>
rosa.fabbricatore@unina.it

Alessandro Magrini, Università di Firenze, Florence, Italy
<https://orcid.org/0000-0002-7278-5332>
alessandro.magrini@unifi.it

Lucio Palazzo, Università di Napoli L’Orientale, Naples, Italy
<https://orcid.org/0000-0001-7529-4689>
lucio.palazzo@unior.it

EDITORIAL MANAGERS

Andrea Marletta, Università di Firenze, Florence, Italy

<https://orcid.org/0000-0002-4050-5316>

andrea.marletta@unimib.it

Domenico Vistocco, Università di Napoli Federico II, Naples, Italy

<https://orcid.org/0000-0002-8541-6755>

domenico.vistocco@unina.it

A.S.A CONTACTS**Principal Contact**

Francesco Palumbo (Editor in Chief)

editor@sa-ijas.org

Support Contact

Domenico Vistocco (Editorial Manager)

ijas@sa-ijas.org

JOURNAL WEBPAGE

<https://www.sa-ijas.org/ojs/index.php/sa-ijas>

Statistica Applicata – Italian Journal of Applied Statistics is associated to the following Italian and international journals:

QTQM – Quality Technology & Quantitative Management (<http://web.it.nctu.edu/~qtqm/>)

SINERGIE – Italian Journal of Management

Statistica Applicata – Italian Journal of Applied Statistics (ISSN:1125-1964, E-ISSN:2038-5587) applies the Creative Commons Attribution (CC BY) license to everything we publish.



Published: April 2026

© 2025 Author(s)

Content license: except where otherwise noted, the present work is released under Creative Commons Attribution 4.0 International license (CC BY 4.0: <https://creativecommons.org/licenses/by/4.0/legalcode>). This license allows you to share any part of the work by any means and format, modify it for any purpose, including commercial, as long as appropriate credit is given to the author, any changes made to the work are indicated and a URL link is provided to the license.

Metadata license: all the metadata are released under the Public Domain Dedication license (CC0 1.0 Universal: <https://creativecommons.org/publicdomain/zero/1.0/legalcode>).

Published by Firenze University Press

Powered by Firenze University Press

Firenze University Press

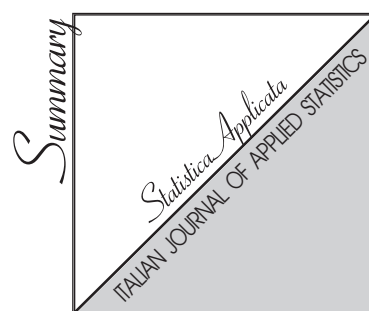
Università degli Studi di Firenze

via Cittadella, 7, 50144 Firenze, Italy

www.fupress.com

Statistica Applicata – Italian Journal of Applied Statistics is a four-monthly journal published by the Associazione per la Statistica Applicata (A.S.A.), Largo Gemelli 1 – 20123 Milano, Italy (phone + 39 02 72342904). Advertising: CLEUP SC, via G. Belzoni, 118/3 – 35128 Padova, Italy (phone +39 049 8753496 – Fax +39 049 9865390), email: info@cleup.it.

Rules for manuscript submission: <https://www.sa-ijas.org/ojs/index.php/sa-ijas/about/submissions>
Subscription: yearly €103.30; single copy €40.00; A.S.A. associates €60.00; supporting institutions: €350.00. Advertisement lower than 70%. Postal subscription Group IV, Milan. Forum licence n. 782/89. CLEUP SC on behalf of ASA, 7 March 2023.



Vol. 38, Number 1

- | | | |
|-----|---|---|
| 13 | <i>Testa, I., Galano, S., Palazzo, L., Ragozzini, G., Boone, W.J.</i> | <i>Thematic issue on “Measurement in stem education” – Editorial</i> |
| 17 | <i>Brancaccio, N., Fraudatario, M.C., Pacca, A., Ragozzini, G., Restaino, M., Vitale M.P.</i> | <i>Personal networks, schools, and higher education: Linking social capital and student pathways beyond high school</i> |
| 57 | <i>Giannantoni, P., Falzetti, P.</i> | <i>Equal opportunities to “count”? A longitudinal population study about choice of stem at university by gender</i> |
| 79 | <i>Catena, D., Testa, I.</i> | <i>Investigating undergraduate and high school students. Identification with physics through structural equation modeling</i> |
| 103 | <i>Bozzi, M., Mazzola, R., Zani, M.</i> | <i>Peer learning in large size format: Is it effective in bolstering problem-solving and metacognitive skills?</i> |
| 131 | <i>Fabbricatore, R., Parola, A., Pepicelli, G., Palumbo, F.</i> | <i>Academic procrastination scale: Psychometric evidence for the Italian version based on CFA and IRT models</i> |

THEMATIC ISSUE ON “MEASUREMENT IN STEM EDUCATION” – EDITORIAL

Italo Testa, Silvia Galano

“E. Pancini” Department of Physics, Federico II University of Naples, Naples, Italy

Lucio Palazzo

Department of Social Sciences and Humanities, L’Orientale University of Naples, Naples, Italy

Giancarlo Ragozini

Department of Political Sciences, Federico II University of Naples, Naples, Italy

William J. Boone

Miami University, Oxford, Ohio, United States

Quantitative methods are indispensable in educational research thanks to their ability to provide objective, reliable and generalizable findings. By employing suitable research designs and statistical analyses, researchers can gain valuable insights into educational processes and outcomes. Over the past twenty years, there has been a growing interest in addressing three broad issues in the field of Science-Technology-Engineering-Mathematics (STEM) education:

- Students’ decreasing interest in pursuing STEM-related careers
- Gender disparity in participation in STEM activities and careers
- Students’ persistence in and dropout from STEM undergraduate courses.

To address these issues within this wider framework, researchers need to analyze several cognitive and non-cognitive constructs that interact with each other in complex ways. As the complexity of the processes involved increases, new methodological and data-driven perspectives are required to provide robust and reliable results.

The aim of this thematic issue is to contribute to the ongoing discussion by presenting original papers that cover the three aforementioned issues and utilize a wide range of statistical methods, including logistic regression, item response theory, structural equation modelling and cluster analysis.

This volume contains five papers that originated from the first Measurement in STEM Education (MESE1) conference, which took place in Naples from 30 January to 1 February 2023. The conference featured seven invited lectures, five contributed and invited symposia, and two workshops.

The first paper (Brancaccio et al., this issue) addresses the issue of declining interest in STEM-related careers by analyzing the role of personal networks — including family, friends, classmates and acquaintances — in shaping students' university degree and course choices. Particular attention is given to the interplay between personal, familiar and social dimensions. Using cluster analysis, the authors identified three profiles ranging from small, inward-looking networks with restricted social capital (Cluster 1) to large, well-connected networks rich in relational resources (Cluster 3).

The second paper (Giannantoni and Falzetti, this issue) and the third paper (Catena and Testa, this issue) mainly address gender disparity in STEM. In particular, the paper by Giannantoni and Falzetti examines the factors influencing boys' and girls' decisions to enroll in STEM subjects, focusing on mathematical ability, as well as individual and family characteristics. Using logistic regression, the authors found that mathematical ability plays a central role in the choice of STEM studies and that a mother's level of education significantly influences girls' intention to enroll in STEM courses. The paper by Catena and Testa examines gender differences in the physics identity of a sample of high school and STEM undergraduate students. Using structural equation modelling, the authors found that gender differences in the dimensions of

the physics identity model are more pronounced among students who have chosen to study physics as an extracurricular activity or as an undergraduate course.

The fourth paper (Bozzi et al., this issue) addresses the third issue: how to encourage first-year students to persist with STEM courses, such as engineering, by implementing a peer-learning approach. The authors found that this approach was effective in improving students' performance in physics and promoting problem-solving skills and conceptual understanding.

Finally, the fifth paper (Fabbricatore et al., this issue) addresses the broader issue of university student dropout by studying academic procrastination. Using confirmatory factor analysis and item response theory, the authors validated a shortened version of a scale to measure this construct. They found that the instrument is a valid and reliable measure of procrastination, demonstrating measurement invariance with respect to gender.

This collection of papers demonstrates that integrating advanced statistical techniques and mixed-methods approaches can enhance the robustness and applicability of STEM educational research.

We thank the Association for Applied Statistics (www.sa-ijas.org) for hosting these papers in its official journal, along with the authors and the referees for making this thematic issue possible.

The Guest Editors

William J. Boone
Silvia Galano
Lucio Palazzo
Giancarlo Ragozini
Italo Testa

PERSONAL NETWORKS, SCHOOLS, AND HIGHER EDUCATION: LINKING SOCIAL CAPITAL AND STUDENT PATHWAYS BEYOND HIGH SCHOOL

Nunzia Brancaccio

Department of Social Sciences, Federico II University of Naples, Naples, Italy

Maria Camilla Fraudatario

National Research Council IRPPS, Italy

Angela Pacca

*"Giuseppe Parenti" Department of Statistics, Computer Science, Applications,
University of Florence, Florence, Italy*

Giancarlo Ragozini

Department of Political Sciences, Federico II University of Naples, Naples, Italy

Marialuisa Restaino

Department of Economics and Statistics, University of Salerno, Salerno, Italy

Maria Prosperina Vitale¹

Department of Political and Social Studies, University of Salerno, Salerno, Italy

Abstract

Understanding the factors that shape students' university choices is essential for analysing educational transitions. Previous research highlights the role of economic, social, and marketing factors, with key influences including personal characteristics, family background, school context, and peers. While much of the literature focuses on peer effects on academic performance, less attention has been paid to how social networks affect decisions about higher education. The paper explores the role of personal networks – comprising family, friends, classmates, and acquaintances – in shaping students' university choices, with particular attention to the interplay between personal, familial, and social dimensions. By conceptualising these networks as a form of social capital, the paper examines how network composition and structure influence students'

¹ Corresponding author (mvitale@unisa.it)

access to support, informational, emotional, and appraisal resources. The findings show that students embedded in highly interconnected networks are better positioned to access both informational and emotional support, which plays a key role in educational transitions. In contrast, students with weaker personal networks face challenges in making informed decisions due to limited access to relevant resources. The study also highlights the importance of both network size and structural diversity in strengthening social capital and facilitating informed decision-making during the transition from high school to university.

Keywords: *Higher education, Peer effects, Egocentric network data, Social support, Soft skills*

1. INTRODUCTION

The transition from high school to university represents a crucial point in educational trajectories, involving significant life challenges for individuals (Mpamhanga et al., 2025). Empirical studies show that university choices involve economic and social dimensions as well as personal decision-making processes. Among others, family and school context (Brooks, 2003; Sandefur et al., 2006; Engberg and Wolniak, 2010; Vandell et al., 2021; Chaudhry et al., 2024), universities' reputation and economic attractiveness (Reay et al., 2001; Bacci and Bertaccini, 2021; Columbu et al., 2021a; Biagi et al., 2024; Priulla et al., 2025), the geographical accessibility of the higher education system (Sá et al., 2006), and other characteristics of the academic environment (Kallio et al., 1995) are recognised as the main factors affecting the transition from high school to university.

Scholars have focused on peer effects in education, highlighting how social support and interactions among adolescents in primary and secondary schools influence student attainment, aspirations, and school-related choices (Rosenqvist, 2018; Raabe and Wölfer, 2019; Raabe et al., 2019; Smith, 2023). In the context of higher education, this line of research has mainly examined how interactions among peers influence individual outcomes, such as well-being, motivation, and academic achievement (Lomi et al., 2011; DeLay et al., 2016; Poldin et al., 2016; Mamas, 2018; Stadtfeld et al., 2019; Vörös et al., 2021; Brouwer et al., 2022; Chuard et al., 2025). To the best of our knowledge, studies investigating peer influences on students' educational choices after high school have received limited attention (Usala et al., 2023).

Research shows that individuals embedded in robust networks tend to exhibit higher performance, stronger social integration, and an overall better quality of life during university (Wilcox et al., 2005; Chaudhry et al., 2024). Socioeconomic background also favours the development of networks in terms of cultural and social capital (Lareau and Cox, 2011; Lareau, 2015) or, conversely, limits access to relational resources, thereby exacerbating existing social inequalities (Walpole, 2003; Pascarella et al., 2004). Institutional guidance programmes can also represent a relevant policy intervention to bridge social capital gaps and support student choices (Upcraft et al., 2005; al.nawas et al., 2015), encouraging the transition to university and mitigating socioeconomic disparities. Hence, identifying the main factors driving students' university choices is essential for developing recruitment strategies and policies in higher education settings.

Moving from this scenario, the purpose of the present contribution is to explore factors affecting students' university decision-making process, characterised by the interplay of individual aspirations and social interactions within family, school, and social environments. Specifically, the multifaceted role of personal, or egocentric, networks (Perry et al., 2018) as a form of social capital (Coleman, 1988; Burt, 2000; Molina et al., 2020) is examined, taking into account their interrelation with socioeconomic and familial backgrounds. Egocentric networks – encompassing family, teachers, and classmates at school, as well as friends and acquaintances in social contexts (Brooks, 2003) – can provide informational, emotional, appraisal, and instrumental support (Malecki et al., 2003), facilitating access to university for low-income individuals (Stanton-Salazar, 2011) and reducing educational inequalities in the transition from high school to university.

An integrated data collection approach, combining a questionnaire with an egocentric network-based research design (Robins, 2015; McCarty et al., 2019; Marsden and Hollstein, 2023), was adopted in a sample of students attending high schools in Southern Italy. The main purpose of the present study was to investigate how different personal networks support students' decision-making process in higher education, focusing on the role of classmates at school – recognised as peers – in providing support beyond that of family members, friends, and acquaintances. Students with limited resources in their networks may experience lower transition rates to university.

First, the composition and structure of egocentric network measures (Krenz et al., 2025) were computed to describe network characteristics and then further analysed through multidimensional exploratory techniques (Lebart et al., 1995),

including principal component analysis (PCA) and hierarchical cluster analysis (HCA), to identify the main individual and network characteristics related to this key educational transition. Additionally, to better illustrate the results of groups derived from HCA, examples of egocentric networks for each cluster were selected to provide a qualitative description in terms of sociodemographic characteristics and network structures of selected egos. Finally, the association between cluster results and students' university choices highlights differences in educational transitions in the presence of varying social support resources.

The remainder of the paper is organised as follows. Section 2 reports the theoretical background, focusing on peer effects on educational choices. Section 3 describes the data collection process and network design. Section 4 presents the main characteristics of egocentric networks, and Section 5 discusses the findings from the factorial analysis. Section 6 provides a brief discussion, outlines the study's limitations, considers policy implications, and suggests directions for future research.

2. THEORETICAL FRAMEWORK

The study of peer effects in educational contexts has developed within various theoretical frameworks. Parsonsian functionalism (Parsons, 1959) emphasised the role of schools as primary sites of socialisation, highlighting the importance of institutional norms and expectations in shaping students' trajectories. Similarly, symbolic interactionism has provided critical insights into how the meanings, identities, and cultures of peer groups influence educational pathways. From this perspective, Willis's (1977) study demonstrated how youth cultures contribute to the reproduction of educational inequalities. Meanwhile, Corsaro's (1985; 2005) studies showed that shared norms and meanings are actively constructed and negotiated through everyday interactions. Together, these contributions reinforce the idea that peer effects do not simply result from the aggregation of individual characteristics but emerge from relational processes situated within specific social contexts, where students continually assemble and renegotiate their educational paths (Piomalli, 2024).

Another key turning point was the Coleman report (Coleman et al., 1966), which highlighted that the social composition of schools and the distribution of students contribute significantly to inequalities in educational opportunities (Rosenqvist, 2018). Coleman's (1986) macro–micro–macro model, known as the 'Coleman boat', further clarified how structural conditions and social interactions

translate into individual outcomes through identifiable mechanisms, linking social phenomena to the constraints and opportunities within which actors operate (Marsden, 2005). Within this framework, social capital plays a central role: embeddedness in peer networks enables students to access information, resources, and support that directly affect their educational outcomes (Coleman, 1988; Portes, 1998). The integration of these perspectives has fostered a multilevel understanding of peer effects, connecting institutional structures, group cultures, and individual behaviours.

Following the structuralist and relational tradition of social capital (Granovetter, 1973; Coleman, 1988; Burt, 1992), this study assumes that educational performance does not depend solely on individual characteristics but also on the quantity and quality of resources embedded in social networks. Social capital is thus defined as the set of resources accessible through family, friendship, or school relations, which can be mobilised to obtain information, support, or opportunities (Woolcock and Narayan, 2000). This conceptualisation comprises three dimensions: the structural dimension, concerning the configuration of networks; the relational dimension, relating to trust, obligations, and norms of reciprocity; and the cognitive dimension, based on shared interpretive codes and meanings. Together, these dimensions provide a robust framework for understanding how peer networks influence students' aspirations, educational choices, and transitions across different stages of their schooling.

Based on reference group theory (Merton and Rossi, 1968), individuals with whom students identify play a key role in shaping their educational paths by providing reference models that guide their aspirations and decisions. According to Kelley (1952), peers can influence students in divergent ways: through conformity, which may promote more ambitious choices, or through social comparison, which may discourage them by highlighting contrasts with others (Smith, 2023). During adolescence, individuals spend a significant amount of time in the school environment, which becomes a crucial space not only for learning but also for personal and social development. Peers therefore become important reference points due to their shared life experiences (Larson and Richards, 1991; Osterman, 2000; Brechwald and Prinstein, 2011). Interactions developed in the classroom often go beyond superficial relationships, evolving into deep and lasting bonds. These social connections can play a significant role in how students navigate important life decisions, including educational choices. In this crucial phase, support from peers with whom students have close relationships can influence their choices, shaping academic and career paths.

Two main processes contribute to the creation and maintenance of peer culture: *peer selection* and *peer socialisation*. The former refers to individuals' tendency to associate with others they perceive as similar, while the latter describes their propensity to be influenced by the attitudes and norms of friends (Raabe and Wölfer, 2019). This tendency to cluster in social groups with similar others is known as the *homophily effect* (McPherson et al., 2001). Students tend to establish friendships primarily with peers who exhibit similar levels of academic achievement, aspirations, and motivations (Raabe et al., 2019, and references therein). Studies on educational transitions have also distinguished between primary and secondary effects. Primary effects are closely associated with academic achievement, where inequalities arise from differences in students' performance. Secondary effects, by contrast, relate to contextual factors such as parental educational level and socioeconomic status (Rosenqvist, 2018; Smith, 2023). These effects influence the formation of social networks, which in turn play a significant role during educational decision-making phases.

In the context of higher education, information-related social capital and support become particularly relevant. Four categories are used to conceptualise social support: *emotional support* based on feelings of trust and mutual esteem; *informational support*, which provides advice or guidance; *appraisal support*, offering evaluative feedback; and *instrumental support*, providing material resources (Malecki and Demaray, 2003). When examining peer effects on educational choices, it is hypothesised that these forms of support play a crucial role in understanding the influence mechanisms at work as high school students face major decisions.

Choices made during major life transitions have significant impacts on income and related social status (Breen and Jonsson, 2005). Students from lower socioeconomic backgrounds often actively seek more information within the school environment due to limited parental guidance and tend to be more affected by peer influence (Roksa and Kinsley, 2019; Smith, 2023).

Based on this theoretical framework, the paper aims to investigate the relevance of social interactions and support mechanisms developed among classmates and friends in high school contexts. This focus is crucial for understanding the broader dynamics at play, particularly the different kinds of assistance students receive from social networks in which they are embedded. Specifically, it explores the role of social networks and the support they provide in shaping students' educational choices, with particular attention to peer influence. The main purpose is to understand how different relational contexts

contribute to decision-making processes regarding the continuation of studies after high school. Two research questions guide the exploratory analysis:

- *RQ1: To what extent do significant social groups influence students' choices during the transition from high school to university? Do classmates play a meaningful role in supporting and assisting students beyond the influence of family members (close ties), friends (strong ties), and acquaintances (weak ties)?*
- *RQ2: How do the social resources embedded in the composition and structure of personal networks affect students' decision-making processes regarding university attendance after high school?*

3. SURVEY DESIGN AND DATA COLLECTION

To address these objectives, a survey was conducted using a non-probabilistic sample of students attending scientific, classical, social science, and vocational high schools in the Campania region (Italy) while participating in a general orientation program. A mixed-method data collection approach was adopted, combining a structured questionnaire with an egocentric network research design (Robins, 2015; McCarty et al., 2019; Marsden and Hollstein, 2023).²

The questionnaire included three main sections.

- **Contextual and sociodemographic characteristics**, including gender, age, municipality and province of residence, type of school attended, and parents' educational level and employment status.
- **University choices**, assessing students' intention to pursue higher education after high school, preferred academic fields (e.g. social sciences, humanities, economics, psychology, law, engineering), and the likelihood of attending a university outside their home region. This section also explored the perceived influence of peers and significant others on students' decisions, identifying key sources of information and the relevance of the opinions of family, friends, classmates, and teachers. Moreover, the role of personal motivations, financial constraints, academic challenges, and peer influence was investigated.

²The questionnaire was administered during guidance laboratories promoted by the Centre of Orientation and Tutoring at the University of Salerno (Italy) during the academic year 2023–2024.

- **Soft skills and individual attitudes**, measured using validated Likert scales. The soft skills self-evaluation scale (3SQ; Lucisano and Du Mérac, 2019) was employed to assess dimensions such as collaboration, empathy, autonomy, problem-solving, and leadership, all critical competencies in decision-making processes. These variables were not incorporated into the analyses presented here.

The relational dimension was captured using an egocentric network protocol implemented with network canvas.³ The egocentric approach is widely used to study social influence and behaviours (Kwon et al., 2014; French et al., 2023), access to resources and social support (Vacca, 2020a, 2020b; Fraudatario et al., 2024), and information flows (Arnaboldi et al., 2016; Smith and Burow, 2020).

To limit the risks associated with self-assessment and recall biases in the identification of alters, the research design incorporated a series of methodological adjustments. Since egocentric networks tend to have a stable core and a more fluid periphery (Perry and Pescosolido, 2012), suggesting that structurally rooted ties are also the most enduring, three social circles of reference (family, school, and wider social context) were predefined. Within these circles, specific key social roles (e.g. parents, siblings, grandparents, classmates, students from other classes, teachers, friends outside school, neighbours, and acquaintances) and the types of resources and support provided by significant alters (informational, emotional, and appraisal support) were recorded. This delimitation of social boundaries facilitated the identification of alters in clearly defined areas, reducing omissions due to salience or selective memory. In addition, interactive sociograms supported data collection by providing respondents with visual ‘anchors’ to more accurately represent their networks in terms of both size and the distribution of social roles (Jeon et al., 2016; Hogan et al., 2020).

Once the boundaries of the relational systems and types of support were defined, network data collection proceeded in three main components:

- **Multiple name generator**, asking: “The educational choice after high school involves an important change in your life. How do the people listed in your three social circles support you in terms of informational support (helping you search for information), appraisal support

³ Further information on Network Canvas (Birkett et al., 2021) is available at <https://networkcanvas.com/>.

(discussing future options), and emotional support (helping you make decisions)?”

- **Name interpreter**, which collected information on each al.ter, including their role within the social circles (parents, relatives, classmates, teachers, friends, acquaintances), demographic attributes (gender, age), type of support provided (informational, emotional, and appraisal), and the perceived strength of the relationship.
- **Alters-by-alters ties**, where respondents reported whether connections existed between pairs of al.ters within each social circle.

The adoption of a mixed-methods design allowed for a nuanced exploration of how the composition and structure of ego-networks shape the educational decision-making process, providing both quantitative and relational perspectives on the social mechanisms influencing students’ transition to higher education.

4. METHODS

The methodological framework described below builds on previous contributions in social network analysis that combine multidimensional data techniques with network data to investigate networks characterized by a single or multiple relations as well as one-mode, two-mode, and multimode networks (Giordano and Vitale, 2007; D’Esposito et al., 2014; Ragozini et al., 2015a,b; Giordano et al., 2019; Genova et al., 2024). These approaches have been successfully applied to both egocentric and complete networks, ranging from the analysis of personal networks and social support (Lumino et al., 2016, 2017) to empirical studies on student mobility choices and, more broadly, on the transition from high school to tertiary education (Columbu et al., 2021b,c; Santelli et al., 2022).

4.1 EGOCENTRIC NETWORK MEASURES

The egocentric networks provide several pieces of information across multiple levels of analysis (Table 1). First, the Ego–Alter Level captures the type of relationships each student (ego) reports for other individuals (alters) in their social circles. At this level, the relationship attributes include the types of support provided, coded as binary variables, and tie strength, measured on a scale from 0 to 4. The al.ter–alter level records the presence or absence of ties between al.ters within al.l social circles. The third level includes al.ter attributes, specifically gender and age group.

Table 1: Ego–alter features of personal networks

Level of analysis	Type of relationship	Attributes	Coding
Ego–alter	Family: kinship type	Informational support Emotional support	1 = support 0 = no support
	School: role (e.g., classmate, teacher)	Appraisal support	
	Social context: role (e.g., friend, neighbour)	Tie strength	3 = very close 2 = close 1 = distant 0 = very distant
Alter–alter	Acquaintance	–	1 = tie exists 0 = no tie
Alter attributes	–	Gender	F = female M = male
		Age group	up to 18 years 18–24 25–29 30–49 50–65 65+

The resulting personal networks were represented as binary, undirected, and unweighted matrices. From these matrices, the composition and structure of each personal network were analysed. Composition was assessed by calculating the size of each social circle, including specific counts of friends and classmates as proxies for peer influence. Additionally, for each social circle, the number of alters providing informational, emotional, or evaluative support was computed, capturing the distribution of social resources across relational domains.

Regarding structural measures, *cohesion* and *centrality* indices of the egocentric networks were calculated using the *egor* package in R (Krenz et al., 2025). Among cohesion measures, *network density* is a basic one that provides an immediate indication of the network structure based on the number of nodes (alters) and the number of ties. Density is computed as the ratio between the actual number of ties in the network and the maximum possible number of ties between

≡

nodes (alters). Density values range from 0 to 1, where 1 indicates a fully connected network, meaning that all possible ties are actually present. The *diameter* (longest shortest path length) was calculated to quantify network “reach” and understand the potential for information diffusion or influence. The *transitivity* (Holland and Leinhardt, 1976) captures the tendency of pairs of nodes to be directly connected when they share a tie with an intermediate node. This measure is particularly useful as it allows the structural properties of the network to be interpreted in terms of social relationship configurations, while also providing an indication of local cohesion. Note that, in the case of personal networks, these measures (density, transitivity, and diameter) were computed on the alter–alter matrix, excluding the ego, as they describe the internal cohesion and overall shape of the alter network.

The second set of network indices consists of centrality measures: *degree*, *betweenness centrality*, *closeness*, and *constraints*. This set of indices captures complementary aspects of network configuration based on the positional properties of nodes (Wasserman and Faust, 1994). *Degree* summarises the number of connections of alters (Borgatti et al., 2013); high values indicate potential popularity or influence due to a greater number of direct connections with other nodes. *Betweenness centrality* (Freeman, 1978) measures the frequency with which a node lies along the shortest paths between other pairs of nodes, thus occupying potential brokerage positions. This positional advantage is often associated with the ability to control the flow of information and other resources within a social subsystem (Brass, 1984). *Closeness centrality*, defined as the reciprocal of the sum of geodesic distances between a node and the alters (Freeman, 1978), reflects how quickly a node can reach, or be reached by, the rest of the network. Burt’s constraint (Burt et al., 2002) measures the extent to which ego’s ties are concentrated within tightly interconnected clusters: high values indicate placement within cohesive and redundant groups, whereas low values suggest a more intermediate position, typical of nodes accessing heterogeneous information sources.

Average degree, betweenness, and closeness were computed for the entire network, including the ego, and subsequently averaged across nodes to capture overall patterns of connectivity, brokerage, and accessibility within the personal network. Constraint was derived as an individual score for the ego and reflects the extent to which their relational environment consists of redundant ties.

Since structural measures may vary as a function of ego-network size, the association between size and all structural measures included in the analysis

(density, diameter, transitivity, average degree, betweenness, closeness, and constraint) was examined. Expected patterns emerged: density and closeness decreased as network size increased, whereas diameter, average degree, and betweenness increased, reflecting the greater structural extent and higher number of potential paths in larger networks. Constraint also showed a negative association with size, indicating reduced redundancy and greater bridging opportunities in larger networks. In contrast, transitivity exhibited no systematic trend, confirming that it captures local clustering dynamics rather than size-induced variations. These patterns are consistent with theoretical expectations for egocentric networks and suggest that size-related variability reflects substantive structural differences rather than distortions arising from the absence of normalisation.

Finally, a gender-based *homophily index* (McPherson et al., 2001) was computed to evaluate individuals' tendency to establish ties with others of the same gender. This index compares the observed frequency of connections between alters of the same gender with that expected randomly, providing a measure of the propensity for gender similarity within the network.

4.2 MULTIDIMENSIONAL EXPLORATORY TECHNIQUES

The data matrix (n, p) , where n corresponds to the total number of ego-networks (295) and p to the composition variables (alter gender, social circle membership, types of support provided) and network structural measures (cohesion and centrality indices), served as the basis for the application of multidimensional exploratory techniques (Lebart et al., 1995). Specifically, PCA and HCA were used.

Through PCA, the optimal number of principal components, linearly associated with the original variables, was identified. Although structural measures were calculated in their standard, non-normalised form, the PCA procedure in R automatically standardises all variables prior to component extraction, ensuring full comparability among indicators measured on different scales. The selection of retained components was based on the eigenvalue >1 criterion (Kaiser, 1960) and inspection of the scree plot (Cattell, 1966).

Factor scores of the retained principal components were subsequently used as input for HCA using Ward's minimum variance method to identify specific network configurations within each derived group. The number of clusters was

determined by examining the dendrogram, the variation in cluster variance, and the substantive interpretability of the resulting solutions. Among the different optimal dendrogram cuts (2, 3, or 6 clusters), a three-cluster structure proved to be the most coherent and analytically informative, allowing a clear and meaningful classification of personal network configurations.

5. RESULTS

A total of 295 students completed both the questionnaire and the egocentric network tool. Within these egocentric networks, students identified 2,874 alters, connected by 4,411 alter-to-alter ties. Following a description of the main characteristics of the student profiles, the networks were analysed in terms of composition variables, including alters' gender, social circles, and types of support provided, as well as structural network measures (i.e. density, centrality indices, and constraint). These variables were subsequently used to perform a PCA and an HCA to identify specific network configurations within each derived group.

5.1 STUDENT PROFILES

Among the respondents, 68.10% were female, and 84.1% attended a lyceum track. Within this group, 26.8% were enrolled in classical lyceums, 22.7% in scientific lyceums, and 34.6% in other lyceum types, while 15.9% attended vocational institutes. Regarding parents' educational levels, 65.75% of mothers held a degree or diploma, compared to 61.94% of fathers. In terms of employment, 53.98% of mothers were permanent employees or self-employed, while 39.45% were housewives. A certain degree of instability emerged in fathers' employment status: 43.60% were employed on fixed-term contracts, 28.77% were self-employed, and only 18.34% held permanent positions.

Concerning university intentions, the majority of students (82%) planned to continue their studies after graduation, 11.20% remained undecided, and only 6.80% did not intend to pursue higher education. Among those aiming to attend university, 61.57% intended to enrol in local institutions near their residence, 19.42% planned to move to other Italian regions or abroad (0.41%), while 18.60%

remained undecided. Regarding the field of study,⁴ 86% were interested in enrolling in a social sciences and humanities degree program, compared to 14% pursuing a program in science, technology, engineering, and mathematics (STEM). Specifically, many students planned to enter health, economics, or law programs, whereas those interested in STEM fields mainly aimed for architecture and engineering programs (Figure 1).

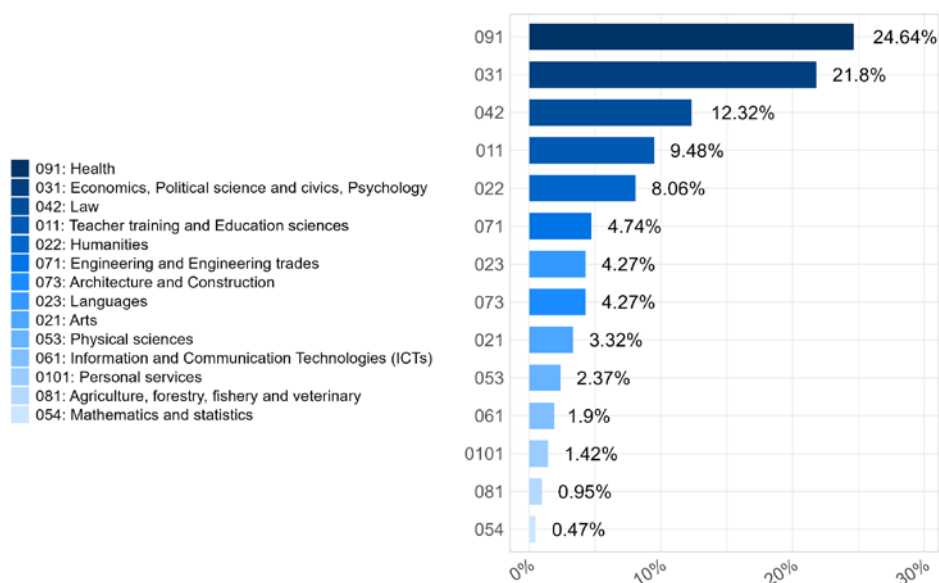


Figure 1: Percentage of respondents by fields of study

Students evaluated a list of factors influencing their university choices using a 5-point Likert scale, ranging from ‘not at all’ to ‘extremely’ (Figure 2). The factors most frequently rated as ‘extremely’ important include individual motivation and aspirations (52.20%), professional opportunities (42.71%), job prospects associated with the degree program (41.01%), the desire to enter the

⁴The degree programs are grouped according to the *International Standard Classification of Education: Fields of Education and Training 2013* (ISCED-F 2013): agricultural, forestry, fishery and veterinary sciences; arts; economics; teacher training and education science; law; health; political sciences and civics; psychology; humanities; languages; personal services; engineering and engineering trades; architecture and construction; physical sciences; information and communication technologies (ICTs); mathematics and statistics.

job market promptly after graduation (21.36%), and interest in a specific discipline (23.05%).

5.2 PERSONAL NETWORK COMPOSITION AND STRUCTURE

The network composition takes into account al.ter's gender, the number and role of each al.ter within specific social circles, and the type of support provided to the ego (Figure 3). Overall, the size of personal networks ranges from a minimum of 2 to a maximum of 24 al.ter, with 63% of al.ter being female. On average, each student is connected to four al.ter, indicating that these networks are relatively small. The average density value of 0.39 reflects a moderate level of cohesion, with many potential connections between al.ter. This degree of interconnectivity facilitates interactions and social support mechanisms, which in turn promote information exchange during the educational transition phase.

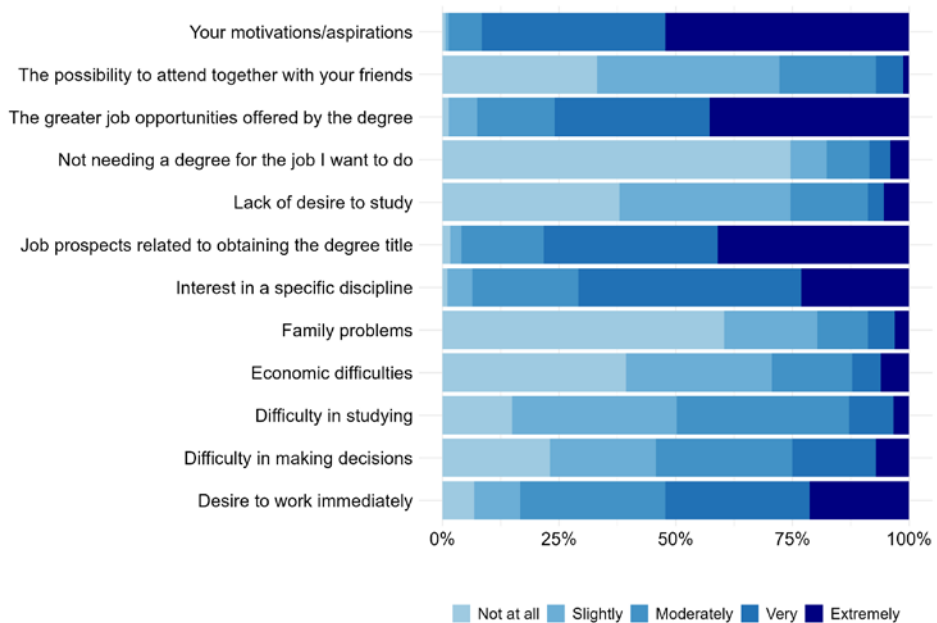


Figure 2: Percentage of respondents by factors influencing university choices

The homophily index, based on gender, shows an average value of 0.66, indicating a high degree of homophily, with students predominantly associating

with same-gender peers. Regarding the role of alters within each social circle,⁵ parents (46.03%) and siblings (19.69%) serve as the primary sources of support for future choices. In the school context, classmates are most frequently mentioned (83.27%), surpassing teachers. Given the relatively low percentage of students referring to their instructors (16.73%), teachers appear to play a limited role in supporting students during university decision-making. Friends (73.94%) are also central to this process, in contrast to romantic partners (14.36%). Concerning the type of support received by the ego, parents, classmates, and friends primarily provide emotional and appraisal support (Figure 3), whereas informational support remains limited across these social circles.

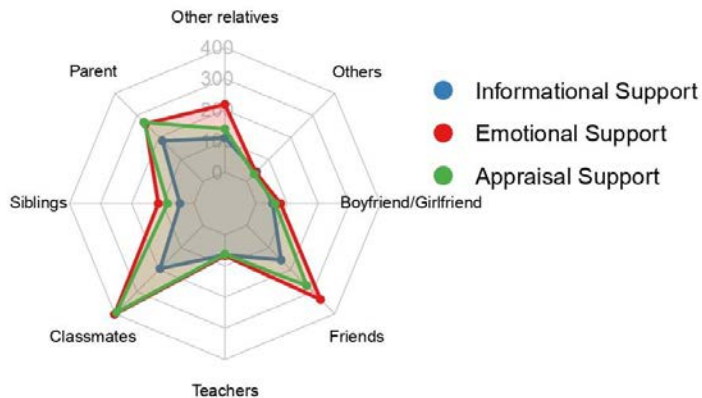


Figure 3: Types of support provided by alters to the ego

The structure of personal networks was examined using centrality measures (betweenness, closeness) and the constraint index, calculated as average values of alter-to-alter ties for each ego. The average betweenness centrality (2.66) indicates that the power of connection is distributed among alters rather than concentrated solely in the egos. The closeness measure is relatively low (0.09), suggesting that alters maintain considerable distance from one another. This

⁵The social circles are described as follows: family, including parents, siblings, and other relatives; school, including classmates, students from other classes, instructors, and administrative staff; and social environment, including friends, romantic partners, acquaintances, neighbours, and university staff.

distance may imply the presence of star networks, which are highly centralised and in which egos play a central role in the flow of information among alters. The constraint index (0.35) reflects a medium level of dependence among alters, indicating limited opportunities to establish new ties. Overall, these networks tend to display relatively closed and constrained structures, with reduced potential for activating new connections and interactions. In terms of transitivity, approximately 50% of possible triads in the networks are closed, indicating a moderate level of clustering. This result fosters a sense of belonging and mutual support among individuals within groups, enhancing overall network cohesiveness. Finally, the network diameter, with an average value of 3.10, indicates a relatively short maximum distance between individuals.

In summary, personal networks exhibit a moderately cohesive structure, characterised by substantial interconnection among alters. The presence of gender homophily and a relatively high constraint index suggests that these networks are composed of individuals with similar characteristics, promoting internal cohesion while limiting the formation of new external ties. Although some central egos play a key role in connecting alters, the networks generally display a tendency towards closure and internal homophily.

5.3 EXPLORATORY ANALYSIS OF PERSONAL NETWORKS

The analytical strategy adopts a two-step approach, designed to identify latent patterns and classify personal networks based on their structural and compositional characteristics. In the first step, PCA was conducted to uncover the latent dimensions underlying the observed variables. This approach integrates network composition characteristics – including alters' attributes such as gender, social roles, and types of support provided to the ego – with structural measures derived from alter-to-alter ties, such as betweenness, closeness, and transitivity. In the second step, the factor scores from the principal components were used as input for HCA, which identifies meaningful clusters of personal networks that share similar profiles in terms of both composition and structure.

The first six principal components with eigenvalues greater than 1 (Kaiser, 1960), obtained through PCA, explain 75.71% of the cumulative variance (Table 2). The first component captures the composition of students' networks, reflecting the number of alters – particularly women – alongside key structural characteristics. The second component highlights emotional and appraisal

support within school and social contexts, emphasising relationships with classmates and friends. The third component focuses on network cohesion and internal structure, as measured by transitivity, diameter, and density. The fourth component underscores the family's role as a multifaceted source of support, providing appraisal, emotional, and informational resources. The fifth component represents gender homophily and the importance of informational support within cohesive and interconnected school networks, particularly among classmates. Finally, the sixth component captures a pattern driven by gender-based homophily. For simplicity, Figure 4 displays the factorial map based on the first two principal components, which account for 49.60% of the total variance.⁶ The analysis emphasises these two components, as they capture the most salient dimensions of network composition and support dynamics. Specifically, the first component differentiates between small, cohesive networks – characterised by high density and the absence of structural holes – and larger, less cohesive, more heterogeneous networks, in which all social circles contribute to exchanges of support relevant to decision-making. The second component distinguishes between social circles, positioning the broader social environment at the higher end and the school context at the lower end, with both settings providing all three forms of social support.

Table 2: Principal component analysis results for the six principal components

Principal Component	Eigenvalue	% of variance	Cumulative variance (%)
1	10.82	37.31	37.31
2	3.56	12.29	49.60
3	2.67	9.22	58.82
4	2.00	6.91	65.73
5	1.54	5.30	71.03
6	1.36	4.69	75.71

⁶The other factorial maps are reported in Appendix A.

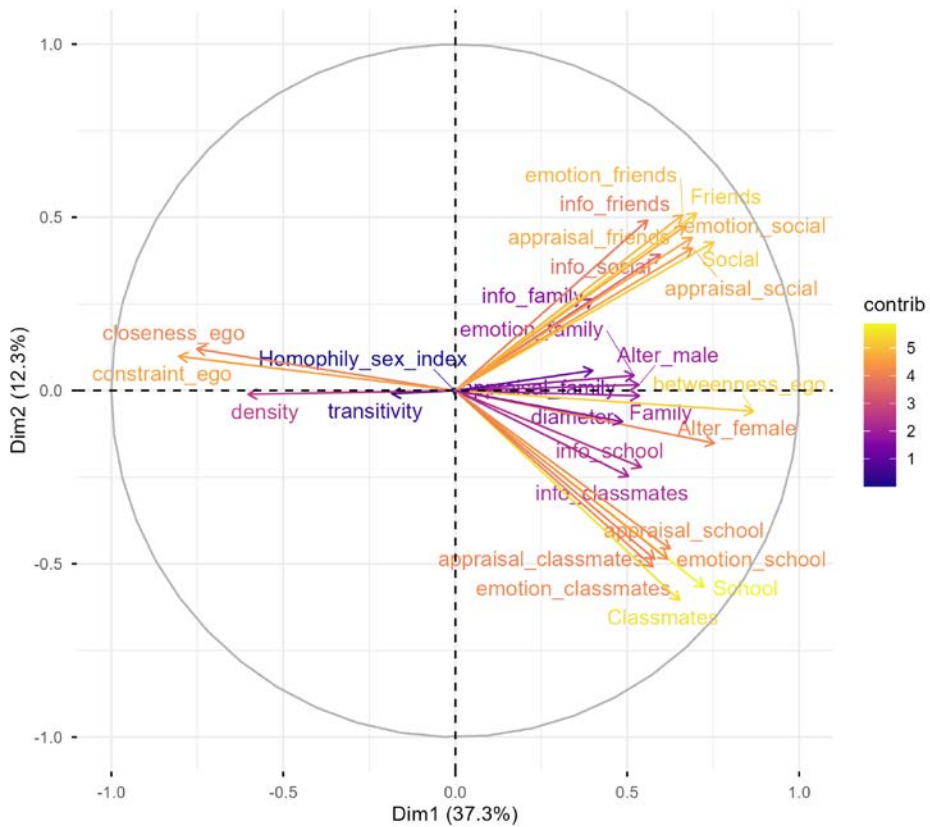


Figure 4: Correlation circle of the first two principal components

5.4 CLUSTERING PERSONAL NETWORKS

Based on the PCA results, an HCA was performed, identifying a three-cluster solution (Figure 5).⁷ The findings reveal distinct personal network configurations with varying types of support and social structures, which are further explored

⁷The Appendix B presents the results of the HCA on principal components conducted using the *FactoMineR* package in R. Tables B.1, B.2, and B.3 report the variables that contribute to the characterisation of Cluster 1, Cluster 2, and Cluster 3, respectively, along with their descriptive statistics and associated test values.

through network measures of structural variables (Table 3) and representative egos within each group (Figure 6).

Personal networks in the **first cluster** (108 students in total – see Table B.1) are characterised by small sizes and a highly cohesive structure within a specific social circle. The closed nature of these networks, combined with redundant ties and limited informational and emotional support from both school and family, may constrain the ego's ability to receive valuable guidance for navigating educational and career pathways. For instance, Ego 284 (Table 3, Figure 6), a female attending a vocational high school, had both parents with a high school diploma; her mother was employed on a fixed-term contract, while her father held a permanent position. Ego 284 was uncertain about university enrolment, mobility choices, and academic discipline. Her personal network consists of five alters, including parents, a relative, one classmate, and her boyfriend. For this student, the family played a prominent role in providing informational and appraisal support, while emotional support came from the classmate and boyfriend. The structural characteristics of Ego 284's network reveal a closed and limited configuration, with a betweenness centrality of 1.17, indicating a relatively low level of connection across alters. The closeness value (0.14) suggests that alters within the social circles are accessible to the ego. The network exhibits a high constraint value (0.41), reflecting redundant ties and limited sources of information. Internal cohesion is strong, with a transitivity value of 1, indicating tight connections within the family circle. The network's diameter (1), combined with relatively low density, points to a compact structure where alters are primarily linked to the ego but do not form connections among themselves, particularly across different social circles. This network configuration limits the ego's access to diverse resources and new information, reducing opportunities for external support and exposure to varied perspectives. Consequently, the ego's social capital appears weak, which could impede her ability to make well-informed decisions and feel adequately supported in her future choices.

In the **second cluster** (176 students in total; see Table B.2), students exhibit broad, open, and moderately cohesive personal networks. These networks are sufficiently large to provide access to a wide range of resources. In terms of support, students benefit from a balanced combination of emotional and appraisal support across school, peer, and family contexts. For example, Ego 201 (Table 3, Figure 6), a female student enrolled in a classical lyceum, had both parents holding a diploma; her mother was a housewife, and her father held a permanent contract. She had decided to enrol at a university in another Italian region,

choosing a STEM program in architecture and construction. Regarding the support received, Ego 201's network is characterised by a well-balanced mix of informational, emotional, and appraisal support. The family provides both informational and emotional support, while the teacher and a friend contribute informational resources, underscoring the relevance of the school context. Appraisal support comes from both family and classmates. The ego thus has access to a variety of sources extending beyond the family circle to the school and broader social spheres. The network's structural measures reveal a betweenness centrality of 3.29, indicating that alters effectively connect different actors within the network. The low closeness value reflects a less centralised configuration, while the diameter (3) suggests a more extended network. The density (0.41) and average degree (6.42) point to a well-connected structure with numerous ties among alters. The constraint value (0.25) indicates a network structure offering greater opportunities to access diverse sources of support. Additionally, the transitivity (0.61) shows a diversified network with cross-circle connections. This configuration provides the ego with access to a range of resources and forms of support, which were instrumental in facilitating her decision to pursue a university path outside her region of residence and within a STEM field.

Finally, the **third cluster** (11 students in total; see Table B.3) comprises students with large personal networks, characterised by a substantial number of alters who play active roles in connecting actors. These alters function as bridging nodes, facilitating the flow of information and reducing the ego's dependence on any single source of support. As a representative case, Ego 230 (Table 3, Figure 6) was a female student attending a classical lyceum, with a favourable socioeconomic and cultural background: both parents held a degree or diploma and occupied permanent job positions. She decided to enrol in a university outside her home region, pursuing a non-STEM degree program. Ego 230's network includes 18 alters, comprising 7 males and 11 females. She receives all forms of support across the three social circles, primarily from parents, other relatives, classmates, and friends. The network exhibits a complex structure in which the ego plays a pivotal bridging role among social circles, as reflected by the high betweenness centrality value (5.53). Despite a low closeness value, which may constrain direct access to some resources offered by alters, the density (0.31) and average degree (6.95) indicate a well-connected network. The relatively low constraint value (0.19) underscores the ego's opportunities to access diverse resources through new connections.

The HCA highlights a progression from constrained, inward-focused networks (Cluster 1) to open, diversified, and resource-rich configurations (Cluster 3), with Cluster 2 representing an intermediate balance. These distinctions underscore the importance of both network size and structural diversity in shaping students' social capital and their ability to make informed, autonomous choices regarding their educational futures.

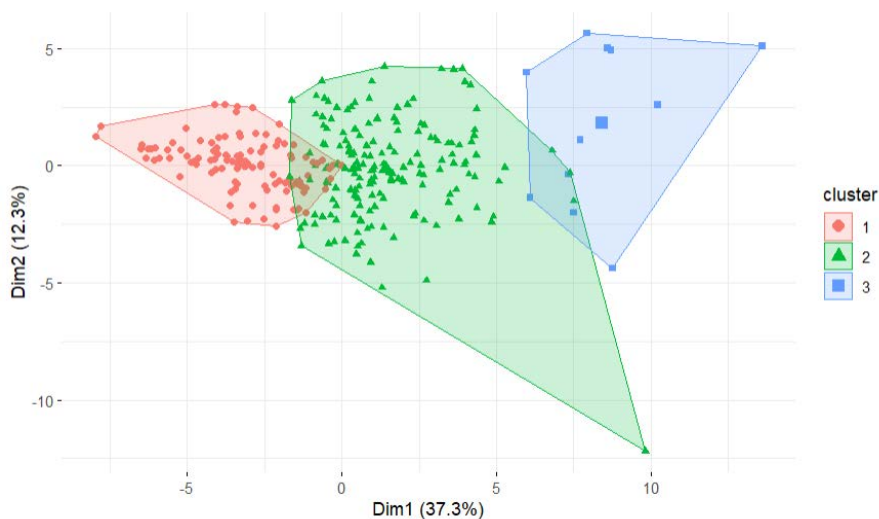


Figure 5: Factorial map and three-cluster solution of personal networks on the first two principal components

Table 3: Structural measures of personal networks in the three cluster solution

	<i>Ego 284</i> (Cluster 1)	<i>Ego 201</i> (Cluster 2)	<i>Ego 230</i> (Cluster 3)
Density	0.30	0.41	0.31
Average degree	2.67	6.43	6.95
Betweenness	1.17	3.29	5.53
Closeness	0.14	0.05	0.04
Diameter	1	3	4
Transitivity	1.00	0.61	0.59
Constraint	0.41	0.25	0.19

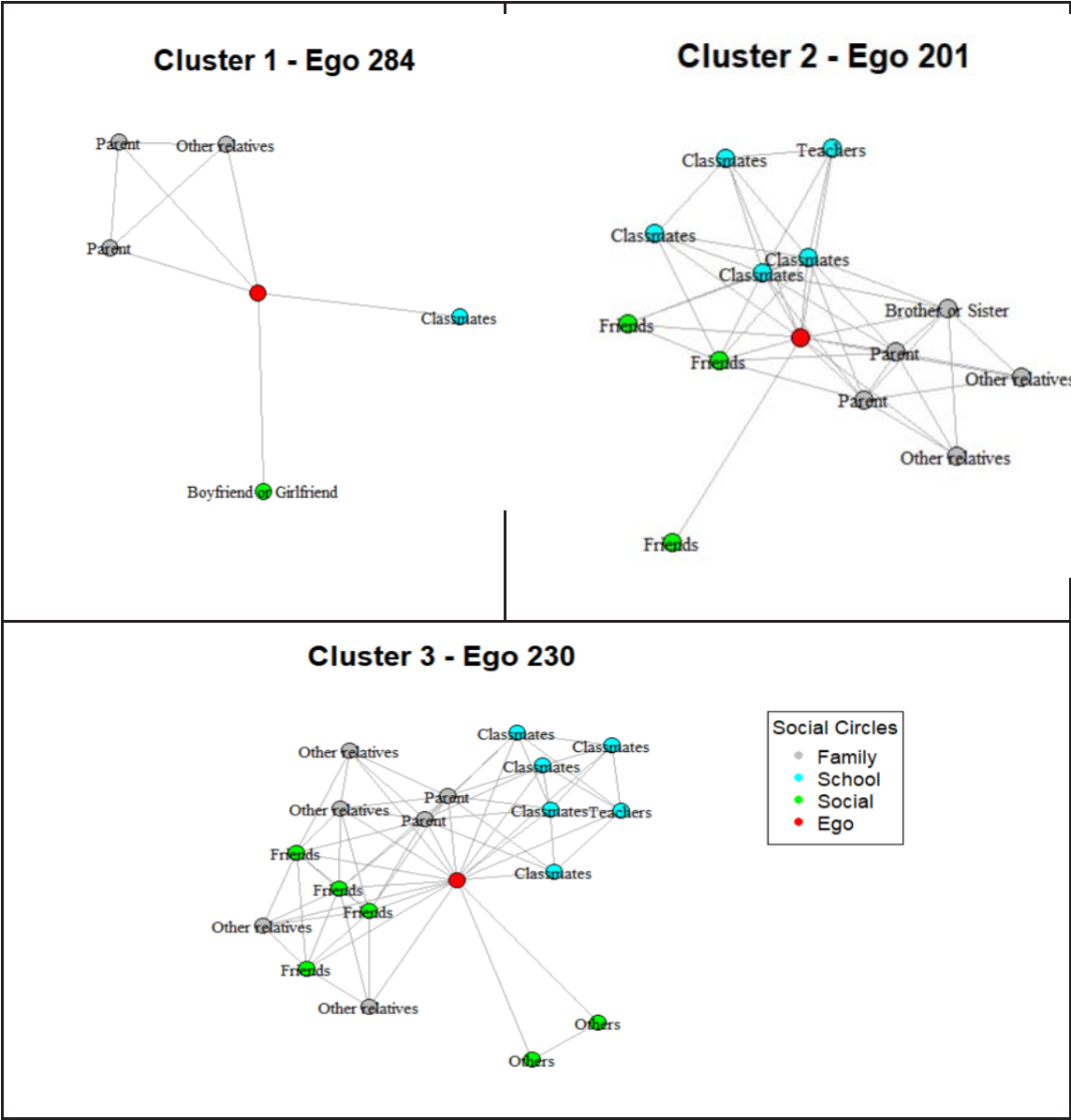


Figure 6: Visualisation of personal networks in the three-cluster solution

5.5 PERSONAL NETWORK CLUSTERS AND UNIVERSITY CHOICES

To examine whether distinct personal network configurations correspond to different orientations towards higher education, differences in university enrolment choices were analysed across the three clusters. The patterns shown in Table 4 indicate that enrolment outcomes are uneven: all students in Cluster 3, which comprises the most open and resource-rich networks, enrol in university, whereas students in Cluster 1, whose networks are more limited in personal resources, show a higher incidence of indecision and non-enrolment. Cluster 2 falls in between. A chi-squared test confirmed a statistically significant association ($\chi^2_4 = 18.04, p = 0.0012$), and Fisher's exact test supported this result ($p = 0.002$), indicating that students' enrolment decisions vary meaningfully across the three network configurations. These findings suggest that the composition and structural properties of personal networks play an important role in shaping students' post-secondary education choices.

Table 4: Distribution of student university choices by personal network cluster

Cluster membership	University enrolment	No university enrolment	Undecided	Total
Cluster 1	70.37	12.96	16.67	100%
Cluster 2	88.07	3.41	8.52	100%
Cluster 3	100.00	0.00	0.00	100%
Total	82.03	6.78	11.19	100%

membership

6. DISCUSSION AND CONCLUDING REMARKS

The paper offers initial insights into the interplay between students' individual characteristics and the structure of their personal networks during key educational transitions. The findings reveal clear patterns in network composition and structure, particularly regarding gender dynamics, social circles, and types of support, suggesting that these configurations significantly influence students'

decisions about university enrolment. Exploratory analysis identified distinct network prototypes, highlighting notable variations in students' available social capital and the support they receive. Across all groups, family members, friends, and classmates emerged as central providers of emotional and appraisal support, whereas informational support remained limited across social circles. Observed patterns of gender homophily and network closure indicate networks that are cohesive but, in many cases, constrained, potentially limiting access to diverse external resources. The three-cluster solution emphasises the existence of markedly different student profiles, ranging from those embedded in small, inward-looking networks with restricted social capital (Cluster 1) to those situated within large, well-connected networks rich in relational resources (Cluster 3). Students embedded in highly interconnected networks appear better equipped to access both informational and emotional support, which is particularly pivotal during educational transitions. In contrast, students with weaker personal networks tend to face challenges in making informed decisions due to limited flows of information and support.

Specifically, the study frames social capital as the number and diversity of individuals that students can mobilise during the decision-making process, aligning with the broader understanding of social capital as the set of resources embedded in social networks that can influence educational pathways. Operationalising social capital through egocentric network data enabled the identification of, among other configurations, a subset of large, low-density personal networks characterised by structural holes and heterogeneity among alters. This approach aligns with Burt's (2000) theory, which posits that individuals bridging disconnected groups can access a broader range of information and opportunities, thereby enhancing their social capital. Similarly, Granovetter's (1973) concept of the strength of weak ties suggests that connections to diverse, loosely connected individuals provide access to novel resources and perspectives. Furthermore, family characteristics and school type emerge as important contextual factors influencing students' decision-making processes, consistent with their treatment as covariates in established models in the literature. This exploratory investigation provides an initial description of how the composition and structure of students' personal networks – spanning family, school, and peer contexts – relate to their access to information and support during educational transitions.

Some limitations should be acknowledged. First, the composition of respondents exhibits self-selection bias. Specifically, because data collection

occurred during guidance programs in the social sciences, the sample predominantly consisted of female students from particular types of high schools. As a result, the respondents' demographic and academic profiles show a degree of homogeneity that may not fully represent other educational tracks. Moreover, the reliance on self-reported data to measure network ties and all forms of support may introduce potential biases affecting data quality. Despite these caveats, the main findings are consistent with those documented in the educational policy literature, suggesting their relevance for evidence-based decision-making. The analytical approach presented can be adapted to diverse educational and social contexts, provided that appropriate measures are taken to account for heterogeneity in student characteristics.

In terms of practical implications, promoting social interactions within school contexts could serve as an effective policy for supporting students during educational transitions. This study underscores the importance of personal networks in shaping educational aspirations and decisions, with social capital playing a pivotal role in influencing students' access to resources during academic choices. These findings could guide policymakers on enhancing social support systems, ensuring that all students – particularly those with lower social capital – have access to necessary resources. To better support students during the critical transition from high school to university, guidance programs should encourage the development of weak ties with external experts, such as university advisors. These weak ties can facilitate access to strategic information – including academic pathways, scholarship opportunities, and career trajectories – which is often unavailable within close-knit family or peer networks. Strengthening these broader connections is particularly important for students from socioeconomically disadvantaged backgrounds, who may lack access to informed guidance within their social environment. Encouraging schools to foster such initiatives could help reduce structural inequalities in the post-secondary decision-making process.

While our main findings highlight the theoretical and policy relevance of weak ties and guidance programs in expanding students' social capital, the mechanisms through which these factors influence educational decisions remain only partially identifiable. Given the cross-sectional nature of the data, we cannot disentangle whether information, motivation, or peer norms represent the primary channels of influence, nor can we rule out potential endogeneity, namely the possibility that highly motivated students actively cultivate more diverse networks. It is therefore likely that social capital and motivation mutually

reinforce each other, in ways shaped by students' initial social and cultural capital. Future longitudinal research will be essential to empirically examine these mechanisms, address endogeneity concerns, and trace how these processes unfold over time.

Looking ahead, future research should implement multilevel models adapted to egocentric network data, accounting for the hierarchical structure in which alters (level 1) are nested within egos (level 2). This approach would allow for a more nuanced understanding of how ego-level and alter-level characteristics, as well as alter-by-alter ties, shape the types of support students receive and the educational choices they make. Furthermore, the data collection method adopted in this study did not allow for the reconstruction of classroom-level (sociocentric) networks. While the egocentric data provided detailed information about each student's personal network, it did not capture whether the same individuals (alters) appeared across multiple students' networks or whether a student could serve as both an ego and an alter in different cases. Consequently, understanding overall peer dynamics within each classroom represents a valuable challenge for future research, particularly in mapping complete sociocentric networks to explore broader patterns of influence and support among students at the school level during higher education choices.

Paper Funding: The paper is financially supported under the National Recovery and Resilience Plan (NRRP), Mission 4, Component 2, Investment 1.1, Call for tender No. 104 published on 2.2.2022 by the Italian Ministry of University and Research (MUR), funded by the European Union – NextGeneration EU– Project Title From high school to university: Assessing peers' influence in educational inequalities and performances – CUP F53D23006150006- Grant Assignment Decree No. 1060 adopted on 07/17/2023 by the Italian Ministry of University and Research (MUR).

REFERENCES

- Alnawas, I. (2015). Student orientation in higher education: development of the construct. *Higher Education*. 69(4): 625-652.
- Arnaboldi, V., Conti, M., La Gala, M., Passarella, A. and Pezzoni, F. (2016). Ego network structure in online social networks and its impact on information diffusion. *Computer Communications*. 76: 26-41.

- Bacci, S. and Bertaccini, B. (2021). Assessment of the university reputation through the analysis of the student mobility. *Social Indicators Research*. 156(2): 363-388.
- Biagi, B., Ciucci, L., Detotto, C. and Pulina, M. (2024). Do diverse degree courses matter for university attractiveness?. *The Annals of Regional Science*. 72(4): 1189-1229.
- Birkett, M., Melville, J., Janulis, P., Phillips II, G., Contractor, N. and Hogan, B. (2021). Network canvas: Key decisions in the design of an interviewer-assisted network data collection software suite. *Social Networks*. 66: 114-124.
- Borgatti S. P., Everett M.G. and Johnson, J.C. (2013). *Analyzing Social Networks*. Sage, Thousand Oaks, CA.
- Brass D.J. (1984). Being in the right place: a structural analysis of individual influence in an organization. *Administrative Science Quarterly*. 29(4): 518-539.
- Breen, R., and Jonsson, J.O. (2005). Inequality of opportunity in comparative perspective: Recent research on educational attainment and social mobility. *Annual Review of Sociology*. 31(1): 223-243.
- Brechwald, W.A., and Prinstein, M.J. (2011). Beyond homophily: A decade of advances in understanding peer influence processes. *Journal of Research on Adolescence*. 21(1): 166-179.
- Brooks, R. (2003). Young people's higher education choices: The role of family and friends. *British Journal of Sociology of Education*. 24(3): 283-297.
- Brouwer, J., De Matos Fernandes, C.A., Steglich, C.E., Jansen, E.P., Hofman, W. and Flache, A. (2022). The development of peer networks and academic performance in learning communities in higher education. *Learning and Instruction*. 80: 101603.
- Burt, R.S. (1992). *Structural Holes: The Social Structure of Competition*. Harvard University Press, Cambridge, MA.
- Burt, R.S. (2000). The network structure of social capital. *Research in organizational behavior*. 22: 345-423.
- Burt, R.S. (2002). The social capital of structural holes. *The New Economic Sociology: Developments in an Emerging Field*. 148(90): 122.
- Cattell, R.B. (1966). The scree test for the number of factors. *Multivariate Behavioral Research*. 1(2): 245-276.
- Chaudhry, S., Tandon, A., Shinde, S. and Bhattacharya, A. (2024). Student psychological well-being in higher education: The role of internal team

- environment, institutional, friends and family support and academic engagement. *Plos One*. 19(1): e0297508.
- Chuard, C., Aerne, A., Eugster, B. and Hodler, R. (2025). Ethnic clustering in schools and early career outcomes. *Economics of Education Review*. 106: 102645.
- Coleman, J.S., Campbell, E., Hobson, C., McPartland, J., Mood, A., Weinfeld, F. and York, R (1966). *Equality of Educational Opportunity*. U.S. Department of Health Education and Welfare, Washington.
- Coleman, J.S. (1986). Social theory, social research, and action. *American Journal of Sociology*. 91(6): 1309-1335.
- Coleman, J.S., (1988). Social capital in the creation of human capital. *American Journal of Sociology*. 94(1): S95-S120.
- Columbu, S., Porcu, M. and Sulis, I. (2021a). University choice and the attractiveness of the study area: Insights on the differences amongst degree programmes in Italy based on generalised mixed-effect models. *Socio-Economic Planning Sciences*. 74: 100926.
- Columbu, S., Porcu, M., Primerano, I., Sulis, I. and Vitale, M.P. (2021b). Geography of Italian student mobility: A network analysis approach. *Socio-Economic Planning Sciences*. 73: 100918.
- Columbu, S., Porcu, M., Primerano, I., Sulis, I. and Vitale, M.P. (2021c). Analysing the determinants of Italian university student mobility pathways. *Genus*. 77(1): 34.
- Corsaro, W.A. (1985). *Friendship and Peer Culture in the Early Years*. Ablex, Norwood, NJ.
- Corsaro, W.A. (2005). *The Sociology of Childhood*. Sage, Thousand Oaks, CA.
- DeLay, D., Zhang, L., Hanish, L.D., Miller, C.F., Fabes, R.A., Martin, C.L., Kochel, K.P. and Updegraff, K.A. (2016). Peer influence on academic performance: A social network analysis of social-emotional intervention effects. *Prevention Science*. 17: 903-913.
- D'Esposito, M.R., De Stefano, D. and Ragozini, G. (2014). On the use of multiple correspondence analysis to visually explore affiliation networks. *Social Networks*. 38: 28-40.
- Engberg, M.E. and Wolniak, G.C. (2010). Examining the effects of high school contexts on postsecondary enrollment. *Research in Higher Education*. 51(2): 132-153.
- Freeman L.C. (1978). Centrality in social networks: Conceptual clarification. *Social Networks*. 1: 215-239.

- French, K.R., Lee, C.D. and Zellner, A. (2023). Beliefs, perception, and change: A study of ego network influence on first-year teachers. *Journal of Educational Change*. 24(4): 1053–1075.
- Fraudatario, M.C., Giordano, G. and Zaccaria, A.M. (2025). Beyond mixed embeddedness: Multilevel personal networks of migrant entrepreneurs in Naples and Manchester. *International Migration*. 63(2): e13262.
- Genova, V.G., Giordano, G., Ragozini, G. and Vitale, M.P. (2024). An analytic strategy for data processing of multimode networks. *Advances in Data Analysis and Classification*. 18(3): 745-767.
- Giordano, G. and Vitale, M.P. (2007). Factorial contiguity maps to explore relational data patterns. *Statistica Applicata*. 19(4): 297-306.
- Giordano, G. and Vitale, M.P. (2011). On the use of external information in social network analysis. *Advances in Data Analysis and Classification*. 5(2): 95-112.
- Giordano, G., Ragozini, G. and Vitale, M.P. (2019). Analyzing multiplex networks using factorial methods. *Social Networks*. 59: 154-170.
- Granovetter, M.S. (1973). The strength of weak ties. *American Journal of Sociology*. 78(6): 1360-1380.
- Hogan, B., Janulis, P., Phillips II, G. L., Melville, J., Mustanski, B., Contractor, N. and Birkett, M. (2020). Assessing the stability of egocentric networks over time using the digital participant-aided sociogram tool network canvas. *Network Science*. 8(2): 204-222.
- Holland, P.W., and Leinhardt, S. (1976). Local structure in social networks. *Sociological Methodology*. 7: 1-45.
- Jeon, G.Y., Ellison, N.B., Hogan, B. and Greenhow, C. (2016). First-generation students and college: The role of Facebook networks as information sources. In *Proceedings of the 2016 ACM Conference on Computer-Supported Cooperative Work and Social Computing*, San Francisco, CA: 887-899.
- Kaiser, H.F. (1960). The application of electronic computers to factor analysis. *Educational and Psychological Measurement*. 20(1): 141-151.
- Kallio, R.E. (1995). Factors influencing the college choice decisions of graduate students. *Research in higher education*, 36(1): 109-124.
- Kelley, H.H. (1952). Two functions of reference groups. In G.E. Swanson, T.M. Newcomb, and E.L. Hartley, editors, *Readings in social psychology*. Holt, Rinehart & Winston, New York: 410-414.

- Krenz, T., Krivitsky, P.N., Vacca, R., Bojanowski, M., Gamper, M., Herz, A. and McCarty, C. (2025). Egor: Import and Analyse Ego-Centered Network Data. <https://CRAN.R-project.org/package=egor>
- Kwon, K.H., Stefanone, M.A. and Barnett, G.A. (2014). Social network influence on online behavioral choices: Exploring group formation on social network sites. *American Behavioral Scientist*. 58(10): 1345-1360.
- Lareau, A. and Cox, A. (2011). Social class and the transition to adulthood: Differences in parents' interactions with institutions. In M.J. Carlson and P. England, editors, *Social Class and Changing Families in an Unequal America*. Stanford University Press, Stanford, CA: 134-164.
- Lareau, A. (2015). Cultural knowledge and social inequality. *American Sociological Review*. 80(1): 1-27.
- Larson, R.W., and Richards, M.H. (1991). Boredom in the middle school years: Blaming schools versus blaming students. *American Journal of Education*. 99(4): 418-443.
- Lebart, L., Morineau, A. and Piron, M. (1995). *Statistique exploratoire multidimensionnelle* (Vol. 3). Dunod, Paris.
- Lomi, A., Snijders, T.A., Steglich, C.E., and Torló, V.J. (2011). Why are some more peer than others? Evidence from a longitudinal study of social networks and individual academic performance. *Social Science Research*. 40(6): 1506-1520.
- Lucisano, P., and Rubat Du Merac, E. (2019). Soft skills self-evaluation questionnaire (3sq), characteristics and psychometric proprieties. In P. Lucisano and A.M. Notti, editors, *Training actions and evaluation processes: Atti del Convegno Internazionale SIRD*. Pensa MultiMedia Editore, Lecce: 609-622.
- Lumino, R., Ragozini, G. and Vitale, M.P. (2016). Investigating social support patterns of single mothers from a social network perspective. *International Review of Social Research*. 6(4): 182-194
- Lumino, R., Ragozini, G., van Duijn, M., and Vitale, M.P. (2017). A mixed-methods approach for analysing social support and social anchorage of single mothers' personal networks. *Quality & Quantity*. 51(2): 779-797.
- Malecki, C.K. and Demaray, M.K. (2003). What type of support do they need? Investigating student adjustment as related to emotional, informational, appraisal, and instrumental support. *School Psychology Quarterly*. 18(3): 231-252.

- Mamas, C. (2018). Exploring peer relationships, friendships and group work dynamics in higher education: Applying social network analysis. *Journal of Further and Higher Education*. 42(5): 662-677.
- Marsden P.V. (2005). The Sociology of James S. Coleman. *Annual Review of Sociology*. 31(1): 1-24.
- Marsden, P.V., and Hollstein, B. (2023). Advances and innovations in methods for collecting egocentric network data. *Social Science Research*. 109: 102816.
- McCarty, C., Lubbers, M.J., Vacca, R., and Molina, J.L. (2019). *Conducting personal network research: A practical guide*. The Guilford Press, New York.
- McPherson, M., Smith-Lovin, L. and Cook, J.M. (2001). Birds of a feather: Homophily in social networks. *Annual Review of Sociology*. 27: 415-444.
- Mpamhanga, K., Jones, H., Orpin, H., Jarvis, J., Mansi, G., Molesworth, C. and Monsey, H. (2025). *Transition Into Higher Education*. Routledge, London.
- Merton, R.K. and Rossi, A.S. (1968). Contributions to the theory of reference group behavior. In R. K. Merton, editor, *Social Theory and Social Structure*. The Free Press, New York: 279-334.
- Molina, J.L., García-Macías, A., Lubbers, M.J., and Valenzuela-Garcia, H. (2020). The embeddedness of social capital in personal networks. *Network Science*. 8(2): 189-203.
- Osterman, K.F. (2000). Students' need for belonging in the school community. *Review of Educational Research*. 70(3): 323-367.
- Parsons, T. (1959), The school class as a social system: Some of its functions in American society. *Harvard Educational Review*. 29(4): 297-318.
- Pascarella, E.T., Pierson, C.T., Wolniak, G.C. and Terenzini, P.T. (2004). First-generation college students: Additional evidence on college experiences and outcomes. *The Journal of Higher Education*. 75(3): 249-284.
- Perry, B. L., Pescosolido, B.A. and Borgatti, S.P. (2018). *Egocentric Network Analysis: Foundations, Methods, and Models* (Vol. 44). Cambridge university press, Cambridge.
- Perry B.L. and Pescosolido, B.A. (2012). Social network dynamics and biographical disruption: The case of fFirst-timers" with mental illness. *American Journal of Sociology*. 118(1): 134-175.
- Piromalli, L. (2024). A tu per tu con le scelte scolastiche: La ricerca interazionista sulle carriere educative. In O. Giancola and L. Salmieri, editors,

- Disuguaglianze educative e scelte scolastiche. Teorie, processi e contesti.* FrancoAngeli, Milano: 118-134.
- Poldin, O., Valeeva, D. and Yudkevich, M. (2016). Which peers matter: How social ties affect peer-group effects. *Research in Higher Education*. 57(4): 448-468.
- Portes A. (1998). Social capital: Its origins and applications in modern sociology. *Annual Review of Sociology*. 24: 1-24.
- Priulla, A., Vittorietti, M., Genova, V.G. and Attanasio, M. (2025). The high school to university transition: exploring the interplay of territory, socioeconomic factors, and gender dynamics. *PloS One*. 20(10): e0326510.
- Raabe, I.J., Boda, Z. and Stadtfeld, C. (2019). The social pipeline: How friend influence and peer exposure widen the STEM gender gap. *Sociology of Education*. 92(2): 105-123.
- Raabe, I.J. and Wölfer, R. (2019). What is going on around you: Peer milieus and educational aspirations. *European Sociological Review*. 35(1): 1-14.
- Ragozini, G. and D'Esposito, M.R. (2015). Archetypal networks. In *ASONAM '15 Proceedings of the 2015 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*: 807-814.
- Ragozini, G., De Stefano, D. and D'Esposito, M.R. (2015). Multiple factor analysis for time-varying two-mode networks. *Network Science*. 3(1): 18-36.
- Reay, D., David, M. and Ball, S. (2001). Making a difference?: Institutional habituses and higher education choice. *Sociological Research Online*. 5(4): 14-25.
- Robins, G.L. (2015). *Doing Social Network Research: Network-Based Research Design for Social Scientists*. Sage, London.
- Rosenqvist, E. (2018). Two functions of peer influence on upper-secondary education application behavior. *Sociology of Education*. 91(1): 72-89.
- Roksa, J. and Kinsley, P. (2019). The role of family support in facilitating academic success of low-income students. *Research in Higher Education*. 60(4): 415-436.
- Sá, C., Florax, R. J. and Rietveld, P. (2006). Does accessibility to higher education matter? Choice behaviour of high school graduates in the Netherlands. *Spatial Economic Analysis*. 1(2): 155-174.

- Sandefur, G.D., Meier, A. M. and Campbell, M.E. (2006). Family resources, social capital, and college attendance. *Social Science Research*. 35(2): 525-553.
- Santelli, F., Ragozini, G. and Vitale, M.P. (2022). Assessing the effects of local contexts on the mobility choices of university students in Campania region in Italy. *Genus*. 78(1): 5.
- Smith, E. (2023). Peer preferences and educational decisions: Heterogeneous associations across student socioeconomic status. *British Journal of Sociology of Education*. 44(2): 374-393.
- Smith, J.A. and Burow, J. (2020). Using ego network data to inform agent-based models of diffusion. *Sociological Methods & Research* 49(4): 1018-1063.
- Stadtfeld, C., Vörös, A., Elmer, T., Boda, Z. and Raabe, I.J. (2019). Integration in emerging social networks explains academic failure and success. *Proceedings of the National Academy of Sciences*. 116(3): 792-797.
- Stanton-Salazar, R.D. (2011). A social capital framework for the study of institutional agents and their role in the empowerment of low-status students and youth. *Youth & Society*. 43(3): 1066-1109.
- Upcraft, M.L., Gardner, J.N. and Barefoot, B.O. (2005). *Challenging and Supporting the First-Year Student: A Handbook for Improving the First Year of College*. Jossey-Bass, San Francisco, CA.
- Usala, C., Porcu, M. and Sulis, I. (2023). The high school effect on students' mobility choices. *Statistical Methods & Applications*. 32(4): 1259-1293.
- Vacca, R. (2020a). Multilevel models for personal networks: Methods and applications. *Statistica Applicata - Italian Journal of Applied Statistics*. 30(1): 59-97.
- Vacca, R. (2020b). Structure in personal networks: Constructing and comparing typologies. *Network Science*. 8(2): 142-167.
- Vandelannote, I. and Demanet, J. (2021). "What's high school got to do with it?" Secondary school composition, school-wide social capital and higher education enrollment. *Research in Higher Education*. 62(5): 680-708.
- Vörös, A., Boda, Z., Elmer, T., Hoffman, M., Mephram, K., Raabe, I.J. and Stadtfeld, C. (2021). The Swiss StudentLife Study: Investigating the emergence of an undergraduate community through dynamic, multidimensional social network data. *Social Networks*. 65: 71-84.
- Walpole, M. (2003). Socioeconomic status and college: How SES affects college experiences and outcomes. *The Review of Higher Education*. 27(1): 45-73.

- Wasserman S. and Faust K. (1994), *Social Network Analysis: Methods and Applications*. Cambridge University Press, Cambridge.
- Wilcox, P., Winn, S. and Fyvie-Gauld, M. (2005). It was nothing to do with the university, it was just the people: The role of social support in the first year experience of higher education. *Studies In Higher Education*. 30(6): 707-722.
- Willis, P. (1977). *Learning to Labour: How Working Class Kids Get Working Class Jobs*. Saxon House, Farnborough.
- Woolcock, M., and Narayan, D. (2000). Social capital: Implications for development theory, research, and policy. *The World Bank Research Observer*. 15(2): 225-249.

Figures A.1–A.2 present the correlation circles from the principal component analysis, illustrating the contribution and orientation of each variable within the multidimensional space defined by the extracted components.

Figure A.1 Correlation circle of third and fourth principal components

Figure A.2 Correlation circle of fifth and sixth principal components

Appendix B

Tables B.1–B.3 report the variables contributing to the definition of the three clusters, based on the results of hierarchical clustering on principal components. For each variable, the tables provide the mean and standard deviation within the cluster, the overall mean and standard deviation in the dataset, the V.test statistic, and the associated p-value.

Table B.1 – Cluster 1

Variables	V. test	Mean in category	Overall mean	SD in category	Overall SD	P-value
Constraint_Ego	11.622	0.500	0.354	0.180	0.164	0.000
Closeness_Ego	10.782	0.146	0.092	0.081	0.066	0.000
Density	8.520	0.540	0.394	0.245	0.223	0.000
Transitivity	3.670	0.590	0.504	0.315	0.305	0.000
Info_Family	-4.895	0.967	1.483	0.941	1.372	0.000
Info_Classmates	-5.009	0.385	0.853	0.647	1.218	0.000
Appraisal_Family	-5.427	1.588	2.154	1.003	1.359	0.000
Info_Friends	-5.723	0.139	0.682	0.337	1.235	0.000
Info_School	-5.745	0.513	1.136	0.742	1.414	0.000
Emotion_Classmates	-6.593	0.928	1.759	1.042	1.641	0.000
Info_Social	-6.626	0.400	1.077	0.610	1.331	0.000
Emotion_Family	-6.8656	1.729	2.594	1.263	1.642	0.000
Emotion_School	-6.998	1.076	2.052	1.104	1.819	0.000
Family	-7.041	2.995	3.836	1.251	1.555	0.000
Appraisal_Classmates	-7.069	0.844	1.720	1.036	1.615	0.000
Appraisal_Friends	-7.099	0.403	1.192	0.567	1.448	0.000
Diameter	-7.120	2.157	3.099	0.910	1.724	0.000
Alter_Male	-7.277	2.293	3.616	1.532	2.369	0.000
Appraisal_School	-7.650	0.981	1.993	1.122	1.723	0.000
Emotion_Friends	-7.748	0.485	1.465	0.714	1.648	0.000
Appraisal_Social	-7.981	0.674	1.608	0.711	1.525	0.000
Classmates	-8.513	1.424	2.594	1.222	1.792	0.000
Emotion_Social	-8.548	0.879	1.989	0.868	1.692	0.000
Friends	-8.776	0.767	1.944	0.876	1.748	0.000
School	-9.474	1.679	3.115	1.272	1.975	0.000
Social	-9.847	1.305	2.629	0.933	1.753	0.000
alter_female	-9.862	3.723	6.034	1.644	3.054	0.000
betweenness_ego	-11.004	1.228	2.657	0.859	1.692	0.000

Table B.2 – Cluster 2

Variables	V. test	Mean in category	Overall mean	SD in category	Overall SD	P-value
Betweenness_Ego	8.531	3.349	2.657	1.394	1.692	0.000
School	7.760	3.851	3.115	1.809	1.975	0.000
Alter_Female	7.759	7.170	6.034	2.725	3.054	0.000
Social	7.434	3.254	2.629	1.571	1.753	0.000
Classmates	7.139	3.208	2.594	1.693	1.792	0.000
Diameter	6.305	3.621	3.099	1.876	1.724	0.000
Friends	6.068	2.453	1.944	1.638	1.748	0.000
Emotion_Social	5.935	2.471	1.989	1.595	1.692	0.000
Alter_Male	5.784	4.273	3.616	2.346	2.369	0.000
Family	5.481	4.244	3.836	1.488	1.555	0.000
Appraisal_School	5.307	2.432	1.993	1.650	1.723	0.000
Emotion_Family	5.048	2.992	2.594	1.592	1.642	0.000
Appraisal_Classmates	5.029	2.110	1.720	1.591	1.615	0.000
Emotion_School	4.955	2.485	2.052	1.833	1.819	0.000
Appraisal_Social	4.861	1.964	1.608	1.420	1.525	0.000
Emotion_Friends	4.813	1.845	1.465	1.598	1.648	0.000
Emotion_Classmates	4.694	2.128	1.759	1.637	1.641	0.000
Appraisal_Family	4.083	2.420	2.154	1.353	1.359	0.000
Appraisal_Friends	3.544	1.438	1.192	1.341	1.448	0.000
Info_School	2.583	1.311	1.136	1.334	1.414	0.009
Info_Social	2.575	1.241	1.077	1.112	1.331	0.010
Info_Family	2.502	1.647	1.482	1.350	1.372	0.013
Transitivity	-3.414	0.454	0.504	0.292	0.305	0.001
Density	-7.331	0.315	0.394	0.159	0.223	0.000
Closeness_Ego	-9.599	0.061	0.092	0.017	0.066	0.000
Constraint_Ego	-10.192	0.274	0.354	0.063	0.164	0.000

Table B.3 – Cluster 3

Variables	V. test	Mean in category	Overall mean	SD in category	Overall SD	P-value
Info_Friends	11.301	4.818	0.682	1.898	1.235	0.000
Info_Social	10.180	5.091	1.077	1.730	1.331	0.000
Appraisal_Friends	8.874	5.000	1.192	1.651	1.448	0.000
Info_Classmates	8.213	3.818	0.853	2.167	1.218	0.000
Info_School	7.920	4.455	1.136	2.105	1.414	0.000
Appraisal_Social	7.706	5.091	1.608	1.564	1.525	0.000
Emotion_Friends	7.239	5.000	1.465	1.651	1.648	0.000
Friends	6.602	5.3634	1.944	1.432	1.748	0.000
Emotion_Social	6.365	5.1812	1.990	1.641	1.692	0.000
Info_Family	5.968	3.909	1.483	1.928	1.372	0.000
Betweenness_Ego	5.888	5.609	2.657	1.420	1.692	0.000
Social	5.788	5.636	2.629	1.367	1.753	0.000
Appraisal_School	5.710	4.909	1.993	1.443	1.723	0.000
Alter_Female	4.985	10.545	6.034	3.367	3.054	0.000
Emotion_School	4.962	4.727	2.052	1.863	1.819	0.000
Appraisal_Classmates	4.953	4.092	1.720	1.831	1.615	0.000
Emotion_Classmates	4.609	4.000	1.759	2.174	1.641	0.000
Emotion_Family	4.384	4.727	2.594	1.286	1.642	0.000
School	3.996	5.455	3.115	1.437	1.975	0.000
Family	3.710	5.545	3.836	1.233	1.555	0.000
Alter_Male	3.526	6.091	3.616	3.0288	2.369	0.000
Appraisal_Family	3.229	3.455	2.154	2.105	1.359	0.001
Classmates	3.161	4.273	2.594	1.814	1.792	0.002
Closeness_Ego	-2.561	0.042	0.092	0.017	0.067	0.010
Density	-2.683	0.216	0.394	0.087	0.223	0.007
Constraint_Ego	-3.160	0.200	0.354	0.053	0.164	0.002

EQUAL OPPORTUNITIES TO “COUNT”? A LONGITUDINAL POPULATION STUDY ABOUT CHOICE OF STEM AT UNIVERSITY BY GENDER

Patrizia Giannantoni¹, Patrizia Falzetti

*National Institute for the Evaluation of the Educational System (INVALSI)
Rome, Italy*

Abstract.

Extensive research has reported that Science Technology Engineering and Mathematics (STEM) disciplines are perceived as “male fields”. This erratic belief is largely responsible for gender gap observed in school and to the underrepresentation of women in STEM majors at university and in the labour market. This paper examines the factors that influence boys’ and girls’ decisions to enrol in STEM, making use of a longitudinal population dataset obtained by joining together Italian data from Ministry of education, University register and INVALSI census test. We investigate the role of math skills, as well as individual and family characteristics, as predictors of STEM academic choices. Key findings highlight the central influence of math ability on the choice of STEM studies, particularly when assessed through teachers’ grades, and reveal an opposite effect of mother’s education on STEM enrolment by gender: increasing the likelihood of STEM selection for girls while decreasing it for boys.

Keywords: *STEM, Gender gap, University, Math skills, Population, Longitudinal data*

¹ Corresponding author (patrizia.giannantoni@invalsi.it)

1. INTRODUCTION

The construction of gender roles, the gender gap in STEM disciplines throughout academic and professional trajectories and occupational segregation are topics widely investigated across multiple research domains. The European Institute for Gender Equality (EIGE, 2021) highlights significant horizontal segregation in higher STEM education, particularly in Southern European countries. This segregation in tertiary education and the labor market represents the culmination of a process that begins in early childhood education and extends throughout the schooling years.

Empirical studies demonstrate that gender disparities in mathematical achievement emerge by adolescence, with girls, on average, scoring lower than boys at age 15 (Borgonovi et al., 2018). This disparity is most pronounced among high-achieving students (Ellison and Swanson, 2010) and substantially influences university major selection, with women being significantly less likely to pursue STEM fields (Turner and Bowen, 1999; Card and Payne, 2021). Cross-national analyses confirm that while gender differences in mathematics performance are negligible in primary school, they expand significantly during adolescence (Contini et al., 2017; Borgonovi et al., 2018; Ellison and Swanson, 2010).

A broad consensus exists that these disparities are not rooted in biological differences but are instead shaped by an array of interrelated social factors (Osborne et al., 2003). Gender stereotypes, differential teacher expectations, familial attitudes and socio-economic constraints contribute to variations in self-confidence, aspirations, and academic choices among boys and girls. Eccles (2007) suggests that girls' lower inclination toward STEM careers is partially attributable to societal expectations that subtly direct them toward non-scientific fields. Similarly, research underscores the role of cultural norms, self-esteem, and perceptions of competence in shaping students' academic trajectories (Eccles and Wang, 2016).

Social representations and stereotypes regarding gender and science influence academic interests and outcomes, as demonstrated by Master and Meltzoff (2011). Their meta-analysis of international datasets, including the Trends in International Mathematics and Science Study (TIMSS) and the Programme for International Student Assessment (PISA), confirms that gender equity in education is crucial not only for female students' mathematical

performance but also for their self-confidence and appreciation of mathematics (Else-Quest et al., 2010).

Understanding the mechanisms through which these stereotypes operate is vital for developing interventions aimed at mitigating gender disparities in STEM education and employment. Crucial moments in educational trajectories, such as transitions between schooling levels, significantly impact students' decisions to pursue STEM pathways. Research employing specialized survey instruments has examined the factors influencing university STEM enrollment decisions (Allegretti et al., 2014).

While numerous studies have examined these issues internationally, relatively few have focused on the Italian context. Priulla et al. (2024) applied machine learning techniques to predict university enrollment patterns based on Italian students' educational backgrounds, uncovering significant gender disparities in STEM participation. Additionally, Priulla and Attanasio (2023) used discrete-time multi-state Markov models to analyze gender differences in university completion rates, demonstrating the impact of prior academic performance on tertiary education outcomes. Barone and Assirelli (2019) further investigated gender segregation in higher education, concluding that school type significantly mediates gender disparities in STEM enrollment.

Given the substantial influence of high school type on university major selection, this study focuses on graduates from scientific lyceums—a subpopulation most likely to pursue STEM disciplines at university. Understanding the educational choices and academic performance of these students is essential for informing policies aimed at fostering gender equity in STEM fields.

2. RESEARCH QUESTIONS

In the existing literature, there are no comprehensive studies on the factors that influence the choice of STEM in tertiary education, especially at the population level in a country with a strong gender gap, such as Italy. The present study aims to gain a better understanding of these factors, focusing on students attending scientific lyceum, who represent the elective subpopulation targeted to pursue a career in highly scientific.

This research aims at answering the following questions: to what extent is maths ability important/predictive of STEM enrolment? Higher scores in math at standardized test, or alternatively higher teachers' evaluation in math, are associated with increased probability of choice of STEM subjects at university? How influential is, in this framework, the social and family context?

Trying to answer these questions, the study sheds light on how all these factors operate jointly according to gender, focusing on similarities and differences in the effects that skills, familial and social factors have of boys and girls in the selection of STEM at university.

3. METHODS

3.1 Sample

An original dataset was created for this research project by combining three different administrative data sources: Italian Ministry of Education, National University Register and National Institute for the Evaluation of the Education and Training System (INVALSI) data². INVALSI's mission is to evaluate the national school system as a whole and to provide schools with guidance for self-evaluation and improvement. The INVALSI surveys consist of standardized tests administered to students in Italian schools at 5 different grade levels. The second and fifth classes of primary school (grades 2 and 5), the third class of lower secondary school (grade 8) and the second and final classes (grades 10 and 12-13) of upper secondary school. The INVALSI tests are designed to measure students' basic skills in subjects such as mathematics, reading comprehension and English. The tests are administered on paper in grades 2 and 5, on paper in grades 8 and 10 until 2017, and on computer in lower and upper secondary schools from 2018 (computer-based tests). The idea behind the construction of the integrated dataset is to make it possible, for the first time with Italian data, to follow a student's academic career from entry into the education system to university

² This study was conducted within the framework of the Agreement "From High School to Job Placement: An Analysis of University Careers and Student Mobility from Southern to Northern Italy" in collaboration with the Italian Ministry of Education and Merit, the Italian Ministry of University and Research and the University of Palermo.

enrolment and graduation. A crucial point of the analysis was the process of matching students in the three data sources: INVALSI has a "student individual code (SIDI)" as a unique identifier for each individual student, while the university register uses the tax code to identify students, and finally the Ministry of Education has both the SIDI and tax code variables in its dataset. Through this two-way correspondence, it was possible to construct a combined, fully anonymized data source that, at the population level, matched the cohort of students enrolled in all Italian universities (public and private) in the 2019/2020 school year with their characteristics and skills in grade 13 of upper secondary school in 2018/19, in a retrospective longitudinal perspective.

3.2 Variables

After the matching process, we obtained a joint dataset with all the information about the students coming from the different sources. INVALSI provided the following variables: students' performance in all subjects (Italian, mathematics and English) from primary school until the last year of upper secondary school (grade 13), information on students' family background (e.g. parents' level of education, family socio-economic status, country of origin, language spoken at home), and information on the geographical location of the school. The Ministry of Education provided: basic demographic information on students, year and grade obtained in the final examination at grade 13. University register provided information related to university's first enrolment: academic year, university course of study, with the specification of the code of the course chosen and the "class of discipline" associated with it. For this research, we selected a subsample of university students composed by all students enrolled in Italian universities (public and private) in the academic year 2019/2020, who attended the last year of scientific lyceum in the previous year (2018/2019). We selected variables related to the following areas:

- Demographic characteristics
- Math proficiency
- Socio-cultural status of student's family
- School characteristics
- STEM enrolment at university

As individual characteristics, we considered, together with gender, the variables: regularity of study (regular/repeating) and migration background (native, I-generation, II-generation), as they have been widely shown to influence school careers and choices (OECD, 2015; Ismu 2020). For each student, demographic information from the Ministry of Education was used as a cross-validation for the same data from the INVALSI dataset.

For the assessment of ability in mathematics, both the teacher's marks in the first quarter of the year and the INVALSI test score (WLE score) in mathematics were considered. Both variables are numerical and were categorised into classes for the analysis. Specifically, the teacher's mark in mathematics is collected with a range from 1 to 10 and was classified as follows: 1 to 5 'insufficient', 6 'sufficient', 7 to 8 'good', 9 and 10 'excellent'. The INVALSI WLE score, elaborated on the principles of Rasch analysis, was classified into levels from 1 to 5 according to cut-off scores based on the descriptive proficiency level approach (Desimoni, 2018).

In relation to socio-economic status, two variables were considered: The index of economic, social and cultural status (ESCS) and the mother's level of education. The ESCS is a synthetic measure of the cultural and economic background of students, constructed through multiple component analysis of three main factors related to the student's family: school attainment of both parents, job of parents, and possession of some goods from the family (Campodifiori et al., 2010). Mother's educational attainment was prioritised over that of the father to account for gender-related disparities in tertiary education (Carneiro et al., 2013). This variable was obtained directly from the INVALSI student questionnaire. As this variable was not collected in 2018/2019 for students in grade 13 of schooling, we matched from the student's questionnaire administered in grade 10 in the year 2015/2016 to the same students. This procedure led to approximately 20% of missing data, and a similar percentage is the proportion of missing data on ESCS.

To ascertain the relationship between missing data and outcome (Stem/Not Stem choice) for both ESCS and mother educational attainment, the non-significance of the missing category in the regression model was tested. Given that the missing data, arising from the discordance between grades 13 and 10, were randomly distributed, only students with valid answers on mother's education and SES items were retained in the sample.

Concerning schools, the geographical location within the macro region (North-East, North-West, Centre, South and Islands) and the management of the school (public/private) were considered.

The outcome variable for this study, i.e. enrolment in a STEM course at university, was created based on the official classification of university courses into five "Areas of discipline": These are: Art and literature, Economics, Law and social sciences, Health studies and STEM. The Ministry of University and Research provides this data in an open data format on their website: <https://dati-ustat.mur.gov.it/dataset/metadati/resource/adb4d6bd-0ad9-4004-8e23-b7baa2e45495>).

The finale sample was composed of 61,574 students, whose characteristics are reported in Table 1.

Table 1 – Characteristics of the sample

	Number	Percentage
Gender		
Female	27,626	44.9
Male	33,948	55.1
Teacher marks in math		
Insufficient	12,602	20.5
Sufficient	16,823	27.3
Good	25,485	41.4
Excellent	6,664	10.8
Level in math score at INVALSI		
Level 1	1,119	1.8
Level 2	2,820	4.6
Level 3	7,419	12.0
Level 4	12,774	20.7
Level 5	37,442	60.8
ESCS-classes		
Low	5,619	9.1
Medium-low	11,657	18.9
Medium-high	14,904	24.2
High	29,394	47.7

Mother education		
Lower secondary school or lower	7,719	12.5
Professional qualification	2,910	4.7
High school diploma	24,609	40.0
University degree or higher	26,336	42.8
Migration background		
Native	59,050	95.9
I generation migrant	828	1.3
II generation	1,696	2.8
Regularity		
Regular attendant	59,945	97.4
Late attendant	1,629	2.6
Geographical area of the school		
North-west	16,526	26.8
North-east	10,931	17.8
Centre	12,778	20.8
South	16,965	27.6
South and islands	4,374	7.1
Management of the school		
Public	58,592	95.2
Private	2,982	4.8
Total	61574	100.0

Source: our elaboration on linked dataset INVALSI-Ministry of Education – University Register

3.3 Data analysis

Descriptive analyses were initially performed on the entire population to illustrate the relationship between typology of upper secondary school and choice of STEM major at university, according to gender. We then compared the proportion of enrolment in STEM according to typology of high school previously attended (scientific lyceum, other lyceum, technical and professional schools) distinctly, for male and female students.

Given that previous school track is the primary predictor of university course selection, a secondary focus was directed exclusively on students from scientific lyceum, as they constitute the "eligible population" for pursuing a

STEM career at university. This subpopulation was subjected to inferential analyses designed to estimate the major drivers for the choice of STEM at university. Given the expectation of differential mechanisms operating on boys and girls, two logistic models stratified by gender were performed to illustrate the gendered effect that each variable has on STEM university choices for the two groups. In instances of multicollinearity (i.e. socio-economic status and mother's education; teacher's mark and INVALSI score in math), a single variable was selected to represent that dimension, based on the higher parsimony of the model, evaluated by the Akaike information criterion (AIC) (Akaike, 1974) and the analysis of VIF values. Eventually, a visual representation was created illustrating the probabilities of selecting STEM courses for boys and girls, according to teachers' evaluations in mathematics. This representation was created by means of the following formula for predicted probabilities:

$$P(Y = 1 | X) = \frac{1}{1 + e^{-(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k)}}$$

where Y is the event “selecting STEM”, and included covariates are gender and teacher's mark in mathematics.

4. RESULTS

4.1 Preliminary analyses

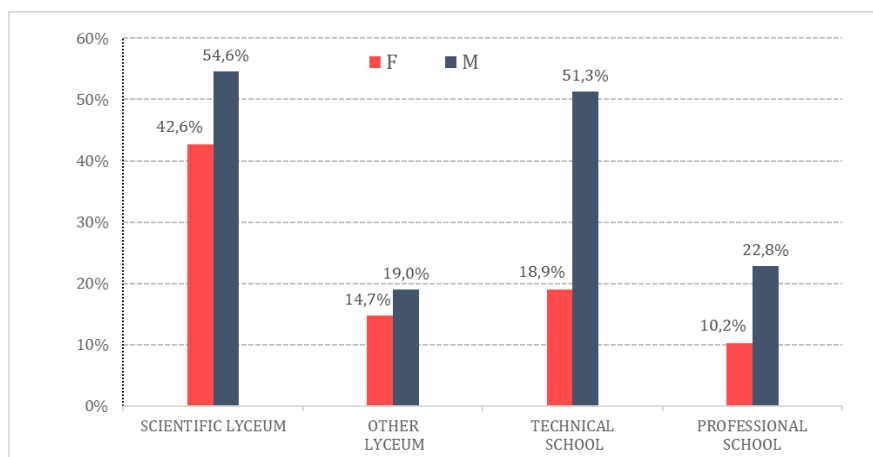
The type of high school previously attended is undoubtedly one of the primary factors influencing the selection of STEM or non-STEM majors at university. As illustrated in Table 2, scientific lyceums and technical schools represent the previous academic background of over 80% of students who have selected STEM at university (55.8% of whom attended scientific lyceums and 25.5% of whom attended technical schools). The chi-squared test for the association between type of high school and STEM field at university has a significance lower than 0.001.

Table 2 – Proportion of STEM choice at university according to previous school track attended – Italy, academic year 2019/2020

Typology high school	University area of enrolment		
	Non Stem	Stem	Total
Scientific lyceum	28.6%	55.8%	37.7%
Other lyceum	45.2%	16.7%	35.7%
Technical school	20.7%	25.5%	22.3%
Vocational school	5.6%	1.9%	4.4%
Total	100.0%	100.0%	100.0%

However, the gender imbalance in STEM choice is evident irrespective of the school track previously attended by the students (Figure 1). The gender disparity in STEM fields persists even when controlled for the inequality in the distribution of students in high-school track by gender (girls more concentrated in humanities, boys in the scientific fields). Notably, the choice of STEM is characterized by a pronounced male predominance among students coming from technical schools, with a gap of approximately 38 percentage points favouring males (51.3% versus 18.9%). Furthermore, among students enrolled in scientific lyceums, the proportion of male students choosing a STEM course at university significantly exceeds that of their female counterparts by more than 12 percentage points (54.6% versus 42.6%), underscoring a pronounced gender imbalance. For all the subgroups (scientific lyceum, other lyceum, technical and vocational school) the difference in the proportion of STEM by gender is significant with chi squared test with a value lower than 0.001.

Figure 1 – Proportion of males and females choosing STEM at University according to previous school track attended – Italy, academic year 2019/2020



The analyses presented hereinafter are thus developed only on the subpopulation of Italian students coming from scientific lyceums, in order to better understand the factors that hinder the propensity of girls in pursuing STEM studies at university, although their possession of scientific academic background.

In a first step of analysis, we focused exclusively on female students at university who attended scientific lyceum, comparing those who selected STEM courses with those who selected non-STEM courses. The analysis encompassed a range of individual characteristics (migration background, regularity during study), family and geographical context (ESCS index, mother educational attainment, geographical area where school is located) and mathematical skills as assessed by teachers and by INVALSI test.

In Table 3 are reported all the variables that exhibited statistical significance at the 0.001 level at the Chi-squared test. The distribution of girls selecting STEM subjects varies significantly with respect to their mathematical skills, as measured by both teachers' grades and INVALSI test scores. A substantial

proportion of girls enrolled in STEM programmes had been assessed as "good/excellent" by their mathematics teachers the previous year (40.8%), compared to only 22.2% of girls in non-STEM programmes. Moreover, nearly two-thirds (63.0%) of "STEM girls" achieved Level 5 (very high competence) on the INVALSI test in grade 13, whereas the corresponding share in the non-STEM group was 48%. Conversely, 10% of girls in the non-STEM group obtained very low/low INVALSI scores, compared to only 4.3% among STEM girls.

With regard to contextual variables, mother's education was the only characteristic showing a significant though numerically modest, difference between the STEM and non-STEM groups, with the expected advantage for girls whose mothers have higher educational attainment. No statistically significant differences emerged for the remaining variables examined.

In this second research line, an inferential perspective was adopted to estimate the effects of all the characteristics previously described on the propensity to select a STEM career, *ceteribus paribus*. The objective of this study was to estimate the relative weight of mathematical ability—hypothesized to exert a legitimate and theoretically grounded influence on the selection of academic orientation—in comparison to the impact of social factors, which are posited to attenuate the propensity of female students to pursue university trajectories in science and technology disciplines.

To this end, we have implemented two logistic regression models, stratified by gender. This approach enabled the estimation of the predictive power of distinct characteristics separately for boys and for girls, facilitating the comparison of similar and different patterns of effects in the two groups.

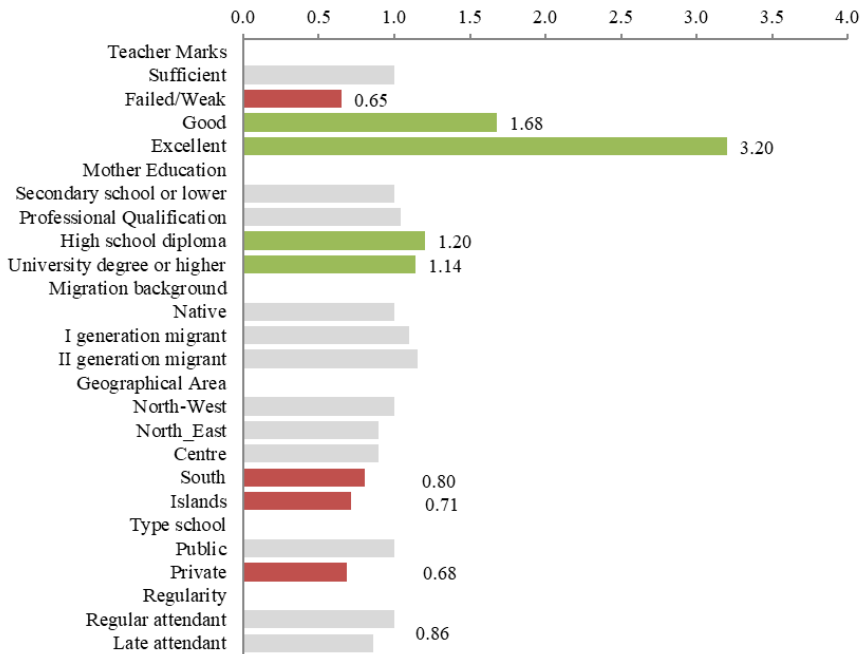
Table 3 – Percentage distribution of girls from scientific lyceum selecting STEM /NON STEM according to the most relevant variables discriminating between the 2 groups

		NOT STEM	STEM
MATH TEACHER MARK	Insufficient	22.0	10.40
	Sufficient	29.8	22.40
	Good	26.0	26.40
	Excellent	22.2	40.80
<i>Total</i>		<i>100.0</i>	<i>100.00</i>
MATH INVALSI SCORE	1 - Very low	2.4	0.90
	2 – Low	7.0	3.40
	3 – Sufficient	16.8	11.60
	4 – High	25.7	21.00
	5 - Very high	48.1	63.00
<i>Total</i>		<i>100.0</i>	<i>100.00</i>
MOTHER EDUCATIONAL ATTAINMENT	Lower secondary school or lower	14.9	12.50
	Professional qualification	5.4	4.50
	High school diploma	40.7	42.50
	University degree or higher	39.0	40.50
<i>Total</i>		<i>100.0</i>	<i>100.00</i>

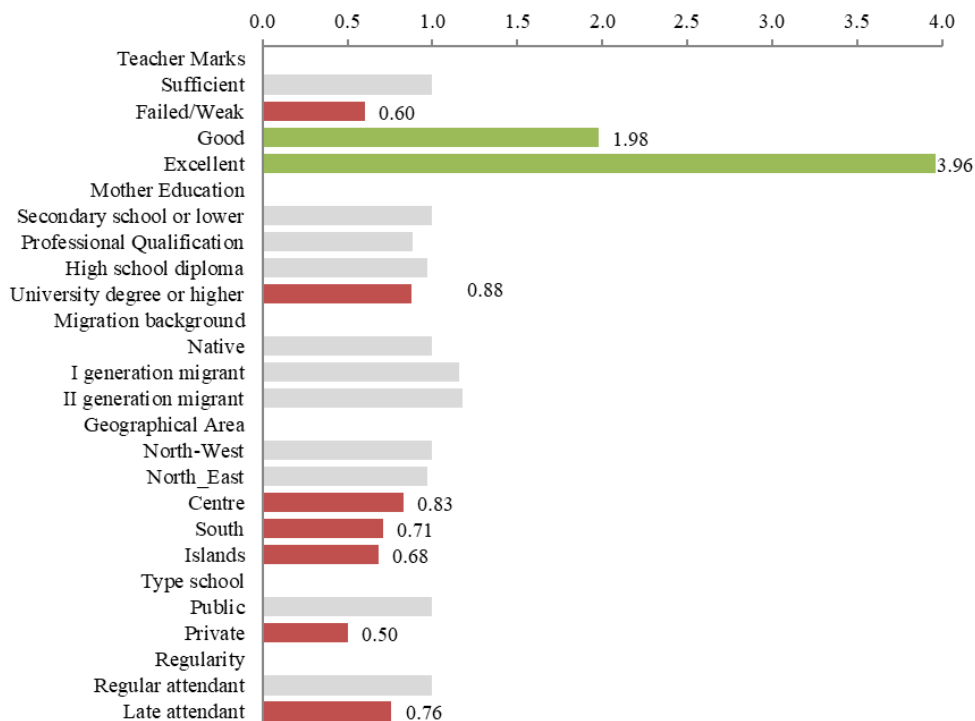
In selecting the variables to be included in the model, it was necessary to choose between pairs of variables capturing the same underlying latent constructs, namely:

- ESCS and mother’s educational attainment, both serving as proxies for the family’s socio-cultural background;
- Teacher-assigned grades and INVALSI test scores, both reflecting students’ mathematical aptitude.

The selection was made based on the Akaike information criterion in the comparison models with the alternative variables. For socio-economic factors, the model with the mother's education resulted more parsimonious than the one with ESCS (AIC = 18979.7 versus 18995.8); with regard to math competencies, the teacher marks resulted as a better predictor for STEM choice than standardized test scores (AIC = 18255.7 versus 18670.5).



(a)



(b)

Figures 2a and 2b –Odds Ratio from Logistic Models predicting choice of STEM at university for girls (a) and boys (b)

In Figure 2 the results from the logistic models are presented for both girls (a) and boys (b). The findings indicate that mathematical skills are the strongest predictor of choosing a STEM degree at university, regardless of gender. For both boys and girls, strong performance in mathematics increases the likelihood of selecting a STEM course by more than threefold. This effect is slightly stronger for boys: their probability of enrolling in a STEM program nearly quadruples when they receive high mathematics evaluations from their teachers, whereas for girls the probability increases by approximately 3.2 times following an excellent mathematics assessment.

Mother's educational attainment appears to exert an opposite effect on girls and boys. Specifically, girls whose mothers hold a university or higher degree

show approximately 20% higher odds of enrolling in a STEM degree compared to those whose mothers completed only upper secondary education (OR 1.17 and 1.20). Conversely, for boys, a higher maternal education is associated with lower odds of selecting a STEM course (OR 0.88 for boys whose mothers had a university degree). The geographical location of the student's high school, which is a good proxy of geographical region of life, exhibits a comparable pattern for both genders: Students from northern regions display a higher propensity to pursue STEM fields compared to their peers from central and southern regions.

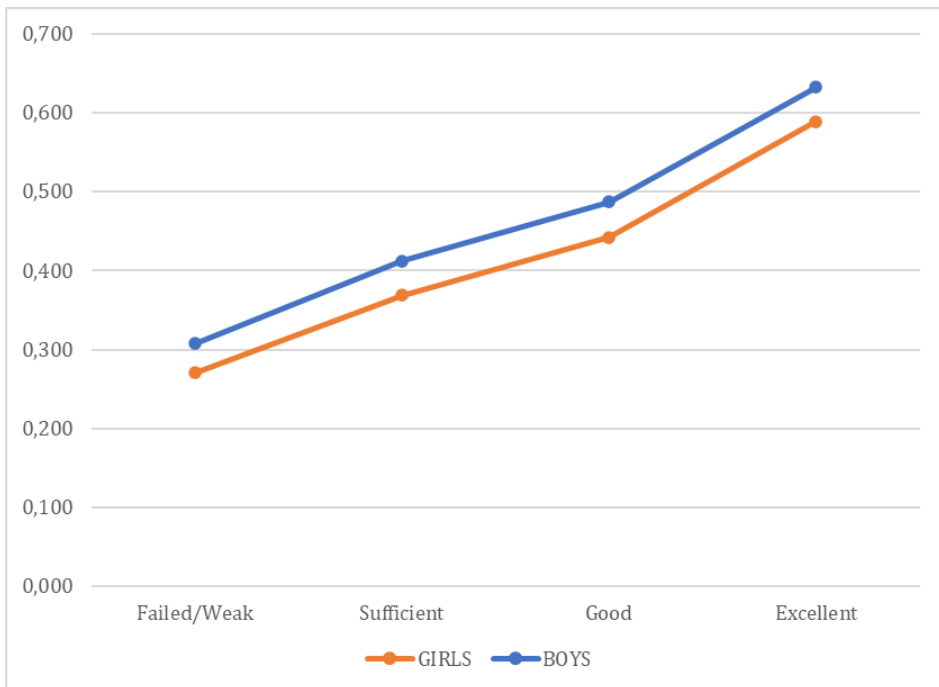


Figure 3–Predictive probability of choosing STEM according to teachers' evaluation in Math by gender -Italy academic year 2019/2020

A visual result that better illustrates the effects shown in the logistic models is the graph of predicted probabilities of enrol in STEM courses at university according to teacher grades received in high school by gender (Figure 3). The probability of choosing STEM majors at university increases steeply as maths grades improve (students with very high mathematics performance, i.e. excellent in the final year of high school have an estimated probability of 0.6 of enrolling in a STEM-university degree, compared to 0.3 for students who have been evaluated as weak in this discipline). A slightly lower, still parallel, trend is observed for girls.

5. DISCUSSION

Although multidisciplinary research on gender and STEM is extensive and ongoing, many aspects of this phenomenon are still unknown, and we have only a partial understanding of the mechanisms that prevent horizontal gender segregation.

A further complexity arises from the fact that the school system and the university system in Italy constitute two markedly distinct institutional contexts. When a student enters the school system, he or she is assigned a number that identifies him or her throughout his or her school career. At the end of the regular school cycle and upon entering the university world, the student receives a new university identification number, which is different from the one used until then. This procedure means that data from the school system and the university system cannot “talk” to each other, as they belong to two different databases. It also excludes the use of the personal identification number, which would make it possible to identify the student, and this would not be allowed under data protection regulations.

The peculiarity of this study lies in the fact that it has been possible to follow students outside the school system and to link them to their choice of university and career.

It is therefore an innovative study that opens the possibility of various more in-depth studies that will also be useful in identifying indicators of predictability of failure in the university world, indicators of inequalities between groups of male and female students, etc., which can then be used at the policy level to intervene promptly in the areas/places where these phenomena are reproduced.

We evaluated the decision-making process from the last year of secondary school to university from a gender perspective, paying particular attention to the most relevant aspects that positively influence the choice of a STEM subject at university. The main finding of the study is that mathematics achievement is a highly influential factor in the choice of STEM subjects, but that the ability as assessed by teachers is much more predictive than the scores obtained in a standardized test designed to measure the same ability.

This effect may be due to two different mechanisms: firstly, the teacher's mark may capture a more multidimensional aspect of the student; secondly, the teacher's mark represents not only a strict evaluation of the student's ability, but also a generalized opinion about the person and his/her potential, which shapes his/her self-esteem and ambitions.

Another important finding is the effect of the mother's educational level on boys and girls in terms of encouraging the choice of STEM studies. For girls, positive effects were found, with a higher propensity to choose STEM if they come from a highly educated family (in terms of mother's title), presumably indicating a weakening of social stereotypes where there are more cultural elements. For boys, on the other hand, the model shows the opposite: those who have a highly educated mother are less likely to end up in a STEM faculty.

This result, that can appear to be contradictory, may represent two side effects of the same phenomenon: the intense social conditioning based on stereotypes, which is relaxed when the cultural level of the family is higher than the norms, allows girls to move towards STEM and boys to leave STEM studies in favour of non-STEM paths.

These findings lend support to the hypothesis that gender stereotypes exert a bidirectional influence, constraining male students' freedom to select non-STEM fields (such as the humanities, social work, or education) in higher education, just as they restrict female students' engagement with STEM disciplines for tertiary education.

Our findings reveal notable gender differences in the factors influencing the choice to enrol in STEM disciplines at university. While teacher-assigned grades play an important role for both male and female students, they appear to have a stronger impact on male students' decisions. Moreover, maternal education, used as a proxy for cultural capital, shows contrasting effects: higher maternal education is associated with a greater likelihood of enrolling in humanities for males, but with an increased likelihood of choosing STEM fields for females. In

addition, we observe significant geographical disparities: female students from southern regions show a lower probability of enrolling in STEM disciplines, while for male students this lower likelihood extends to both central and southern regions. These results partially align with Priulla and Attanasio (2023), who emphasized the complex interplay between gender and academic background in shaping university pathways, highlighting how social and cultural resources differently affect male and female students' educational choices. Similarly, Priulla et al. (2022), using a machine learning approach, demonstrated that high school performance indicators are significant predictors of university enrolment choices, though without explicitly addressing the gender-specific dynamics of teacher-assigned grades or territorial variations that we observe. Finally, Barone and Assirelli (2020) directly investigated gender segregation in higher education, testing multiple explanatory mechanisms including the role of cultural capital, social expectations, and regional contexts. Their study acknowledged the importance of geographical inequalities, but our findings refine this understanding by showing how regional effects intersect with gender, particularly disadvantaging female students from the South in accessing STEM fields. Our results thus add a novel perspective to the understanding of gendered and territorial pathways in higher education.

Furthermore an innovative aspects is given by the fact that findings come from a population study that jointly considers different sets of variables: geographical, economic and cultural background, individual characteristics and mathematics ability, to give an integrated picture of the mechanisms that influence university choice. An important limitation, however, is that we have only looked at enrolment, with no information on subsequent academic careers. For the success of the academic career, the results of the INVALSI test could play a much more important role than the teacher's assessment. In fact, the latter may be influenced by gender stereotypes and risk directing students towards disciplines not based on their best qualities, but on social expectations linked to gender.

More boys enrol in STEM subjects, but it is still unclear whether they complete their studies as quickly and successfully as their female counterparts. Girls, who are more highly selected at entry, may achieve more brilliant results in less time. Future lines of research should look in these directions, focusing on post-enrolment careers and estimating factors for higher 'success' in STEM from a gender perspective.

REFERENCES

- Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19(6), 716–723.
- Allegri, A., Pellegrini, G. and Segafredo, C. (2015). Italian female and male students' choices: STEM studies and motivations. In: E. Henriksen, J. Dillon, J. Ryder (eds) *Understanding Student Participation and Choice in Science and Technology Education*. Springer, Dordrecht.
- Barone, C. and Assirelli, G. (2020) Gender segregation in higher education: An empirical test of seven explanations. *Higher Education: The International Journal of Higher Education Research*, 79, 55–78.
- Borgonovi, F., Choi, A. and Paccagnella, M. (2018). The evolution of gender gaps in numeracy and literacy between childhood and adulthood. *OECD Education Working Papers*, 184, OECD Publishing, Paris.
- Campodifiori, E., Figura, E., Papini, M. and Ricci R. (2010) Un indicatore di status socio-economico-culturale degli allievi della quinta primaria in Italia. *INVALSI Working Paper*, n. 2.
- Carneiro, P., Meghir, C., and Parey, M. (2013). Maternal education, home environments, and the development of children and adolescents. *Journal of the European Economic Association*, 11(s1), 123-160.
- Card, D. and Payne, A.A. (2021), High school choices and the gender gap in stem. *Economic Inquiry*, 59, 9-28.
- Contini, D., Di Tommaso, M.L., Mendolia, S. (2017). The gender gap in mathematics achievement: Evidence from Italian data. *Economic Education Review* 58, 32-42.
- Desimoni, M. (2018). *I livelli per la descrizione degli esiti delle Prove INVALSI*. Roma: INVALSI.
- Eccles, J. S. (2007). Where are all the women? Gender differences in participation in physical science and engineering. In S. J. Ceci, W. M. Williams (Eds.), *Why Aren't More Women in Science? Top Researchers Debate the Evidence* (pp. 199-210). American Psychological Association.
- Eccles, J. S. and Wang, M. T. (2016). What motivates females and males to pursue careers in mathematics and science? *International Journal of Behavioral Development*, 40(2), 100-106.
- Ellison, G. and Swanson, A. (2010). The gender gap in secondary school mathematics at high achievement levels: Evidence from the American

- mathematics competitions. *Journal of Economic Perspectives*, 24 (2), 109-28.
- Else-Quest, N. M., Hyde, J. S. and Linn, M. C. (2010). Cross-national patterns of gender differences in mathematics: A meta-analysis. *Psychological Bulletin* 136, 103–127.
- European Institute for Gender Equality (EIGE), (2021). *Browse Gender Statistics | Gender Statistics Database*. European Institute for Gender Equality
- Hill, C., Corbett, C. and St. Rose, A. (2010). *Why So Few? Women in Science, Technology, Engineering, and Mathematics*. American Association of University Women.
- Ismu (2020). *Alunni con background migratorio in Italia. Le opportunità oltre gli ostacoli*. Fondazione ISMU.
- Makarova, E., Aeschlimann, B. and Herzog, W. (2019) The gender gap in STEM fields: The impact of the gender stereotype of math and science on secondary students' career aspirations. *Frontiers in Education*. 4,60.
- Master, A. and Meltzoff, A. N. (2020). Cultural stereotypes and sense of belonging contribute to gender gaps in STEM. *International Journal of Gender, Science and Technology*, 12(1), 152–198.
- Osborne, J., Shirley, S. and Collins, S. (2003). Attitudes towards science: A review of the literature and its implications. *International Journal of Science Education*, 25(9), 1049-1079.
- OECD (2015), *Immigrant Students at School: Easing the Journey towards Integration*, OECD Reviews of Migrant Education, OECD Publishing, Paris.
- Priulla, A., Albano, A., D'Angelo, N. and Attanasio, M. (2024). Investigating the association between high school outcomes and university enrolment choices: A machine learning approach. In *Proceedings of the Statistics and Data Science 2024 Conference*
- Priulla, A. and Attanasio, M. (2023). Unveiling gender disparities in university pathways: Insights from Italy's master's level. *European Journal of Higher Education*, 13(4), 345–360
- Turner, S.E. and Bowen W.G. (1999). Choice of major: The changing (unchanging) gender gap. *Industrial and Labor Relations Review*, 52(2), 289–313.

INVESTIGATING UNDERGRADUATE AND HIGH SCHOOL STUDENTS' IDENTIFICATION WITH PHYSICS THROUGH STRUCTURAL EQUATION MODELING

Danilo Catena

Department of Mathematics, Physics and Computer science, University of Udine, Udine, Italy

Italo Testa¹

E. Pancini, Department of Physics University of Naples Federico II, Naples, Italy

Abstract. *The disciplinary identity framework has been increasingly used to investigate students' intentions to pursue Science, Technology, Engineering and Mathematics (STEM) related careers. However, much less is known about whether groups of students with different career orientations towards a particular discipline differ in their disciplinary identity. Using physics as a relevant context, the aim of this study was to investigate the validity of an identity framework that includes the following dimensions: identity, interest, recognition and self-efficacy. The analysis was based on a Likert-scale survey conducted online among N = 1135 (female students = 479) Italian high school and undergraduate students, divided into four groups: high school students participating in generic extracurricular vocational activities; high school students participating in specific extracurricular vocational activities focused on physics contents; first-year computer science and biomedical engineering students; first-year undergraduate physics students. We used structural equation modelling to validate the physics identity framework. Differences across groups and between genders were examined by means of multigroup analysis. Results show that for high school students, the effect of self-efficacy on physics identity is fully mediated by interest and recognition. The direct effect of self-efficacy on physics identity is significant only for undergraduate students. Our results show that gender differences in the constructs of the identity model are stronger for students who chose physics as an extracurricular activity or as an undergraduate course. Thus, our results have implications for understanding the mechanism underlying the promotion of students' identity development in physics.*

Keywords: *Pphysics identity, Structural equation modeling, Gender gap*

¹ Corresponding author (italo.testa@unina.it)

1. INTRODUCTION

Despite higher education plays an increasingly fundamental role in Europe in the construction of social resources for the growth of the individuals, the number of tertiary education graduates is still unsatisfactory (Eurostat, 2022). Italy, in particular, has the lowest percentage of young people holding a university degree (e.g., in 2015, 25.2% of the population aged 25–34 years had a university) and is still currently facing a significant shortage of university graduates, which poses a significant challenge for Italy's future economic growth and development (ANVUR, 2023). This trend can be identified also in the Science – Technology – Engineering – Mathematics (STEM) field², despite being a field with a shorter transition into the labour market with respect to all other fields of studies (Rocca and Quintano, 2024).

Research has thoroughly demonstrated that the intention to pursue a STEM-related career is influenced by the extent to which students perceive themselves as a STEM person, namely their perceived STEM *identity* (Calabrese Barton et al., 2013; Carlone and Johnson, 2007; Grimalt-Alvaro et al., 2022). One of the most common models used in science education research posits that identity in a given discipline is predicted mainly by self-efficacy and interest in the discipline, as well as the sense of perceived recognition by others (Hazari et al., 2010; 2020). This identity model has been applied to different disciplines such as biology and chemistry (Potvin and Hazari, 2013), mathematics (Cribbs et al., 2015), computer science (Mahadeo et al., 2020), engineering (Godwin et al., 2016), and STEM (Dou and Cian, 2022).

However, there is still a lack of studies that address how students aged 14–19 develop their STEM identity and whether students with little or no specific experience in a particular discipline differ in their disciplinary identity from those who have already chosen to study that discipline. The main reason for focusing on these questions is that it is at the high school level that students begin to develop a clear idea of their own identity in relation to the discipline by gaining

² In this work, we refer to the discipline categorization reported in *Gazzetta Ufficiale della Repubblica Italiana*, available here:

https://www.gazzettaufficiale.it/do/atto/serie_generale/caricaPdf?cdimg=22A0076300100010110001anddgu=2022-02-07andart.dataPubblicazioneGazzetta=2022-02-07andart.codiceRedazionale=22A00763andart.num=1andart.tiposerie=SG

a deeper understanding of the subject matter related to that discipline and of the job opportunities offered by a particular university course focused on that discipline (Liu et al., 2023; Lockart et al., 2022; Oon and Subramaniam, 2013). Furthermore, exploring how high school students develop their own disciplinary identities may also shed light on the reasons for gender gaps in certain professional areas of STEM, such as physics, where a widespread and generalised underrepresentation of women has been observed in comparison to other sciences – both in university courses and in professional careers – as it has also been highlighted in recent international reports at the European level (EU, 2021; 2022).

To address these issues, we applied the general STEM identity model developed by Dou and Cian (2022) to the discipline of physics, using the same measurement tool to analyse differences in identity development between university and high school students, and to explore whether gender affects identity constructs differently depending on whether students are STEM undergraduates or high school students.

2. BACKGROUND

2.1 THE IDENTITY FRAMEWORK IN THE STEM FIELD

In STEM education, the identity framework has often been used to investigate the extent to which students engage in STEM-related activities and, in particular, to better understand their career choices. In particular, STEM identity has been identified as a significant construct that positively influences students' learning engagement and career aspirations in STEM (Liu et al., 2023), as well as their learning approaches (Hazari et al., 2010). Furthermore, STEM identity has been found to be the most accurate predictor of high school students pursuing a STEM undergraduate major (Cha, 2013). In addition, the literature indicates that at the university level, there is a bidirectional relationship between identity and academic achievement. For example, students who reported higher levels of physics identity tended to get better grades, and students who perform better in their exams generally report higher levels of physics identity at the end of the academic year (Seyranian et al., 2018).

According to this framework, self-identification can be defined as the process by which “people, using the reflexive aspect of the self, name themselves in terms of positional designations or labels” (Burke and Stets, 2009). In other

words, an individual's identity is the result of a dynamic process based on personal and social negotiations with members of the same STEM-related community (Gee, 2000). In this framework, self-identification is considered as a way to characterise a specific aspect of a more general social identity. As such, the process of identity development in a particular discipline is also influenced by the student's other personal identities (e.g., gender) and personal experiences related to the discipline (Kim et al., 2018). For example, identification with physics is a part of the student's identity that only partially overlaps with personal and social identities (Hazari et al., 2010). Therefore, the trajectory through which a student's STEM identity develops can change dynamically over time, as the creation of a personal identity in STEM is a process of recursive social construction (Gee, 2000). This process unfolds in two ways: students shape the identity of the kind of STEM person they want to be through interaction with scientific phenomena and social negotiation of the meaning of STEM concepts (Li, 2012), but also their life experiences with school, perceived sense of belonging and social interactions with relevant others in their family may directly influence their STEM identity (Liu et al., 2023; Verdín, 2021). Previous findings in the STEM field have also shown that when students have a strong sense of identity, they invest more in learning content since they highly value the consequence of such a behaviour, thus maintaining the sustainability of the learning process (Liu et al., 2023).

Based on the above theoretical considerations, to explore the mechanism underlying identity development, most existing models (Dou et al., 2019; Dou and Cian, 2022) adopt a four-factor structure (Fig. 1), in which disciplinary identity, considered as an independently measured construct (Grimalt-Alvaro et al., 2022), is predicted by the dimensions of interest in the discipline, perceived recognition – namely, how others view the student in relation to the discipline – and performance/competence – namely, the *belief* in the ability to perform specific tasks and the ability to understand the content knowledge related to the discipline (Carlone and Johnson, 2007).

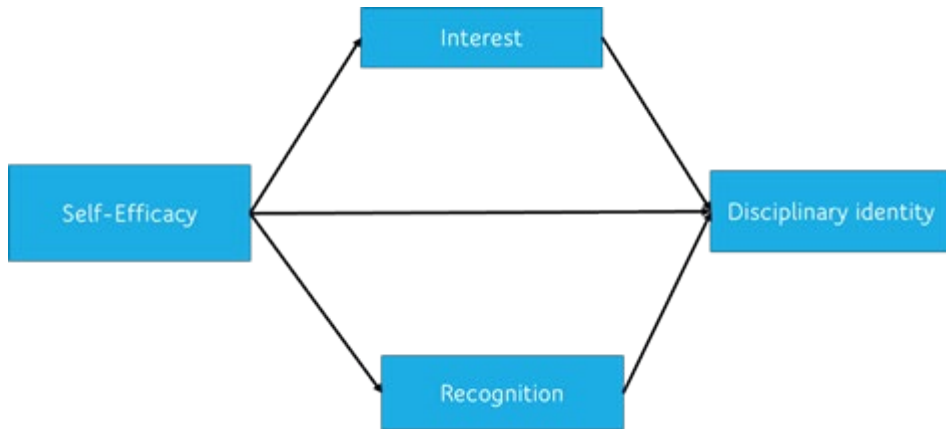


Figure 1: Adopted model of disciplinary identity

The combination of performance and perceived competence into a single category is suggested by previous findings, making this construct closely related to *self-efficacy*. Self-efficacy can be defined as an individual's belief in his or her ability to achieve a certain level of performance or to influence events that affect his or her life (Bandura, 1991). Among those proposed by Bandura, the component of mastery experience, namely the belief in one's ability to handle specific tasks that are similar to previously encountered ones, is particularly relevant for STEM education (Wieselmann et al., 2022).

A large body of research has also shown that interest and self-efficacy – particularly at university level – are strong predictors of STEM identity and disciplinary identity in STEM (Dou and Cian, 2022). Furthermore, findings from recent longitudinal studies are consistent with reciprocal and time-dependent correlational relationships between self-efficacy and academic performance, and between perceived recognition and academic performance (Bottomley et al., 2023). This confirms the 'feedback loop' hypothesized by Kalender and colleagues (Kalender et al., 2019) and is consistent with the finding that early performance influences academic self-concept and past self-concept influences future performance (Marsh et al., 2002).

In the case of physics, which is the focus of this study, previous studies have also thoroughly confirmed that individuals' self-identification with physics is predicted by their interest, sense of recognition, and – indirectly – by their self-

efficacy, and that the structural relationships in the models are moderated by gender (Hazari et al., 2010; 2020).

2.2 THE ROLE OF GENDER IN STEM AND PHYSICS IDENTITY

While many studies in the Italian context have addressed the gender gap at university level in STEM field (Barone and Assirelli 2020; Priulla and Attanasio, 2023), the role gender in STEM identity, and specifically physics identity, has been studied mainly in the US context. Specifically, previous studies have reported that women report significantly lower STEM and physics identity than men, suggesting that this perception of the self may contribute to the underrepresentation of women in physics (Dou and Chan, Hazari et al., 2010). Such underrepresentation may in turn affect one's identity, as previous studies have shown that students who belong to underrepresented groups experience a lower sense of belonging than their peers, with a consequent negative impact on self-efficacy and performance in disciplinary coursework, which may affect their disciplinary identity construction over time (Liu et al., 2023). Other studies have also shown that the physics identity of undergraduate physics students declines over time, but this decline does not significantly interact with gender (Bottomley et al., 2023). However, there does appear to be a significant interaction between gender and academic semester in perceived recognition, as gender differences in perceived recognition were greater in the second semester than in the first, regardless of undergraduate year.

Finally, recent results have shown that gender stereotypes can influence individuals' disciplinary identity (Galano et al., 2023; Liu et al., 2023), especially in physics (Marchand and Taasobshirazi, 2013). For example, stereotypical views of physics typically include perceptions of masculinity (Kessels et al., 2006). Thus, it is likely that women do not identify with physics as a discipline to the same extent as men, not only because of their lower levels of self-efficacy and perceived recognition (Bottomley et al., 2023; Kalender et al., 2019), but also because they do not perceive physics as fitting their female identity.

3. RESEARCH QUESTIONS

While literature has thoroughly validated the model of physics identity for undergraduate students, there is a lack of research on whether the model also

holds for secondary school students. In particular, there is little evidence on whether the structural relationships of the model are moderated by different types of out-of-school time experiences focused on physics. Furthermore, there is little evidence as to whether the model is also valid for students enrolled in a typical undergraduate STEM course other than physics (e.g. engineering).

Therefore, the specific research questions that guided this part of the study were:

RQ1. Does the model reported in Figure 1 accurately describe the physics identity of secondary and undergraduate students?

RQ2.a Does the physics identity construct differ between secondary school students who attended out-of-school activities in physics and other students?

RQ2.b Does the physics identity construct differ between physics and engineering undergraduates?

Finally, given the role of gender in identity construction, especially in physics where there is a significant under-representation of women in professional careers compared to other STEM fields, we also investigated whether the gender variable moderates the relationships between the constructs of the physics identity model (self-efficacy, interest, perceived recognition and identity) for the sample of secondary school and undergraduate students we included in the study.

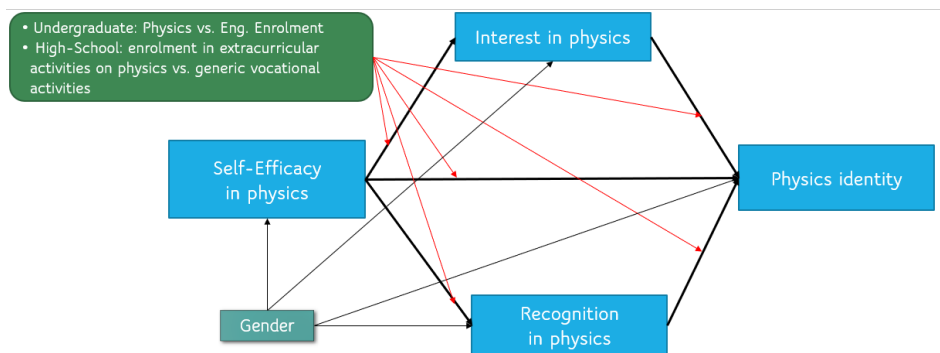


Figure 2: Modified model of disciplinary identity

Therefore, the final research question that guided our study was:

RQ3. Does gender affect the relationships of the physics identity construct? If so, what are the differences between secondary school students who attended extra curriculum activities in physics and other students who did not attend these activities and between physics undergraduates and engineering undergraduates?

Figure 2 summarises the hypothesized effects of the out-of-school activities, the chosen undergraduate course and gender.

4. METHODS

4.1 SAMPLE

Overall, 1135 Italian students participated in the study and completed the final survey (see next section). The sample consisted of four groups:

G1) 169 high school students (female students: 48.5%) who participated in general out-of-school vocational activities (duration = 15 hours);

G2) 177 high school students (female students: 59.9%) who participated in specific out-of-school vocational activities (duration = 15 h) focused on physics contents (quantum mechanics, astrophysics, optics);

G3) 427 first-year students of computer science and biomedical engineering (female students: 37.5%);

G4) 362 first-year physics students (female students: 36.2%).

All groups of high school students (G1 and G2) attended school streams where physics is a compulsory subject, with the same curriculum and the same amount of time devoted to physics (4 hours per week). G3 and G4 students had already attended an introductory physics course of 48 and 96 hours, respectively. The percentage of female students in the G3 and G4 groups reflects that of these university courses at national level.

4.2 INSTRUMENT

We used a 12-item instrument with a 5-point Likert scale: not at all, not very, fairly, mostly, completely. The instrument aimed at measuring the constructs of the model in Figure 1, and specifically:

Physics identity: we used a single item: *do you see yourself as a physics person?* from Hazari et al. (2010). This item has been used in several studies, both for physics (Bottomley et al., 2023) and STEM identity (Dou and Cain, 2022), as it was found to be a good proxy for overall physics identity (Potvin and Hazari, 2013).

Perceived recognition in physics: we used a 2-item scale that measured the perceived recognition as a physicist from others, namely teachers and classmates using the same structure for identity measurement: *'do ... see you as a physics person?'*

Interest in physics: we used 3 items to investigate to what extent students were interested in topics and research results in physics, as well as in physics-related hobbies.

Self-efficacy in physics: we used 6 items targeting how students rate themselves in: explaining the topics they studied (1 item); solving 'easy' and 'difficult' problems (2 items); designing an experiment and performing it alone/guided by teachers (3 items).

The data collection took place between 2020 and 2021 in remote teaching modality. Students completed the survey online and were informed that their participation was voluntary. Gender was indicated by the students themselves and was added to the survey prior to anonymizing the data for analysis.

4.3 DATA ANALYSIS

First, we explored how the four involved groups differed in their physics identity by means of an analysis of variance (ANOVA). Post-hoc analysis was then carried out to inspect differences between the four groups. Gender differences in each group were first explored through a series of *t-tests* and then investigated through a 2-way ANOVA testing the interaction between the group variable (categorical with 4 values) and the gender variable (binary, 1 = female student, 0 = male student).

Then, to answer our research questions, we performed a structural equation modeling (SEM) based on the relationships hypothesized in Figure 2. SEM is a theory-driven statistical approach aimed at: testing hypotheses; inferring causal relationships; estimating direct and indirect effects. In other words, with SEM it is possible to test if a model in which well-determined relationships between variables are consistent with the empirical data. As in the confirmatory factor analysis, a χ^2 test assesses how consistent the model is with empirical data (acceptable values of $\chi^2/\text{d.o.f.} < 3$) and it is also used to calculate the goodness indices of the fit (Schreiber et al., 2006). In this study, we also used two further

indices, namely the comparative fit index (CFI), which compares the hypothesized model with the null model, and the Tucker – Lewis index (TLI), which compares a χ^2 calculated on the hypothesized model with a χ^2 calculated on the null model, i.e. without the item-factor correlations. Both CFI and TLI should be above 0.95 (Hooper et al., 2008). To establish goodness of fit, we also used the root mean squared error of approximation (RMSEA), which refers to the estimate of the approximation error of the correlation matrix between the items of the instrument. The RMSEA index addresses the problem of the strong dependency of the χ^2 test when the sample size N is high. Typically, acceptable values of RMSEA for a good-model fit are lower than 0.08 (Hooper et al., 2008; Schreiber et al., 2006). Finally, we also assessed convergent validity, which indicates whether the dominant latent factor is extracted from the items, by calculating the average variance extracted (AVE) for each latent factor of the model. AVE values greater than 0.50 indicate that more than half of the variance of the items is explained by the latent factor.

To explore the role of extracurricular activities and of the chosen undergraduate course as moderating variables (see Figure 2), we conducted two separate multigroup analyses, comparing the G1 and G2 groups and the G3 and G4 groups, respectively. For both multigroup analyses, we hypothesized that the role of gender would influence the four variables in the models (self-efficacy, interest, recognition, identity). The significance of indirect effects was assessed using the bootstrap bias-corrected percentile method, which provides confidence intervals for the effect estimate. The magnitude of indirect effects was assessed by the ratio of the indirect effect to the direct effect (Sobel, 1982). To estimate the parameters of the hypothesized model and the goodness of fit indices, we followed the covariance-based approach using the maximum likelihood method in IBM SPSS Amos 28.

5. RESULTS

5.1 DIFFERENCES IN PHYSICS IDENTITY ACROSS GROUPS

Table 1 reports the descriptive statistics and correlations for all measured variables. For the self-efficacy, interest and recognition scales we report the average scores. Cronbach's alpha of each scale is also reported on the table's diagonal.

Table 1. Complete descriptive statistics for the measured variables (N = 1135)

Variable	1	2	3	4
1. Physics identity	-	-	-	-
2. Self-efficacy	.371**	0.86 ^a	-	-
3. Interest	.573**	.441**	0.84 ^a	-
4. Recognition	.492**	.490**	.464**	0.88 ^a
Mean	2.87	3.45	3.65	3.32
SD	1.08	0.69	0.82	0.83
Kurtosis	-0.567	-0.137	-0.099	0.155
Asymmetry	0.143	0.036	-0.342	-0.114

^a Cronbach's alpha; ** $p < .01$

Note that the physics identity scale has no Cronbach's alpha associated since it is a single item scale. Table 2 reports the mean scores for female and male students for the four groups involved.

The results of the ANOVA show that the four groups significantly differed for their level of physics identity, Welch's $F(3, 482.855) = 60.596$, $p < .001$, $\eta^2 = .13$. Post-hoc analysis shows that differences between the groups are all statistically significant, with physics students (G4) scoring highest (mean = 3.33, SD = 1.07), followed by the high school students who followed also out-of-school activities focused on physics (G2, mean = 3.15, SD = 1.08), the engineering students (G3, mean = 2.60, SD = 0.99) and the high school students who followed generic out-of-school activities (G1, mean = 2.30, SD = 0.84).

Table 2. Mean scores of the measured variables for the four groups of the sample

Variable	G1 (N = 169)	G2 (N = 177)	G3 (N= 427)	G4 (N=362)
1. Physics identity	2.30	3.15	2.60	3.33
Female students	2.23	3.00	2.60	3.04
Male students	2.37	3.38	2.60	3.49
<i>t</i>	-1.408	-2.265*	-0.030	-3.929***
Cohen's <i>d</i>	-0.16 ^a	-0.36 ^b	-0.00 ^a	-0.43 ^b
2. Self-Efficacy	3.05	3.72	3.67	3.22
Female students	2.94	3.58	3.63	3.11
Male students	3.16	3.94	3.70	3.29
<i>t</i>	-2.352*	-3.70***	-1.124	-2.629**
Cohen's <i>d</i>	-0.36 ^b	-0.57 ^c	-0.11 ^a	-0.29 ^b
3. Interest	3.01	3.77	3.68	3.87
Female students	2.86	3.63	3.81	3.72
Male students	3.15	3.97	3.60	3.95
<i>t</i>	-2.610*	-2.407*	2.878**	-2.915**
Cohen's <i>d</i>	-0.40 ^b	-0.37 ^b	0.29 ^b	-0.32 ^b
4. Recognition	2.94	3.80	3.37	3.19
Female students	2.90	3.71	3.56	3.02
Male students	2.96	3.95	3.27	3.29
<i>t</i>	-0.311	-2.181*	3.378***	-3.54
Cohen's <i>d</i>	-0.05 ^a	-0.33 ^b	0.34 ^b	-0.39 ^b

* $p < .05$; ** $p < .01$; *** $p < .001$; Cohen's *d* magnitude: a $\rightarrow$ negligible (0.0-0.2); b $\rightarrow$ small (0.2-0.5); c $\rightarrow$ moderate (0.5-0.8).

5.2 GENDER DIFFERENCES IN PHYSICS IDENTITY AND GENDER - GROUP INTERACTION

The *t-test* statistics show that, independently on the group, female students have lower physics identity and lower self-efficacy than their male counterpart. The magnitude of such differences is moderate for the G2 and G4 groups, namely high school students involved in out-of-school activities focused on physics and physics undergraduates, respectively. Girls of G1, G2 and G4 groups also scored significantly lower than boys in the interest and recognition scales, while for the G3 group girls report higher interest and perceive higher recognition.

Results of the 2-way ANOVA on physics identity show a significant effect of the group variable, $F(3; 1127) = 50.843$; $p < .001$; $\eta^2 = .12$, as well as of gender, $F(1; 1127) = 13.598$, $p < .001$, $\eta^2 = .01$. The interaction between gender and groups is also statistically significant, $F(3; 1127) = 3.515$; $p = .015$; $\eta^2 = .01$. Simple effects analysis shows that differences between groups are significant for both male students, $F(3; 1127) = 47.242$; $p < .001$, and female students, $F(3; 1127) = 14.355$, $p < .001$, although the effect is greater for male students ($\eta^2 = .11$) than for female student ($\eta^2 = .04$). Simple effect analysis within each group show that gender differences are stronger for the G4 group, $F(1; 1127) = 16.983$, $p < .001$, $\eta^2 = .015$ and the G2 group, $F(1; 1127) = 6.141$, $p = .013$, $\eta^2 = .005$, while there are no significant gender differences for the G1 and G3 groups.

5.3 ASSESSMENT OF THE DISCIPLINARY IDENTITY MODEL THROUGH STRUCTURAL EQUATION MODELING

The results of the SEM analysis carried out for the first-year students (G3 and G4) and the high school students (G1 and G2) are shown in Figures 3 and 4 respectively. The AVE values for the three predictors of disciplinary identity, namely self-efficacy, interest and recognition by others are: 0.53, 0.65 and 0.79, thus convergent validity is confirmed for both models. The variables reported in rectangles represent the items of the instrument, except for those related to self-efficacy scale, for which we performed item parcelling for a more efficient parameter estimation (Bandalos, 2002). Specifically, we averaged the two items about students' perceived self-efficacy in solving 'easy' and 'difficult' problems, and the three items about students' perceived self-efficacy in designing an experiment and performing it alone/guided by teachers.

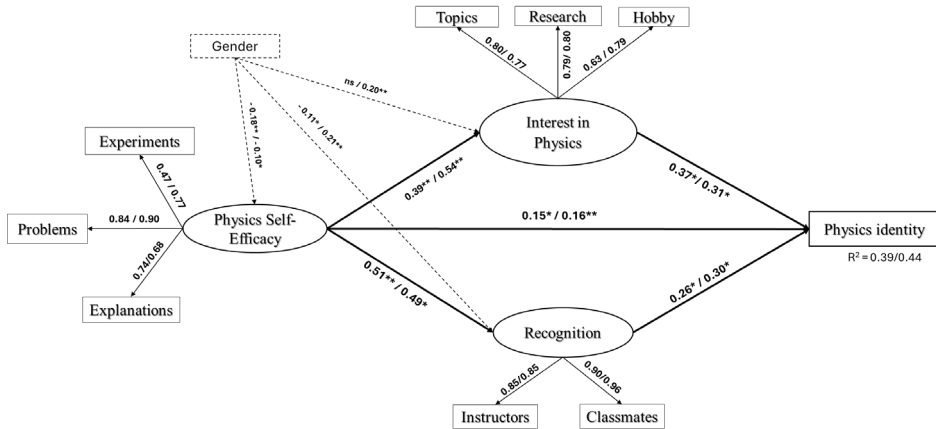


Figure 3. Standardized parameters estimate of relations of the physics identity model for undergraduate students. Results are reported in format *Physics/Engineering*. Note: * $0.01 < p \leq 0.05$; ** $0.001 < p \leq 0.01$

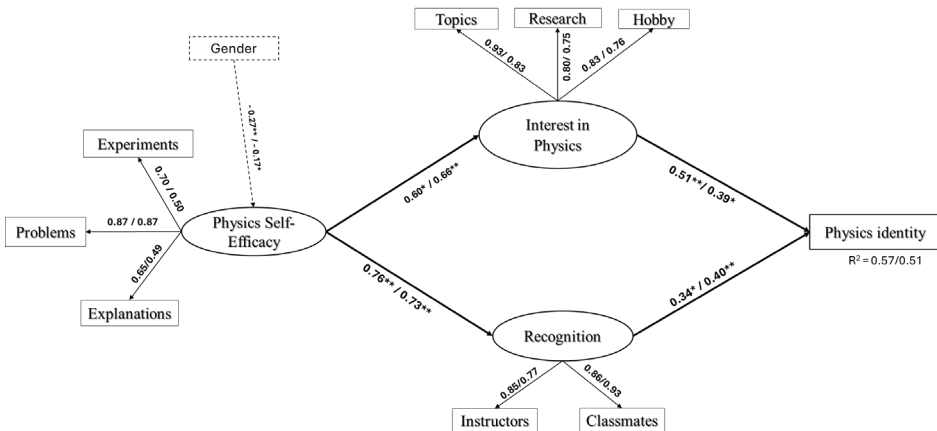


Figure 4. Standardized parameters estimates of relations of the physics identity model for high school students. Results are reported in format *Physics out-of-school/Generic vocational activities*. Note: * $0.01 < p \leq 0.05$; ** $0.001 < p \leq 0.01$

Both models have optimal fit indices (Model for G1 and G2 groups: $\chi^2/df = 1.711$, $p < 10^{-3}$, RMSEA = 0.045, CFI = 0.97, TLI = 0.96; Model for G3 and G4 groups: $\chi^2/df = 2.510$, $p < 10^{-3}$, RMSEA = 0.044, CFI = 0.97, TLI = 0.96). The

model for G1 and G2 explains 57% and 51% of the variance of the physics identity, respectively, and shows measurement and structural invariance, $\Delta\chi^2(7) = 9.171$, $p = 0.241$, $\Delta\chi^2(10) = 12.005$, $p = 0.285$, respectively. Differently, the model for G3 and G4 groups explains 39% and 44% of the variance of the physics identity, respectively, while measurement and structural invariance are not supported, $\Delta\chi^2(8) = 22.098$, $p = 0.005$, $\Delta\chi^2(14) = 55.746$, $p < 0.001$, respectively.

Unstandardised indirect effects are reported in Table 3.

Table 3. Unstandardised indirect effects for SEM diagrams of Figures 3 and 4.

Path	<i>B</i>	CI	<i>R_M</i>
Self-efficacy → Interest → Identity			
<i>G1</i>	0.55*	[0.278 – 1.008]	7.64
<i>G2</i>	0.60*	[0.327 – 0.889]	3.33
<i>G3</i>	0.28*	[0.169 – 0.425]	1.06
<i>G4</i>	0.37**	[0.231 – 0.577]	0.95
Self-efficacy → Recognition → Identity			
<i>G1</i>	0.62**	[0.363 – 1.149]	8.61
<i>G2</i>	0.51*	[0.290 – 0.848]	2.83
<i>G3</i>	0.25*	[0.160 – 0.335]	0.94
<i>G4</i>	0.34*	[0.161 – 0.544]	0.88

Note: CI = Confidence Interval; *R_M* = Ratio of unstandardised indirect to unstandardised direct effect; * $p < 0.05$; ** $p \leq 0.01$. G1 = high school students who participated in generic out-of-school vocational activities; G2 = high school students who participated in specific out-of-school vocational activities focused on physics contents; G3 = freshman computer science and biomedical engineering students; G4 = freshman undergraduate physics students.

Indirect effects of self-efficacy are significant for the path mediated by interest and the path mediated by recognition, for all groups. However, self-efficacy has a significant direct positive effect on physics identity only for

undergraduate students. After accounting for such mediating paths, gender has no significant direct effect on physics identity for all groups in the sample, but its effect is fully mediated by self-efficacy (all groups) and recognition (G3 and G4 groups). Specifically, gender has a negative effect on self-efficacy and recognition for the G4 group (physics students), while it has a weak negative effect on self-efficacy and a positive effect on recognition and interest for the G3 group (computer science and biomedical engineering students). This means that, in general, being a female student has a negative impact on perceived self-efficacy in physics, regardless of the chosen undergraduate course. In contrast, being a female engineering student increases both interest and recognition in physics. Similarly, for high school students, gender has a negative effect on self-efficacy of both groups. This means that, in general, being a female student has a negative impact on perceived self-efficacy in physics, regardless of attending extracurricular activities.

6. DISCUSSION

In the following, we briefly describe the results according to the research questions of the study.

6.1 RQ1. Does the model reported in Figure 1 accurately describe the physics identity of secondary and undergraduate students?

Our results confirm that the disciplinary identity model in Figure 1, developed from previous research, accurately describes the physics identity of the students in our sample, namely high school students with different career orientations towards physics and university students who had chosen two different STEM subjects, physics and engineering, respectively. In particular, our results support the use of three latent constructs, namely interest, recognition and self-efficacy, to measure physics identity for our population. Consistent with previous models (Dou and Cian, 2022; Hazari et al., 2010), our results support a significant direct effect of interest and recognition on physics identity and a significant indirect effect of self-efficacy mediated by interest and recognition.

However, some differences in the strength of structural relationships emerge when comparing our findings with previous research. First, in our case, the recognition construct makes a smaller contribution to physics identity, contrary

to what happens for STEM identity in Dou and Cian (2022). As physics is a specific area of STEM, we can argue that being interested in physics may play a more important role in contributing to identity development than recognition from others, as classmates or other students may be perceived as poorly informed about one's physics knowledge (Kalender et al., 2019). The same argument may apply to recognition from teachers or lecturers, as their judgement may be perceived as less relevant, given that the high school and undergraduate students involved in the study already had an interest in STEM (Starr et al., 2020). The second difference is that self-efficacy has a direct effect on physics identity only for undergraduate students. This finding may be related to our decision to include in the survey items on problem solving, oral presentation, and conducting experiments, which may be more familiar to undergraduate STEM students than to high school students.

6.2 RQ2.a Does the physics identity construct differ between secondary school students who attended out-of-school activities in physics and other students?

We found that, from a measurement and structural point of view, our model of identity is invariant with respect to the type of extracurricular activities engaged in, despite significant differences in identity scores between the two groups of students. This means that interest and recognition by others significantly predict physics identity independently of involvement in physics activities. Furthermore, self-efficacy has only an indirect effect on identity, which is fully mediated by both interest and recognition. This is in line with the findings of Dou and Chan (2021), who found that self-efficacy does not directly influence STEM identity. This result can be interpreted by considering that the self-efficacy items were related to specific teaching practices in physics (e.g. designing an experiment), which are hardly implemented at high school level.

6.3 RQ2.b Does the physics identity construct differ between physics and engineering undergraduates?

Our results support the validity of the proposed physics identity model for undergraduate students. In particular, self-efficacy, interest and recognition independently predict career identity, but self-efficacy also has a direct effect on identity for undergraduate physics students. The different role played by self-

efficacy is in line with previous studies which have shown that self-efficacy has an impact on the identity of the person in a particular discipline (Stout et al., 2011). The significant differences between physics and engineering students in the measurement and structural model can be explained by taking self-efficacy into account. In particular, we found that self-efficacy significantly affects interest, recognition and identity, as measured by our items that target the perceived ability to solve problems, design and perform an experiment, and orally discuss curriculum content. This perception may be very different, as physics students may be more familiar with these practices than engineering students.

6.4 RQ3 Does gender affect the relationships of the physics identity construct? If so, what are the differences between secondary school students who attended extra curriculum activities in physics and other students who did not attend these activities and between physics undergraduates and engineering undergraduates?

Our results show that gender does not directly affect physics identity, but only indirectly. For all students in the sample, gender negatively affects self-efficacy, i.e. female students are less confident in their knowledge, while for undergraduates, gender also affects recognition, with different directions, positive for engineering students, negative for physics students. Another direct effect, on interest, was only observed for engineering students. Overall, our results are consistent with previous studies in physics education, which have shown that even when learning experiences are similar, stereotypical associations of competence with men can lead women to be underconfident about their performance (Galano et al., 2023; Li and Singh, 2021). Notably, in the physicist group, gender negatively affects self-efficacy and recognition, whereas in the engineering group, gender negatively affects self-efficacy but positively affects recognition and interest. A possible explanation for this result is that, in the engineering courses included in our sample (computer science and biomedical engineering), female students are more interested in physics than boys and are recognised as better students by classmates and teachers due to the different type of school attended, namely technical school or scientific lyceum. However, further studies with a more representative sample are needed to support our interpretation.

7. CONCLUSIONS AND IMPLICATIONS

Our results provide further evidence for the validity of the three-factor structure of identity predictors (Hazari et al., 2010; Hazari et al., 2020; Dou and Cian, 2022), while supporting the use of our instrument to measure physics identity. However, due to the use of a convenience sample, better sampling approaches are needed to confirm the emerging direct and indirect effects. Furthermore, the model should be improved by including other predictors and mediators in the model that take into account the cultural milieu in which the identity develops (e.g., utility value of physics, family background, socio-economic status, school context). Following Gee (2000) and Bottomley (2023), this would also require longitudinal studies, although this construct appears to be quite stable over time (Starr et al., 2020).

These findings also have implications for high school teachers and university instructors, in particular the role they can play in influencing perceived recognition as a physicist as a mediator of physics identity. As suggested by Wang and Hazari (2018), explicit and implicit attempts to increase students' perceptions of being recognised 'as a physicist' can be internalised by students. This can take the form of explicitly telling students that they are capable of setting tasks that make students feel recognised without explicitly telling them so (Bottomley et al., 2023). However, further research is needed to explore whether these findings can be extended to undergraduate students.

Some limitations of the study have to be highlighted. First, the convenience sample was based on a voluntary online participation, and this could have biased the results. Future studies may have investigated if the relationship found in this study can be extended to a representative sample of the population involved in this study. Second, reproducibility of the results should be investigated, as there is a lack of this kind of quantitative research involving Italian students. Furthermore, the role of confounding variables – such as performance in physics or social indicators – as mediator in the physics identity model should be considered as a further research step. Third, the study was carried out during remote teaching due to the COVID-19 pandemic, and this may have affected the measured variables. Specifically, in one of our previous studies in the Italian university context (Marzoli et al., 2021), we found that the transition to remote teaching led to a significant decrease in interest towards physics as well as in the students' perceived self-efficacy in physics. Hence, further research is needed to find whether the decrease in these dimensions affected also the relationships of

physics identity model. Finally, we focused on the classification of gender into two categories (male and female). While we recognise this as a limitation of the present study due to adequately power our analysis, the evolving experience of those students who identify as non-binary is an interesting question for future research.

REFERENCES

- ANVUR (2023). Rapporto sul sistema della formazione superiore e della ricerca. Retrieved from <https://www.anvur.it/wp-content/uploads/2023/06/Sintesi-Rapporto-ANVUR-2023.pdf>
- Bandalos, D. L. (2002). The effects of item parceling on goodness-of-fit and parameter estimate bias in structural equation modeling. *Structural Equation Modeling: A Multidisciplinary Journal*, 9,(1), 78–102. doi:10.1207/S15328007SEM0901_5
- Bandura, A. (1991) Self-efficacy mechanism in physiological activation and health-promoting behavior. In J. Madden IV (ed.) *Neurobiology of Learning, Emotion and Affect* , 229–269, Raven, New York.
- Barone, C., and Assirelli, G. (2020) Gender segregation in higher education: An empirical test of seven explanations.” *Higher Education*, 79, 55–78. <https://doi.org/10.1007/s10734-019-00396-2>
- Bottomley, E., Kohnle, A., Mavor, K., Miles, P. and Wild, V. (2023). The relationships between gender and performance in undergraduate physics students: The role of physics identity, perceived recognition and self-efficacy. *European Journal Physics*, 44, 025701.
- Burke, P. J. and Stets, J. E. (2009). *Identity Theory*. Oxford University Press.
- Calabrese Barton, A., Kang, H., Tan, E., O'Neill, T. B., Bautista-Guerra, J. And Brecklin, C. (2013). Crafting a future in science: Tracing middle school girls' identity work over time and space. *American Educational Research Journal*, 50, (1), 37–75. <https://doi.org/10.3102/0002831212458142>.
- Carlone, H. B. and Johnson, A. (2007). Understanding the science experiences of successful women of color: Science identity as an analytic lens. *Journal of Research in Science Teaching*, 44, 8, 1187–1218.
- Cha, Y. (2013). Overwork and the persistence of gender segregation in occupations, *Gender and Society*, 27, 158–184.

- Cheryan, S., Ziegler, S. A., Montoya A. K. and Jiang, L. (2017). Why are some STEM fields more gender balanced than others? *Psychological Bulletin* 143, (1), 1-35.
- Cribbs, J., Hazari, Z., Sonnert G. And Sadler, P. M. (2015). Establishing an explanatory model for mathematics identity. *Child Development* 86,(4), 1048-1062.
- Dou R., Hazari Z., Dabney K., Sonnert G., and Sadler P., (2019). Early informal STEM experiences and STEM identity: The importance of talking science. *Science Education*, 103, (3), 623 – 637
- Dou, R., and Cian, H. (2022). Constructing STEM identity: An expanded structural model for STEM identity research. *Journal of Research in Science Teaching*, 59,(3), 458–490.
- European Commission (EU, 2021). *She Figures 2021: Gender in Research and Innovation: Statistics and Indicators*. Luxembourg: Publication office of the European Union. <https://data.europa.eu/doi/10.2777/06090>
- European Commission (EU, 2022). *2022 Report on Gender Equality in the EU* <https://doi.org/10.2838/94579>
- Eurostat (2016). *Tertiary Education Statistics*. Retrieved from: https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Tertiary_education_statistics/it
- Galano, S., Liccardo, A., Amodeo, A. L., Crispino, M., Tarallo, O. and Testa, I. (2023). Endorsement of gender stereotypes affects high school students' science identity. *Physical Review Physics Education Research*, 19, 010120.
- Gee, J. (2000). Identity as an analytic lens for research in education. *Review of Research in Education*, 25, 99–125.
- Godwin, A., Potvin, G., Hazari Z. and Lock R. (2016) Identity, critical agency, and engineering: An affective model for predicting engineering as a career choice, *Journal of Engineering Education*, 105, 312-340.
- Grimalt-Alvaro, C., Couso, D., Boixadera-Planas, E. and Godec, S. (2022). I see myself as a STEM person. Exploring high school students' self-identification with STEM. *Journal of Research in Science Teaching*, 59, (5), 720–745.
- Hazari Z., Chari, D., Potvin, G. and Brewé, E. (2020). The context dependence of physics identity: Examining the role of performance/competence, recognition, interest, and sense of belonging for lower and upper female

- physics undergraduates, *Journal of Research in Science Teaching*, 57, 1583–1607.
- Hazari, Z., Sonnert, G., Sadler, P. M. and Shanahan, M.-C. (2010). Connecting high school physics experiences, outcome expectations, physics identity, and physics career choice: A gender study. *Journal of Research in Science Teaching*, 47, (8), 978–1003.
- Hooper, D., Coughlan, J. and Mullen, M. R. (2008). Structural equation modelling: Guidelines for determining model fit. *Electronic Journal of Business Research Methods*, 6, (1), 53–60
- ISTAT (2018). *Livelli di istruzione e ritorni occupazionali*. Retrieved from: https://www.istat.it/it/files/2019/07/Report-Livelli-di-istruzione-e-ritorni-occupazionali_2018.pdf
- Kalender, Z. Y., Marshman, E., Schunn, C. D., Nokes-Malach, T. J. and Singh, C. (2019). Gendered patterns in the construction of physics identity from motivational factors. *Physical Review Physics Education Research*, 15, 020119
- Kalender, Z. Y., Marshman, E., Schunn, C. D., Nokes-Malach, T. J. and Singh, C. (2019). Why female science, technology, engineering, and mathematics majors do not identify with physics: They do not think others see them that way, *Physical Review Physics Education Research*, 15, 020148
- Kessels, U., Rau, M. and Hannover, B. (2006). What goes well with physics? Measuring and altering the image of science. *British Journal of Educational Psychology*, 76, 761–780.
- Kim, A. Y., Sinatra, G. M. and Seyranian, V. (2018). Developing a STEM identity among young women: A social identity perspective. *Review of Educational Research*, 88, (4), 589–625.
- Li, S. L. (2012). *Learning in a Physics Classroom Community: Physics Learning Identity Construct Development, Measurement and Validation*, Degree of Doctor of Philosophy Granting Oregon State University, Corvallis, Oregon, USA,.
- Li, Y. and Singh, C. (2021) Effect of gender, self-efficacy, and interest on perception of the learning environment and outcomes in calculus-based introductory physics courses. *Physical Review Physics Education Research*, 17 010143
- Liu, S., Xu, S., Li, Q., Xiao, H., Zhou, S. (2023). Development and validation of an instrument to assess students' science, technology, engineering, and

- mathematics identity. *Physical Review Physics Education Research*, 19, 010138.
- Lockhart, M. E., Kwok, O. M., Yoon, M. and Wong, R. (2022). An important component to investigating STEM persistence: The development and validation of the science identity (SciID) scale. *International Journal of STEM Education*, 9, 34
- Mahadeo, J., Hazari, Z. and G. Potvin (2020). Developing a computing identity framework: Understanding computer science and information technology career choice. *ACM Transactions on Computer Education*, 20, 1.
- Marchand, G. C. and Taasobshirazi, G. (2013). Stereotype threat and women's performance in physics. *International Journal of Science Education*, 35, 3050–3061
- Marsh, H.W., Hau, K. T. and Kong C. K. (2002). Multilevel causal ordering of academic self concept and achievement: Influence of language of instruction (English compared with Chinese) for Hong Kong students. *American Educational Research Journal*, 39, 3, 727–763.
- Marzoli, I., Colantonio, A., Fazio, C., Giliberti, M., Scotti di Uccio, U. and Testa I. (2021). Effects of emergency remote instruction during the COVID-19 pandemic on university physics students in Italy. *Physical Review Physics Education Research*, 17, 020130
- Miller, C. J. and Crouch, J. G. (1991). Gender differences in problem solving: Expectancy and problem context. *The Journal of Psychology*, 125, (3), 327-336
- Oon, P. T. and Subramaniam, R. (2013). Factors influencing Singapore students' choice of physics as a tertiary field of study: A Rasch analysis, *International Journal of Science Education*, 35, 86 – 118
- Potvin, G. and Hazari, Z. (2013). The development and measurement of identity across the physical sciences. Paper presented at *Physics Education Conference 2013*, Portland, OR, <https://doi.org/10.1119/perc.2013.pr.058>
- Priulla, A., and Attanasio, M. (2023). Unveiling gender disparities in university pathways: insights from Italy's master's level. *European Journal of Higher Education*, 14(4), 574–597. <https://doi.org/10.1080/21568235.2023.2247189>
- Rocca, A. and Quintano, C. (2024). Success stems from STEM fields: An analysis of Italian graduates. *Italian Economic Journal*, 10, 761–793 <https://doi.org/10.1007/s40797-023-00255-1>

- Schreiber, J. B., Nora, A., Stage, F. K., Barlow, E. A. and King, J. (2006). Reporting structural equation modeling and confirmatory factor Analysis results: A review. *Journal of Educational Research*, 99, 6, 323 – a37
- Seyranian, V., Madva, A., Duong, N., Abramzon, N., Tibbetts, Y. and Harackiewicz, J. M. (2018). The longitudinal effects of STEM identity and gender on flourishing and achievement in college physics. *International Journal of STEM Education*, 5, 1–14.
- Sobel, M. E. (1982). Asymptotic confidence intervals for indirect effects in structural equation models. *Sociological Methodology*, 13, 290–312
- Starr, C. R., Hunter, L., Dunkin, R., Honig, S., Palomino, R. and Leaper, C. (2020). Engaging in science practices in classrooms predicts increases in undergraduates' STEM motivation, identity, and achievement: A shortterm longitudinal study. *Journal of Research in Science Teaching*, 57, (7), 1093–1118. <https://doi.org/10.1002/tea.21623>.
- Stout, J. G., Dasgupta, N., Hunsinger, M. and McManus, M. A. (2011). STEMing the tide: Using ingroup experts to inoculate women's self-concept in science, technology, engineering, and mathematics (STEM). *Journal of Personality and Social Psychology*, 100,(2), 255–270. <https://doi.org/10.1037/a0021385>
- Verdín, D., (2021). The power of interest: Minoritized women's interest in engineering fosters persistence beliefs beyond belongingness and engineering identity, *International Journal of STEM Education*, 8, 33.
- Wang, J. and Hazari, Z. (2018). Promoting high school students' physics identity through explicit and implicit recognition. *Physical Review Physics Education Research*, 14, 020111.
- Wieselmann, J. R., Dare, E. A., Roehrig, G. H. and Ring-Whalen, E. A. (2022). Measurement instruments of STEM affective learning: A systematic review. In *International Encyclopedia of Education: Fourth Edition* (pp. 421-443). Elsevier. <https://doi.org/10.1016/B978-0-12-818630-5.13014-3>

PEER LEARNING IN LARGE SIZE FORMAT: IS IT EFFECTIVE IN BOLSTERING PROBLEM-SOLVING AND METACOGNITIVE SKILLS?

Matteo Bozzi¹, Roberto Mazzola, Maurizio Zani

Department of Physics, Politecnico di Milano, Milan, Italy

Abstract. *Active methods tend to bolster university students' learning. Previous studies highlighted that they generally improve learners' outcomes in concept inventories focused on conceptual understanding. In the context of a physics course offered to some Italian engineering freshmen in academic year 2021-2022, a quasi-experiment was implemented to investigate the effectiveness of a new pedagogical approach, based on the synergistic use of different educational practices and tools, and strengthened by the use of technology. Its effectiveness was measured investigating the freshmen' rate of success in their physics course final exam. This study highlights that this methodology is effective in increasing students' learning in the context of large size formats and a conservative estimate quantifies in 9.4% the size of this effect.*

Keywords: *Peer learning, Peer discussion, Student response system, Technology, STEM*

1. INTRODUCTION

The transition from secondary to tertiary education is a challenging step which requires students to take a true quality leap in terms of effort, determination and perseverance. In science, technology, engineering and mathematics (STEM) faculties introductory physics, mathematics and chemistry courses are a stumbling block to undergraduates, who face the risk of modest results in their final examination or outright failure (Bozzi et al., 2019, p. 2241; Medicine et al., 2016; Örnek et al., 2007; Redish et al., 2006).

¹ Corresponding author (matteo.bozzi@polimi.it)

Koch (2017) defines high-enrolment, foundation courses, characterised by low success rate, gateway courses. They often constitute barriers which learners need to overcome if they want to achieve graduation (Gasiewski et al., 2012; Medicine et al., 2016). A negative learning experience in these gateway courses, indeed, represents one of the principal reasons for dropping out (Seymour and Hunter, 2019).

Since the end of the 20th century it has been pointed out that the adoption of active methodologies in STEM courses bolsters students' learning (Hake, 1998). These findings have been confirmed by both a meta-analysis which examined 225 studies involving a wide range of active instructional practices in different disciplines (Freeman et al., 2014), and a number of recent studies (Balta et al., 2017; Bauer et al., 2022; Bozzi et al., 2021). In a recent book, Hattie (2023) provides a synthesis of more than 2,100 meta-analyses covering about 130,000 studies which engaged over 200 million students at all levels of education, from primary school to tertiary level. The authors categorise about 350 factors related to learning outcomes from very positive effects to very negative influences and the average effect size taken into account (Cohen's d is .40).

Hattie's investigations (Hattie, 2008, 2023; Hattie et al., 2015) show that some widespread active methodologies have a positive effect size, although lower or equal to the mean value. For instance, Cohen's d for problem-based learning and for cooperative learning is respectively .12 and .40. On the contrary, such factors as classroom discussion ($d = .82$), positive peer influences ($d = .53$) and peer tutoring ($d = .53$) show a higher effect size than the average value. With specific reference to the teaching of physics, student-centred activities for large enrolment undergraduate programs (Beichner, 2007; Foote et al., 2016; Foote et al., 2014), problem-based learning (Duch, 1996) and peer instruction (Crouch and Mazur, 2001; Dancy et al., 2016; Korpas et al., 2019; Rudolph et al., 2014; Wood et al., 2016; Zhang et al., 2017) are among the most noticeable pedagogical strategies designed and implemented in numerous universities.

Educational research allows arguing that students' conceptual understanding thrives when active methods are employed in STEM courses (Freeman et al., 2014; Hake, 1998; Wieman, 2014). Furthermore, these interactive-engagement methods foster learners' interest, self-efficacy (Dou et al., 2018), metacognitive regulation (De Backer et al., 2015), motivation

and engagement (Schunk and Zimmerman, 2012), which improve their academic outcomes (Delaney and Royal, 2017; Schunk and Zimmerman, 2012; Williams et al., 2019).

Freeman et al.'s (2014) seminal investigation highlights, however, that active learning improves learners' outcomes in concept inventories more than in final course examinations. Moreover, it has been empirically demonstrated that these educational approaches foster a reduction in the achievement gap which could afflict under-represented groups in STEM faculties (Seymour and Hunter, 2019; Theobald et al., 2020).

Nevertheless, the use of active methods in the gateway courses of STEM faculties remains sparse even despite an increased awareness of the need to implement them (Rasmussen et al., 2019; Seymour and Hunter, 2019; Stains et al., 2018). Instructors' dearth of time, their need to cover completely the syllabus content of their academic course, limited resources available, and a bland departmental as well as institutional support hamper their spread (Deslauriers et al., 2019; Silverthorn et al., 2006). A further barrier which may hamper the adoption of active learning is large size classes (Apkarian et al., 2021; Hornsby and Osman, 2014; Palfreyman et al., 2011; Scott, 1995; Shadle et al., 2017; Sturtevant and Wheeler, 2019). It is no wonder, thus, that studies on active methods undertaken in Europe frequently involve small or medium size formats (Leinonen et al., 2017; Schiltz et al., 2019; Wood et al., 2014).

Although the Italian context is even more complicated by the hegemonic role played by traditional academic lectures, the effects of innovative teaching methods like peer learning (PL) on students' conceptual understanding has already been investigated in the context of Italian STEM large size classes (Bozzi et al., 2018, 2021; Commission's Expert Group on Quality Investment in Education and Training et al., 2022; Directorate-General for Education, 2022; Raffaghelli et al., 2018).

In this study we wish to further investigate the issue of PL effectiveness in large size formats to examine whether this innovative educational methodology, already illustrated in previous studies (Bozzi et al., 2021, 2023), may contribute to improve not only the learners' conceptual understanding, but more generally their learning strategies and bolster different skills, like problem-solving and metacognitive skills.

2. METHOD

2.1 RESEARCH DESIGN

In the Italian context, undergraduates' achievements in introductory physics, mathematics and chemistry courses are generally assessed through instructor written exams which focus on problem-solving skills rather than conceptual understanding. Although this type of final examination is probably not the most favourable to evaluate the effectiveness of active instructional practices implemented in gateway courses, this is a widespread option adopted by many Italian universities, departments and instructors to measure learners' achievements. Thus, a wider adoption of innovative teaching methods requires taking up the challenge of demonstrating that they are effective in improving undergraduates' results in final exams based also on skills different from conceptual understanding. As a result, in the present study we identified the pass rate of a physics course as the dependent variable, comparing freshmen who attended or not attended the peer learning (PL) sessions and evaluating the amount of attendance.

In this regard, an innovative teaching methodology was implemented in the context of an introductory physics course attended by some Politecnico di Milano freshmen in academic year 2021-2022. Its main characteristics were the synergistic use of different educational practices and tools, like concept tests with double vote within the "Bring Your Own Device" (BYOD) framework, students response system (SRS), discussions among peers and explanations provided by an expert, along with the integration of active methods enhanced by the use of technology, traditional lectures, and exercise sessions in large size formats. A research question was formulated:

(RQ) *Does this integrated teaching methodology improve the students' final examination pass rate in a physics course?*

2.2 RESEARCH CONTEXT AND PARTICIPANTS

In the first term of the academic year, Politecnico di Milano freshmen enrolled on both Chemical Engineering and Materials and Nanotechnology Engineering are required to attend an introductory physics course called "Experimental Physics A+B" ("Fisica Sperimentale A+B" in Italian language). In this gateway course students address a number of mechanics and electromagnetism topics and are granted ten university educational credits if they pass the final exam.

In academic year 2021-2022 we conducted a quasi-experiment which involved 202 first year students enrolled on “Experimental Physics A+B”. In terms of gender, 62 of them identified themselves as females and the others 140 as males. According to a student survey, conducted by our university at the end of the term, the attendance at this academic course was high. Only 5.4% of the participants attended less than 25% of the activities, namely lectures and exercise sessions. Moreover, only 11.4% of the learners attended less than 50% of this course.

Moreover, all of them accepted to participate in the study. As the students were numerous, the size of the class may be regarded as large.

2.3 DESIGN OF THE TEACHING/LEARNING EXPERIENCE

The overall duration of the “Experimental Physics A+B” course was 100 hours, from September 2021 to December 2021. Traditional academic lectures within a theoretical framework covered 57 hours, 40 hours consisted in exercise sessions aimed at fostering learners’ problem-solving skills, and 3 hours were devoted to PL activities which were divided into 7 sessions focused on mechanics (4 sessions) and electromagnetism topics (3 sessions). Offered to the learners on average every ten days, the PL sessions were included within seven traditional lectures and the setting was the same large lecture classroom. They were conducted by the same instructor who gave classic academic lectures and their attendance was not mandatory in the same way as every lecture or exercise session. Moreover, the students did not know when the PL sessions would take place beforehand.

As illustrated in Bozzi et al. (2023), each PL session developed over four steps, the former three were repeated three times and the latter only once at the end of the session. In every PL session a questionnaire of three multiple choice items given in a sequence was administered to the learners through Socrative, an online portal whose effective use as a SRS has been extensively investigated (Balta et al., 2018; Gómez-Espina et al., 2019; Guarascio et al., 2017; Lim, 2017; Wood and Shirazi, 2020). Since the BYOD approach has been documented to be successful in terms of students’ engagement and outcomes (Afreen, 2014; Blasco-Arcas et al., 2013; Caldwell, 2007; Hattie, 2008, 2023; Mayer et al., 2009; McLean, 2016), we asked participants to answer every question by using their own smartphones, tablets or laptops. Each item provided four possible

alternatives, one correct and three incorrect, and the learners had 1.50 min to give their answer.

Before describing the design of the PL session in detail, it is worth noting that these items were constructed ad hoc by the researchers to highlight some widespread misunderstandings or conceptual errors which freshmen often make about some topics discussed in our academic course. They are frequently related to how these issues were addressed in their physics lectures at high school. These closed-ended questions were generally based on the researchers' professional background and experience in teaching physics and their awareness of how these topics were proposed in some popular Italian high school physics textbooks. Moreover, they were not selected from or inspired by published concept inventories or research-based textbooks (see Appendix 1 for one of these items).

As aforementioned, each PL session developed over four steps repeated three times. The first step began when the instructor showed to the participants the first item and gave them 1.50 min to answer it. After this first question was completed, and without any feedback on the answer, a peer discussion phase was developed in small groups (3–4 people) for few minutes (3–5 min). This was the second step. For ease of arrangement, participants were instructed to work in groups with students sitting next to them. At the end of this discussion, they answered the same question individually again. The time allotted to this third step was again 1.50 min. This way, the students debated the issue with their classmates in the small group and soon afterwards could ponder personally on it and draw their own conclusions.

This sequence was repeated in succession twice more so that the freshmen could address also the second and the third item of the questionnaire. Finally, there was one last step, i.e. the explanation from an expert, which lasted between 3 and 5 minutes. The instructor gave the correct answer to every item in the concept test and explained why the other alternatives were incorrect. Moreover, the teacher showed the percentage of participants who had chosen the different options in each question. These different percentages were shown only at the end and not after the first and the third step, not to influence the peer discussion phase and the possible contribute of each participant to it.

In conclusion, we would like to emphasise again that PL should not be intended as peer discussion alone, considering that our overall innovative active method combined a plurality of educational tools and instructional practices, among others peer discussion.

2.4 MODELLING OF THE PROBLEM AND INDEPENDENT VARIABLES

To assess the effectiveness of our innovative teaching method, we considered the freshmen' rate of success in their physics course final exam during the first examination session (dependent variable), which took place between January 2022 and February 2022. The final examination was an instructor written exam. It lasted two hours and consisted in four problems, two focused on mechanics topics and the other two on electromagnetism subjects (see Appendix 2 for an example of these problems).

This pass rate was conceivably influenced by multiple factors, like students' motivation, engagement, interest, initial knowledge in physics and, we hope, our integrated educational approach. To evaluate their role and quantify their weight on the pass rate, two independent variables, were defined, namely "Attendance at the PL sessions" (V1) and "Exposure dose to the PL sessions" (V2).

V1 was a dichotomous variable which aggregated the learners on the basis of their attendance or non-attendance at the PL sessions. These data were collected during the same PL sessions. The criterium adopted was 50%, hence freshmen who had attended more than 50% of the PL sessions, from 4 to 7 sessions, were considered to have been exposed to PL and constituted the first group (G1). Those who were classified as not exposed to PL (between 0 and 3 sessions attended) were assigned to the second group (G2). It should be emphasised that factors like students' motivation, engagement and interest could be substantially assumed constant into G1 and G2, while they showed a discontinuity in the transition from G2 to G1. It seems reasonable, indeed, to assume that learners who attended nought PL sessions rather than one or, symmetrically, six sessions rather than seven did not have a truly different motivation, engagement and interest. Therefore, V1 was sensitive to these factors but, at the same time, was not extremely sensitive to exposure to the PL methodology due to the fact that being a dichotomous variable, it condensed the eight possible exposure doses, which were in themselves few, into just two categories (exposed vs non-exposed).

V2 was an ordinal variable which aggregated the freshmen on the basis of the number of PL sessions which they had attended and it ranged from 0 to 7. It was blatantly sensitive to exposure to PL methodology, but it modelled the other factors worse than V1. Finally, on account of the high course attendance rates

(see Section 2.2), V1 and V2 may not be extremely sensitive to the students general attendance at traditional lectures and exercise sessions.

3. RESULTS

The data collected were analysed through the statistical opensource software R (version 4.2.2) in the integrated development environment RStudio (<https://rstudio.com/> accessed on 8 November 2023). The statistical significance was beforehand established at level $\alpha = .05$.

3.1 IMPACT OF THE STUDENTS' INITIAL KNOWLEDGE IN PHYSICS

A first investigation stage consisted in determining the initial level of knowledge in physics of all the students enrolled on “Experimental Physics A+B” and involved in this study at the beginning of the course. For this purpose, an initial questionnaire on some physics topics was administered to the participants during one of the first lectures of the academic course (see Appendix 3 for an example of these problems).

This questionnaire consisted of 18 multiple choice items focused on some mechanics (6), thermodynamics (6) and electromagnetism (6) subjects. Not taken from published concept inventories or research-based textbooks, it was substantially a concept test without any numeric problems. Its items, which were different from the questions offered during the subsequent PL sessions, allowed to test or check some possible freshmen’s misunderstandings, conceptual errors or misconceptions. The learners gained one point for each correct answer and nought points in all other cases, therefore the maximum possible score achieved by the participants was 18 and the minimum 0. The questionnaire was administered online and freshmen answered it using their own electronic devices (BYOD approach).

Through this instrument we could compare the initial level of knowledge in physics of G1 and G2. Table 1 shows some descriptive score statistics about the findings in this initial questionnaire regarding the two cohorts.

Table 1: Descriptive score statistics for G1 and G2 in the initial questionnaire.

Cohort	Number of students	Average percentage of correct answers	Mean score (max. 18)	Standard error
G1	63	30.78	5.54	.24
G2	41	30.78	5.54	.27

To check the normality of the population distributions a Shapiro-Francia test was implemented. Given that the populations were not normally distributed, we concluded that the most appropriate inferential tests to compare the different mean scores in the initial questionnaire were a robust ANOVA as well as the Mann-Whitney U-test. We conducted both these tests, which highlighted that the difference between the mean score of the two cohorts was not statistically different, i.e. the findings of the two groups were equivalent. Table 2 summarises the outcomes of these statistical tests.

Table 2: Robust ANOVA and Mann-Whitney U-test results for G1 and G2.

Test	Parameter	df	p-value	Effect size
Robust ANOVA	F = 0.1089	1; 53.36	.74	.0924
Mann-Whitney U	Z = 0.18563		.85	(negligible)

3.2 IMPACT OF THE STUDENTS' MOTIVATION, ENGAGEMENT AND INTEREST

To evaluate the possible impact of the learners' participation and engagement, we compared the pass rate of G1 and G2 students in their physics course final exam. To turn it round, we analysed if there was an association between the dependent variable "Rate of success in physics course final exam" and the independent variable V1. Although motivation and interest should be investigated through specific questionnaires and could not be completely identified with participation, V1 was related to them. As a consequence, this analysis may provide a first indication of their possible impact.

Descriptive statistics appeared to highlight possible major differences between the pass rate of the two cohorts of freshmen. In G1, the learners who passed the exam were many more than anticipated if there were no association

between the pass rate and the attendance at the PL sessions, and the learners who failed were largely fewer than expected. The opposite happened in G2. Tables 3 and 4 are the contingency tables and highlight respectively the observed and the expected distribution, i.e. the number of learners who passed or failed the examination and the number of learners who may be expected to pass or fail it if there were no association between the two variables.

Table 3: Observed distribution of students who passed or failed the final exam for cohorts G1 and G2.

Cohort	Pass number	Failure number
G1	64	33
G2	24	81

Table 4: Expected distribution of students who may have passed or failed the final exam for cohorts G1 and G2.

Cohort	Pass number	Failure number
G1	42.26	54.74
G2	45.74	59.26

It should be pointed out that all the students enrolled on “Experimental Physics A+B”, ($n = 202$ as shown in table 3) were considered in our analysis. The freshmen who took the physics course final exam in the first session were 161 (79.7%), aligned with the previous academic years. In fact, the average exam attendance rate had been equal to 79.0% between the academic year 2014-2015 and 2020-2021. As the present study focuses on evaluating if our methodology might determine an increase in the number of students who passed the exam within the beginning of the second academic term, the fact that some students failed the exam or did not take it was considered equivalent.

To investigate the possible association between the outcome of the exam and the students’ attendance of the PL activities, we implemented a Pearson’s chi-square test. It highlighted a statistically highly significant association between the two variables ($\chi^2 = 38.13533$, $df = 1$, $p\text{-value} = 6.60 \cdot 10^{-10} \ll .05$). Moreover, the intensity of the effect was evaluated through both odds ratio (OR) and Cramer’s V , which were respectively equal to 6.545 (large or strong (Rosenthal, 1996)) and .434 (relatively strong (Rea and Parker, 2014)). Thus, we

might argue that not only was there a statistically highly significant association between the outcome of the final exam and V1, but the strength of this association was extremely significant.

3.3 IMPACT OF THE INNOVATIVE INSTRUCTIONAL PRACTICE

To assess the possible impact of our innovative educational methodology, we compared the pass rate of students who attended a different number of PL sessions in their physics course final exam. To turn it round, we analysed if there was an association between the dependent variable “Rate of success in physics course final exam” and the independent variable V2. Moreover, the pass proportion was calculated with reference to all the possible exposure doses to PL (P_i). P_i was defined as the ratio between the number of students who both attended i PL sessions and passed the final exam and the total number of learners who attended i PL sessions. The P_i expected in case of no association between the dependent variable and V2 was .436. Table 5 shows the descriptive score statistics regarding students who were differently exposed to PL and their results in the final exam.

Table 5: Descriptive score statistics regarding students differently exposed to the PL and their results in the final exam.

	V2								
	0	1	2	3	4	5	6	7	Total number
Pass number	9	6	5	4	10	17	12	25	88
Failure number	39	12	18	12	9	9	10	5	114
Total number	48	18	23	16	19	26	22	30	202
P_i	.188	.333	.217	.250	.526	.654	.545	.833	.436

Descriptive statistics appeared to highlight possible major differences between the pass rate of the diverse cohorts of freshmen. It should be emphasised that the lowest P_i regarded students who had never attended PL sessions, while the highest P_i concerned learners who had participated in all the PL activities. Furthermore, every cohort previously assigned to G1 (from 4 to 7 PL sessions

attended) showed a P_i superior to .436, which was the value expected in event of no association between the two variables studied. Conversely, each group which had been classified in G2 had a P_i inferior to .436.

Employing a Pearson's chi-square test ($\chi^2 = 45.53499$, $df = 7$, $p\text{-value} = 1.08 \cdot 10^{-7} < .05$) we could argue that there was a statistically highly significant difference between at least two of these pass proportions. Moreover, Cramer's V effect size was extremely significant ($V = .475$, relatively strong (Rea and Parker, 2014)).

To analyse if the dependent variable "Rate of success in physics course final exam" depended on exposure to our instructional method, which was only one of the factors that might influence V2, we investigated if there was a linear trend among the pass proportions. In our model all the other possible factors, for instance students' motivation, engagement and interest, levelled out among the cohorts of students corresponding to G1 and G2, thus a linear trend might suggest an association between the exposure to PL and the rate of success. From this perspective and in analogy with a variety of carcinogenicity studies, toxicological risk assessment and clinical trials based on dose-response relationships, we carried out a dose-response experiment as we were interested in investigating a possible dose-related trend, hypothesising an increasing trend in pass proportions (response) with increasing attendance to the PL sessions (dose).

As the Cochran-Armitage (Armitage, 1955; Cochran, 1954) is the most frequently adopted trend test for binomial proportions (Chen et al., 1997; Neuhauser and Hothorn, 1999), it was implemented for this purpose and three different cases were analysed. Firstly, we considered the eight pass proportions related to every possible number of PL sessions attended by the participants (from nought to seven). Secondly, we carried out the test considering only the pass proportions corresponding to nought-three PL sessions attended (G2) and finally we investigated the possible linear trend intra G1. Table 6 summarises the Cochran-Armitage test findings.

Table 6: Cochran-Armitage test findings.

Case	Parameter	dim	p-value	Linear trend
From 0 PL to 7 PL	$Z = 6.2588$	8	$3.88 \cdot 10^{-10}^{***}$	Yes
From 0 PL to 3 PL (G2)	$Z = 0.49276$	4	.6222	No
From 4 PL to 7 PL (G1)	$Z = 2.0095$	4	.04448 *	Yes

* $p < .05$, *** $p < .001$

The Cochran-Armitage test highlighted an overall linear trend among the pass proportions, but, at the same time, it pointed out that this trend concerned only G1.

4. DISCUSSION

To evaluate if the integrated teaching methodology which we developed was effective in improving the participants' final examination pass rate of the academic course "Experimental Physics A+B", a quasi-experimental research study was conducted. We preliminarily investigated possible differences among the freshmen involved in this study in terms of their initial knowledge of physics. Although the students had not been randomly assigned to G1 and G2, it was demonstrated that their knowledge of physics was equivalent at the beginning of the academic course. It should be emphasised that their mean score in the initial questionnaire (5.54, corresponding to 30.78% of correct answers) appeared to be much lower than the maximum score potentially achievable, i.e. 18. It implies that all the learners appeared to be slightly weak in physics.

Interestingly, a distinctive peculiarity of this study was that not only both G1 and G2 had the same instructor, but they attended the same lectures and exercise sessions at the same time. As a consequence, notwithstanding that the teacher is recognised to play a paramount role in the students' learning process (Hattie, 2012), the possible positive or negative impact of the lecturer was substantially equivalent for every group and each learner.

The Pearson's chi-square test pointed out a statistically highly significant association between the rate of success in the physics course final exam and the independent variable "Attendance of the PL sessions". Furthermore, the size of the effect might be classified as large or relatively strong with reference to odds ratio and Cramer's V respectively. However, this independent variable was influenced by many factors, like students' motivation, engagement, interest, and involvement in our innovative instructional practice ("Exposure to the PL methodology"). As a consequence, it is not possible to argue that the learners' increased success rate was the result of our teaching method. Rather, it depended on all these factors and their synergistic action. Furthermore, if it was not possible to quantify the contribute of each of them to an increase in the freshmen's success rate in their final exam, likewise it was not possible to evaluate how much every factor contributed to the odds ratio and Cramer's V overall values. However, since the independent variable V1 was not very

sensitive to the factor “Exposure to the PL methodology”, we assumed very conservatively that the intensity of the effect was entirely related to all the other factors and considered the calculated odds ratio and Cramer’s V as a baseline.

In a further step, a statistically highly significant correlation was found between the rate of success in the physics course final exam and the second independent variable V2 “Exposure dose to the PL sessions”. Like V1, V2 did not purely depend on the students’ exposure to the PL methodology, but it was also influenced by all the other aforementioned factors. However, it ought to be remembered that in our model these further factors, like engagement, interest and motivation, should be considered substantially constant within G1 and G2 and they may show a discontinuity, a sudden increase in the transition from G2 to G1 owing to the binary nature of V1.

Therefore, with regard to the influence of these factors if the “Exposure to the methodology” was ineffective or its effects were negligible, the model would forecast that they equally affect the rate of success in the final exam of learners who respectively had attended from nought to three PL sessions (G2) and from four to seven PL sessions (G1). Thus, the P_i values should show a discontinuity in the transition from three and four PL sessions attended, but they ought to be flat simultaneously within G2 and G1. However, the P_i values were flat within G2, but not within G1, where a linear upward trend emerged. As a result, we might conclude that these further factors could not determine the observed upward trend in the P_i , while this trend is consistent with our modelling of the factor “Exposure to the methodology”. Thus, our instructional practice appeared to be effective in fostering participants’ learning.

However, it is reasonable to assume that there is an efficacy threshold of this instructional practice, i.e. our innovative educational method may be generally effective only if a student attends a minimum number of PL sessions. The findings of the Cochrane-Armitage tests involving G1 and G2 confirmed this hypothesis and showed that this threshold corresponded to four PL sessions.

If instructors would like to try out our PL methodology in their academic courses, they should be conscious that it needs regularity in its administration to the students. With regard to time, however, the efficacy threshold found corresponded to only about 1.5%-2% of the overall duration of our physics course, a modest investment in time. Moreover, PL sessions were included within traditional lectures which were not based on physics education research. The setting was the same large lecture classroom and the PL sessions were conducted by the same lecturer who gave classic academic lectures. Furthermore, only three items per session were used and an instructor could

decide to employ concept questions already available in research-based textbooks.

A conservative quantitative assessment of the PL effectiveness in increasing the students' success rate in their final exam could be done by comparing the intensity of the effect when V1 and V2 were considered. Cramer's V was respectively equal to .434 and .475, corresponding to a percentage increase of 9.4%. On balance, our innovative teaching methodology might improve the freshmen's success rate by about 9.4% and this appears to be a cautious estimate.

This result is not immediately comparable with the findings achieved in Hattie's studies (Hattie, 2008, 2023) due to the fact that this percentage cannot be expressed in terms of Cohen's d. However, according to Freeman et al. (2014) the students who attend academic courses consisting in traditional lectures are 1.5 times more likely to fail their final exams than students attending courses with active learning, as the risk ratio for failure (RR) is 1.5. In our study, if the freshmen who had attended 0 PL sessions are assumed as a baseline on account of their being the only learners who were not at all exposed to PL methodology, the RR with reference to respectively all the other students (from 1 to 7 PL sessions attended), the other freshmen of G2 (between 1 and 3 PL sessions attended) and the G1 students (from 4 to 7 PL sessions attended) is respectively $1.67 > 1.5$ (+11.3%), $1.10 < 1.5$ (-26.7%) and $2.74 \gg 1.5$ (+82.7%). A prudent conclusion is that learners who attend only traditional lectures and exercise sessions are 1.67 times more likely to fail their final exams than freshmen exposed to our PL methodology and this value is 11.3% higher than the average RR that they would have if other active methods are considered.

It is, however, worth highlighting that although there seems to be a correlation between the innovative teaching method that we propose in the present paper and our participants' rate of success, it is not possible to establish a causal effect between them, considering that other factors, like the participants' general course attendance, motivation, engagement, and interest may have contributed to their success rate.

5. CONCLUSIONS

In the context of a physics course offered to some Politecnico di Milano engineering freshmen across the first term of academic year 2021-2022, a quasi-experiment was implemented to investigate the effectiveness of a new pedagogical practice, based on the integration of PL activities strengthened by the use of technology, traditional lectures and exercise sessions. Importantly, it has been highlighted that this innovative educational methodology is effective in increasing the students' learning in the context of large size formats and a

conservative estimate quantifies the size of this effect as 9.4%. Furthermore, it appears on average more effective than other active methods in improving the students' rate of success in their final exam when this final exam consists in problems to solve.

As the physics course final exam focused on problems which required a clear conceptual understanding of the topics involved, we argue that our PL methodology bolsters problem-solving skills and conceptual understanding. It should be emphasised that active learning generally tends to improve learners' outcomes in concept inventories focused on conceptual understanding more than to boost the students' findings in instructor-written course examinations (Freeman et al., 2014). It is worth pointing out that we did not measure specifically the improvement of freshmen conceptual understanding through an ad-hoc concept test in this study. However, we had already highlighted the effectiveness of our PL methodology in increasing student conceptual understanding in physics in a previous research (Bozzi et al., 2021; Directorate-General for Education, 2022, p. 48).

Finally, it has been confirmed the hypothesis that there is an efficacy threshold of the PL educational method. Student learning begins to improve already after four PL sessions, corresponding to about 1.5 h - 2 h out of 100 h which is the overall duration of our physics course. It would seem a low threshold and this might encourage instructors to experiment with this innovative educational methodology.

Lastly, it is worth emphasising that traditional lectures and exercise sessions offered during "Experimental Physics A+B" were not physics education research-based and the items employed during PL sessions were not inspired by published concept inventories or research-based textbooks. As a consequence, it is unlikely that the effectiveness of our innovative teaching methods is related to physics education research findings, like the activation of some productive students' resources during PL sessions. On the contrary, the positive results achieved may be due to the method which fosters the learners' metacognitive regulation and their self-regulated learning. The research on this issue has shown that metacognitive regulation improves the students' performance (Meijer et al., 2006; Schraw and Moshman, 1995; Zimmerman and Schunk, 2011). Moreover, peer interaction is extremely effective in triggering regulatory skills, like planning, monitoring and evaluation (De Backer et al., 2015; Schraw and Moshman, 1995). Considering this perspective, it could be argued that our innovative teaching methodology is also effective in bolstering metacognitive skills.

REFERENCES

- Afreen, R. (2014). Bring your own device (BYOD) in higher education: Opportunities and challenges. *International Journal of Emerging Trends & Technology in Computer Science*, 3(1), 4.
- Apkarian, N., Henderson, C., Stains, M., Raker, J., Johnson, E. and ,M. (2021). What really impacts the use of active learning in undergraduate STEM education? Results from a national survey of chemistry, mathematics, and physics instructors. *PLoS ONE*, 16(2 February). Scopus. <https://doi.org/10.1371/journal.pone.0247544>
- Armitage, P. (1955). Tests for linear trends in proportions and frequencies. *Biometrics*, 11(3), 375–386. <https://doi.org/10.2307/3001775>
- Balta, N., Michinov, N., Balyimez, S. Ayaandz, M. F. (2017). A meta-analysis of the effect of peer instruction on learning gain: Identification of informational and cultural moderators. *International Journal of Educational Research*, 86, 66–77. <https://doi.org/10.1016/j.ijer.2017.08.009>
- Balta, N., Perera-Rodríguez, V.-H. and Hervás-Gómez, C. (2018). Using Socrative as an online homework platform to increase students' exam scores. *Education and Information Technologies*, 23(2), 837–850. <https://doi.org/10.1007/s10639-017-9638-6>
- Bauer, T., Biehler, R. and Lankeit, E. (2022). Concept tests in undergraduate real analysis: Comparing peer discussion and instructional explanation settings. *International Journal of Research in Undergraduate Mathematics Education*. <https://doi.org/10.1007/s40753-022-00167-y>
- Beichner, R. J. (2007). The student-centered activities for large enrollment undergraduate programs (SCALE-UP) project. In *Research-Based Reform of University Physics* (Vol. 1, pp. 1–42). E. F. Redish and P. J. Cooney (eds). <https://www.per-central.org/items/detail.cfm?ID=4517&Attached=1>
- Blasco-Arcas, L., Buil, I., Hernández-Ortega, B. and Sese, F. J. (2013). Using clickers in class. The role of interactivity, active collaborative learning and engagement in learning performance. *Computers & Education*, 62, 102–110. <https://doi.org/10.1016/j.compedu.2012.10.019>
- Bozzi, M., Balossi, B., Sieno, L. D., Ganzer, L., Gondoni, P., Genco, I., Minnai, C., Pini, A., Rezoagli, F., Zanoletti, M. and Zani, M. (2019). Securing

- freshmen's learning through a physics refresher course: A breakthrough experience at Politecnico di Milano. *ICERI2019 Proceedings*, 2237–2243. <https://library.iated.org/view/BOZZI2019SEC>
- Bozzi, M., Mazzola, R. and Zani, M. (2023). Peer learning in higher education: An example of practices. *Il Nuovo Cimento C*, 46(6), 1–11. <https://doi.org/10.1393/ncc/i2023-23206-7>
- Bozzi, M., Raffaghelli, J. E. and Zani, M. (2021). Peer learning as a key component of an integrated teaching method: Overcoming the complexities of physics teaching in large size classes. *Education Sciences*, 11(2), Article 2. <https://doi.org/10.3390/educsci11020067>
- Bozzi, M., Raffaghelli, J. and Zani, M. (2018). Peer learning for large size physics lectures in higher education: Yes, we can. *ICERI2018 Proceedings*, 8739–. <https://library.iated.org/view/BOZZI2018PEE>
- Caldwell, J. E. (2007). Clickers in the large classroom: Current research and best-practice tips. *CBE—Life Sciences Education*, 6(1), 9–20. <https://doi.org/10.1187/cbe.06-12-0205>
- Chen, J. J., Kodell, R. L. and Pearce, B. A. (1997). Significance levels of randomization trend tests in the event of rare occurrences. *Biometrical Journal*, 39(3), 327–337. <https://doi.org/10.1002/bimj.4710390307>
- Cochran, W. G. (1954). Some methods for strengthening the common χ^2 tests. *Biometrics*, 10(4), 417–451. <https://doi.org/10.2307/3001616>
- Commission's Expert Group on Quality Investment in Education and Training, Directorate-General for Education, Y., Fack, G., Agasisti, T., Bonal, X., De Witte, K., Dohmen, D., Haase, S., Hylen, J., McCoy, S., Neycheva, M., Pantea, M. C., Pastore, F., Pausits, A., Poder, K., Puukka, J. and Velissaratou, J. (2022). *Interim Report of the Commission Expert Group on Quality Investment in Education and Training*. Publications Office of the European Union. <https://data.europa.eu/doi/10.2766/37858>
- Crouch, C. H. and Mazur, E. (2001). Peer instruction: Ten years of experience and results. *American Journal of Physics*, 69(9), 970–977. Scopus.
- Dancy, M., Henderson, C. and Turpen, C. (2016). How faculty learn about and implement research-based instructional strategies: The case of peer instruction. *Physical Review Physics Education Research*, 12(1), 010110. <https://doi.org/10.1103/PhysRevPhysEducRes.12.010110>
- De Backer, L., Van Keer, H. and Valcke, M. (2015). Promoting university students' metacognitive regulation through peer learning: The potential

- of reciprocal peer tutoring. *Higher Education*, 70(3), pp. 469–486.
<https://doi.org/10.1007/s10734-014-9849-3>
- Delaney, M. L. and Royal, M. A. (2017). Breaking engagement apart: The role of intrinsic and extrinsic motivation in engagement strategies. *Industrial and Organizational Psychology*, 10(1), 127–140.
<https://doi.org/10.1017/iop.2017.2>
- Deslauriers, L., McCarty, L. S., Miller, K., Callaghan, K., & Kestin, G. (2019). Measuring actual learning versus feeling of learning in response to being actively engaged in the classroom. *Proceedings of the National Academy of Sciences of the United States of America*, 116(39), pp. 19251–19257. Scopus. <https://doi.org/10.1073/pnas.1821936116>
- Directorate-General for Education (2022). *Investing in Our Future: Quality Investment in Education and Training*. Publications Office of the European Union. <https://data.europa.eu/doi/10.2766/45896>
- Dou, R., Brewe, E., Potvin, G., Zwolak, J. P. and Hazari, Z. (2018). Understanding the development of interest and self-efficacy in active-learning undergraduate physics courses. *International Journal of Science Education*, 40(13), 1587–1605.
<https://doi.org/10.1080/09500693.2018.1488088>
- Duch, B. J. (1996). Problem-based learning in physics: The power of students teaching students. *Journal of College Science Teaching*, 15(5), pp. 326–329.
- Foote, K., Knaub, A., Henderson, C., Dancy, M. and Beichner, R. J. (2016). Enabling and challenging factors in institutional reform: The case of SCALE-UP. *Physical Review Physics Education Research*, 12(1), 010103. <https://doi.org/10.1103/PhysRevPhysEducRes.12.010103>
- Foote, K. T., Neumeyer, X., Henderson, C., Dancy, M. H. and Beichner, R. J. (2014). Diffusion of research-based instructional strategies: The case of SCALE-UP. *International Journal of STEM Education*, 1(1), 10. <https://doi.org/10.1186/s40594-014-0010-8>
- Freeman, S., Eddy, S. L., McDonough, M., Smith, M. K., Okoroafor, N., Jordt, H. and Wenderoth, M. P. (2014). Active learning increases student performance in science, engineering, and mathematics. *Proceedings of the National Academy of Sciences of the United States of America*, 111(23), 8410–8415. Scopus. <https://doi.org/10.1073/pnas.1319030111>

- Gasiewski, J. A., Eagan, M. K., Garcia, G. A., Hurtado, S. and Chang, M. J. (2012). From gatekeeping to engagement: A multicontextual, mixed method study of student academic engagement in introductory STEM courses. *Research in Higher Education*, 53(2), 229–261. <https://doi.org/10.1007/s11162-011-9247-y>
- Gómez-Espina, R., Rodríguez-Oroz, D., Chávez, M., Saavedra, C. and Bravo, M. J. (2019). Assessment of the Socrative platform as an interactive and didactic tool in the performance improvement of STEM university students. *Higher Learning Research Communications*, 9(2).
- Guarascio, A. J., Nemecek, B. D. and Zimmerman, D. E. (2017). Evaluation of students' perceptions of the Socrative application versus a traditional student response system and its impact on classroom engagement. *Currents in Pharmacy Teaching and Learning*, 9(5), 808–812. <https://doi.org/10.1016/j.cptl.2017.05.011>
- Hake, R. R. (1998). Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses. *American Journal of Physics*, 66(1), 64–74. <https://doi.org/10.1119/1.18809>
- Hattie, J. (2008). *Visible Learning: A Synthesis of Over 800 Meta-Analyses Relating to Achievement* (1st Edition). Routledge. <https://www.routledge.com/Visible-Learning-A-Synthesis-of-Over-800-Meta-Analyses-Relating-to-Achievement/Hattie/p/book/9780415476188>
- Hattie, J. (2012). *Visible Learning For Teachers: Maximizing Impact on Achievement*. Routledge. <https://visible-learning.org/2013/01/visible-learning-for-teachers-book-review/>
- Hattie, J. (2023). *Visible Learning: The Sequel: A Synthesis of Over 2,100 Meta-Analyses Relating to Achievement*. Routledge. <https://www.routledge.com/Visible-Learning-The-Sequel-A-Synthesis-of-Over-2100-Meta-Analyses-Relating/Hattie/p/book/9781032462035>
- Hattie, J., Masters, D. and Birch, K. (2015). *Visible Learning into Action: International Case Studies of Impact* (1st Edition). Routledge. <https://www.routledge.com/Visible-Learning-into-Action-International-Case-Studies-of-Impact/Hattie-Masters-Birch/p/book/9781138642294>

- Hornsby, D. J. and Osman, R. (2014). Massification in higher education: Large classes and student learning. *Higher Education*, 67(6), pp. 711–719. Scopus. <https://doi.org/10.1007/s10734-014-9733-1>
- Koch, A. K. (2017). It's about the gateway courses: Defining and contextualizing the issue. *New Directions for Higher Education*, 2017(180), 11–17. <https://doi.org/10.1002/he.20257>
- Korpas, G. S., Kahrs, M. S. and Andersen, T. H. (2019). Peer learning and concept tests in physics. *Varietas Delectat... Complexity Is the New Normality*, 651–657. <https://researchportal.tuni.fi/en/publications/proceedings-of-the-sefi-47th-annual-conference-2019-varietas-dele>
- Leinonen, R., Asikainen, M. A. and Hirvonen, P. E. (2017). Peer discussions in lecture-based tutorials in introductory physics. *Physical Review Physics Education Research*, 13(1). Scopus. <https://doi.org/10.1103/PhysRevPhysEducRes.13.010114>
- Lim, W. N. (2017). Improving student engagement in higher education through mobile-based interactive teaching model using Socrative. *2017 IEEE Global Engineering Education Conference (EDUCON)*, 404–412. <https://doi.org/10.1109/EDUCON.2017.7942879>
- Mayer, R. E., Stull, A., DeLeeuw, K., Almeroth, K., Bimber, B., Chun, D., Bulger, M., Campbell, J., Knight, A. and Zhang, H. (2009). Clickers in college classrooms: Fostering learning with questioning methods in large lecture classes. *Contemporary Educational Psychology*, 34(1), pp. 51–57. <https://doi.org/10.1016/j.cedpsych.2008.04.002>
- McLean, K. J. (2016). The implementation of bring your own device (BYOD) in primary [elementary] schools. *Frontiers in Psychology*, 7. <https://www.frontiersin.org/articles/10.3389/fpsyg.2016.01739>
- Medicine, N. A. of S., Engineering, and, Engineering, N. A. of, Affairs, P. and G., Workforce, B. on H. E. and, Education, D. of B. and S. S. and, Education, B. on S., & Degrees, C. on B. and O. in C. 2-Y. and 4-Y. S. (2016). *Barriers and Opportunities for 2-Year and 4-Year STEM Degrees: Systemic Change to Support Students' Diverse Pathways*. National Academies Press.
- Meijer, J., Veenman, M. V. J. and van Hout-Wolters, B. H. A. M. (2006). Metacognitive activities in text-studying and problem-solving:

- Development of a taxonomy. *Educational Research and Evaluation*, 12(3), pp. 209–237. <https://doi.org/10.1080/13803610500479991>
- Neuhäuser, M. and Hothorn, L. A. (1999). An exact Cochran–Armitage test for trend when dose–response shapes are a priori unknown. *Computational Statistics & Data Analysis*, 30(4), 403–412. [https://doi.org/10.1016/S0167-9473\(98\)00091-7](https://doi.org/10.1016/S0167-9473(98)00091-7)
- Örnek, F., Robinson, W. R. and Haugan, M. R. (2007). What makes physics difficult? *Science Education International*, 18(3), pp. 165–172.
- Palfreyman, D., Tapper, T. and Thomas, S. (2011). Series editors’ introduction: International studies in higher education. *Accountability in Higher Education: Global Perspectives on Trust and Power*, 1–270. Scopus. <https://doi.org/10.4324/9780203846162>
- Raffaghelli, J., Ghislandi, P., Sancassani, S., Canal, L., Micciolo, R., Balossi, B., Bozzi, M., Di Sieno, L., Genco, I., Gondoni, P., Pini, A. and Zani, M. (2018). Integrating MOOCs in physics preliminary undergraduate education: Beyond large size lectures. *Educational Media International*, 55(4), pp.301–316. <https://doi.org/10.1080/09523987.2018.1547544>
- Rasmussen, C., Apkarian, N., Hagman, J. E., Johnson, E., Larsen, S. and Bressoud, D. (2019). Brief report: Characteristics of precalculus through calculus 2 programs: Insights from a national census survey. *Journal for Research in Mathematics Education*, 50(1), 98–111. <https://doi.org/10.5951/jresmetheduc.50.1.0098>
- Rea, L. M. and Parker, R. A. (2014). *Designing and Conducting Survey Research: A Comprehensive Guide* (4th Edition). Jossey-Boss. <https://www.wiley.com/en-us/Designing+and+Conducting+Survey+Research%3A+A+Comprehensive+Guide%2C+4th+Edition-p-9781118767030>
- Redish, E. F., Scherr, R. E. and Tuminaro, J. (2006). Reverse-engineering the solution of a ‘simple’ physics problem: Why learning physics is harder than it looks. *Physics Teacher*, 44(5), 293–300. Scopus. <https://doi.org/10.1119/1.2195401>
- Rosenthal, J. A. (1996). Qualitative descriptors of strength of association and effect size. *Journal of Social Service Research*, 21(4), 37–59. https://doi.org/10.1300/J079v21n04_02
- Rudolph, A. L., Lamine, B., Joyce, M., Vignolles, H. and Consiglio, D. (2014). Introduction of interactive learning into French university physics

- classrooms. *Physical Review Special Topics - Physics Education Research*, 10(1), 010103. <https://doi.org/10.1103/PhysRevSTPER.10.010103>
- Schiltz, G., Feldman, G. and Vaterlaus, A. (2019). *Active-Learning Settings and Physics Lectures: A Performance Analysis*. 1286(1). Scopus. <https://doi.org/10.1088/1742-6596/1286/1/012019>
- Schraw, G. and Moshman, D. (1995). Metacognitive theories. *Educational Psychology Review*, 7(4), 351–371. <https://doi.org/10.1007/BF02212307>
- Schunk, D. H. and Zimmerman, B. J. (Eds.). (2012). *Motivation and Self-Regulated Learning: Theory, Research, and Applications*. Routledge.
- Scott, P. (1995). *The Meanings of Mass Higher Education*. Open University Press.
- Seymour, E. and Hunter, A.-B. (Eds.). (2019). *Talking about Leaving Revisited: Persistence, Relocation, and Loss in Undergraduate STEM Education*. Springer International Publishing. <https://doi.org/10.1007/978-3-030-25304-2>
- Shadle, S. E., Marker, A. and Earl, B. (2017). Faculty drivers and barriers: Laying the groundwork for undergraduate STEM education reform in academic departments. *International Journal of STEM Education*, 4(1), Article 1. <https://doi.org/10.1186/s40594-017-0062-7>
- Silverthorn, D. U., Thorn, P. M. and Svinicki, M. D. (2006). It's difficult to change the way we teach: Lessons from the integrative themes in physiology curriculum module project. *American Journal of Physiology - Advances in Physiology Education*, 30(4), pp. 204–214. Scopus. <https://doi.org/10.1152/advan.00064.2006>
- Stains, M., Harshman, J., Barker, M. K., Chasteen, S. V., Cole, R., DeChenne-Peters, S. E., Eagan, M. K., Jr., Esson, J. M., Knight, J. K., Laski, F. A., Levis-Fitzgerald, M., Lee, C. J., Lo, S. M., McDonnell, L. M., McKay, T. A., Michelotti, N., Musgrove, A., Palmer, M. S., Plank, K. M., ... Young, A. M. (2018). Anatomy of STEM teaching in North American universities. *Science*, 359(6383), pp. 1468–1470. Scopus. <https://doi.org/10.1126/science.aap8892>
- Sturtevant, H. and Wheeler, L. (2019). The STEM Faculty Instructional Barriers and Identity Survey (FIBIS): Development and exploratory results.

- International Journal of STEM Education*, 6(1), Article 1.
<https://doi.org/10.1186/s40594-019-0185-0>
- Theobald, E. J., Hill, M. J., Tran, E., Agrawal, S., Arroyo, E. N., Behling, S., Chambwe, N., Cintrón, D. L., Cooper, J. D., Dunster, G., Grummer, J. A., Hennessey, K., Hsiao, J., Iranon, N., Jones, L., Jordt, H., Keller, M., Lacey, M. E., Littlefield, C. E., ... Freeman, S. (2020). Active learning narrows achievement gaps for underrepresented students in undergraduate science, technology, engineering, and math. *Proceedings of the National Academy of Sciences*, 117(12), pp. 6476–6483. <https://doi.org/10.1073/pnas.1916903117>
- Wieman, C. E. (2014). Large-scale comparison of science teaching methods sends clear message. *Proceedings of the National Academy of Sciences*, 111(23), 8319–8320. <https://doi.org/10.1073/pnas.1407304111>
- Williams, E. A., Zwolak, J. P., Dou, R. and Brewster, E. (2019). Linking engagement and performance: The social network analysis perspective. *Physical Review Physics Education Research*, 15(2), 020150. <https://doi.org/10.1103/PhysRevPhysEducRes.15.020150>
- Wood, A. K., Galloway, R. K., Donnelly, R. and Hardy, J. (2016). Characterizing interactive engagement activities in a flipped introductory physics class. *Physical Review Physics Education Research*, 12(1), 010140. <https://doi.org/10.1103/PhysRevPhysEducRes.12.010140>
- Wood, A. K., Galloway, R. K., Hardy, J. and Sinclair, C. M. (2014). Analyzing learning during peer instruction dialogues: A resource activation framework. *Physical Review Special Topics - Physics Education Research*, 10(2). Scopus. <https://doi.org/10.1103/PhysRevSTPER.10.020107>
- Wood, R. and Shirazi, S. (2020). A systematic review of audience response systems for teaching and learning in higher education: The student experience. *Computers and Education*, 153. Scopus. <https://doi.org/10.1016/j.compedu.2020.103896>
- Zhang, P., Ding, L. and Mazur, E. (2017). Peer instruction in introductory physics: A method to bring about positive changes in students' attitudes and beliefs. *Physical Review Physics Education Research*, 13(1), 010104. <https://doi.org/10.1103/PhysRevPhysEducRes.13.010104>

Zimmerman, B. J. and Schunk, D. H. (2011). Self-regulated learning and performance: An introduction and an overview. In *Handbook of Self-Regulation of Learning and Performance* (pp. 1–12).
Routledge/Taylor & Francis Group.

APPENDIX 1

Example of an item administered to the students during a peer learning session. The correct answer is in bold.

An aeroplane runs through equal spaces in equal times

- A. only if the velocity vector is constant over time;
- B. only if its trajectory is straight;
- C. only if the tangent component of the acceleration is zero and the centripetal component is constant;
- D. only if the tangent component of the acceleration is zero.**

APPENDIX 2

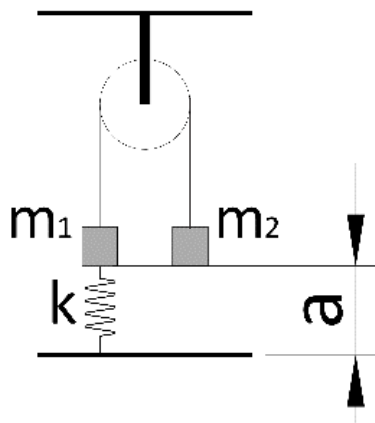
Example of a problem given to the students during their final exam (exercise 1, 21/01/2022)

Two blocks of masses $m_1 = 1 \text{ kg}$ and $m_2 = 2 \text{ kg}$ are attached to opposite ends of an ideal rope which can run through the throat of an ideal pulley, like in the figure. Block 1 (m_1) is connected to the ground through an ideal vertical spring which has a relaxed length $h_0 = 10 \text{ cm}$.

a) Determine the spring constant in such a way that the two blocks are initially at the same height $a = 15 \text{ cm}$ from the ground.

Block 2 (m_2) is lowered vertically by 5 cm and then is released:

b) Calculate the acceleration of the two blocks when block 2 is released.



APPENDIX 3

Example of an item administered to the students during the initial questionnaire. The correct answer is in bold.

The force of static friction exerted by a horizontal floor on a crate at rest on it

- A. has always a magnitude higher than the magnitude of the force of dynamic friction exerted by the horizontal floor on the crate if the latter is moving;
- B. has a magnitude directly proportional to the magnitude of the normal contact force exerted by the horizontal floor on the crate;
- C. is never null;
- D. is opposite to the sum of all other external forces lying on the horizontal plane.**

ACADEMIC PROCRASTINATION SCALE: PSYCHOMETRIC EVIDENCE FOR THE ITALIAN VERSION BASED ON CFA AND IRT MODELS

Rosa Fabbriatore

*Department of Economics and Statistics, Federico II University of Naples,
Naples, Italy*

Anna Parola

Department of Humanities, Federico II University of Naples , Naples, Italy

Giuliana Pepicelli, Francesco Palumbo

Department of Political Sciences, Federico II University of Naples, Naples, Italy

Abstract. *This study aimed to translate and validate the Academic Procrastination Scale – Short Form (APS-SF) in the Italian context. Specifically, this study focused on the psychometric properties of validity and reliability in a sample of N=340 university students. Confirmatory factor analysis (CFA) and item response theory (IRT) models were performed. CFA and IRT models showed the unifactorial model proposed in the original version of the scale. The IRT approach also provided evidence for satisfactorily test information and the measurement invariance for gender. These findings confirm that the Italian version of APS-SF is a valid and reliable measure to assess procrastination in Italian academic students.*

Keywords: *Procrastination, Higher education, Confirmatory factor analysis, Item response theory.*

1. INTRODUCTION

Procrastination refers to needlessly task delaying that is followed by the experience of subjective discomfort (Solomon and Rothblum, 1984). From a historical perspective, the term procrastination derives from the Latin adjective *crastinus*, translated as "of tomorrow"; it did not necessarily have a negative connotation until the mid-eighteenth century (Ferrari et al., 1995). Nowadays, it

is considered a dysfunctional tendency with moral connotations (Sabini and Silver, 1982) and psychological consequences as experienced guilt, decreasing performance outcomes, and social impact such as being considered unreliable. In attempting to understand the procrastination construct, researchers generally considered procrastination a stable disposition, a fixed personal trait that does not vary according to tasks, contexts, and time (Ferrari et al., 1995; Schouwenburg, 1995; Van Eerde, 2000).

Everyone may procrastinate sometimes but, according to this perspective, procrastination is defined as an intra-individual process depending by own norms and moral principles and consisting of delaying actions or decisions even if it is not appropriate to a particular situation (Burka and Yuen, 1983; Milgram et al., 1998). However, not every late-performing task is procrastination: indeed, the delay can be used as a wise strategy if purposely planned. Instead, procrastination refers to the act of postponing needless the implementation of what was planned (Van Eerde, 2003). Moreover, other studies presented empirical support for its temporal and situational variability. They demonstrated that procrastination is better described as a situation-specific behavior that changes over time and depends on the tasks and contexts interaction (e.g., Blunt and Pychyl, 2000; Lonergan and Maher, 2000; Milgram et al., 1988).

The amount of procrastination literature originated in studies with college students because they tended to delay writing term papers, preparing for tests, studying for examinations, and keeping up with weekly reading assignments (Burka and Yuen, 1983; Ellis and Knaus, 1977; Onwuegbuzie and Jiao, 2000). Several studies showed that over 70% of college students engage in procrastination (Ellis and Knaus, 1977) and regularly procrastinate (Ferrari, O'Callaghan, and Newbegin, 2005; Rabin et al., 2011). In this field of research, academic procrastination has been defined as the students' tendency to delay academic tasks intentionally, although it will carry out negative consequences (Simpson and Pychyl, 2009; Steel, 2007). Furthermore, other studies found a strong relationship between academic procrastination and academic performance (e.g., Beswick et al., 1988; Steel et al., 2001). Several scholars underlined the adverse outcomes of procrastination in (higher) educational settings such as psychological distress, anxiety, reduced well-being, low academic performance, and avoidance of social relations (Kim and Seo, 2015; Krause and Freund, 2014). In a ground theory study examining the process by which academic procrastination occurs, students involved in the research reported that they

procrastinate to manage their time better, reduce boredom, and work more efficiently. Moreover, students had a low perception that procrastination can negatively impact learning achievement (Schraw et al., 2007). Several studies investigated the relationship between academic procrastination and perceived self-efficacy. They pinpointed that perceived self-efficacy is negatively related to procrastination (Steel, 2007). Accordingly, procrastinating students presented weak self-efficacy beliefs (Lindblom et al., 2015). Instead, students' perception of a high self-efficacy in an academic context seemed to be strictly linked to a positive quality of studying and academic performance (Caprara et al., 2008).

Recent studies in the Italian academic context have shown that procrastination represented a failure in regulating students' behavior (Limone et al., 2020). Regarding the outcomes of academic procrastination, some studies show a strong relationship between procrastination and students' academic performance in secondary school (De Paola and Scoppa, 2015) and university (De Paola and Scoppa, 2015, Limone et al., 2020). Furthermore, in a recent study, Licata et al. (2024) suggested that the harms of procrastination are not limited to academic performance, but also include effects on health and well-being. In this sense, an assessment of the academic procrastination experienced by students turns out to be a good starting point for conceivable programs to enhance a student's academic success and well-being.

McCloskey (2011) proposed the first theorization of academic procrastination and the academic procrastination scale (APS), consisting of six main dimensions: psychological beliefs about abilities, attention distractions, social factors, time management skills, personal initiative, and laziness. Starting with this scale (25 items). The APS was validated and used in the Italian context by Limone et al. (2020). Yockey (2016) has tested the academic procrastination scale - short form (APS-SF) that consisted of five items. The original factor analysis showed a single-factor solution, a good factorial load of the five items, and a Cronbach's alpha of .87. The measurement invariance suggests the validity for males and females.

As things stand, the APS-SF has been tested with samples of American (Yockey, 2016) and Spanish-speaking (Brando Garrido et al., 2020) academic students. To our knowledge, there is no Italian validation study of the short form of the APS-SF. In recent years, the need for short and valid assessment instruments has been highlighted in research practice. The 25 items of the long version of the APS, despite their undoubted usefulness, do not always allow for

a short and routine assessment of academic procrastination. It can be burdensome to administer, making it difficult to include in longer assessment batteries (Moreira et al., 2020). In line with the validation of the APS-SF in other countries (e.g. Brando-Garrido et al., 2020, Yockey, 2016), the purpose of the short version of the procrastination measure is to provide a measure that can be answered by subjects in a short time, especially in contexts where the administration time cannot be unduly prolonged, in a long psychological battery with several other instruments, and to avoid fatigue or exhaustion among participants.

Therefore, this study aimed to validate the APS-SF in the Italian context by examining the psychometric properties (factor structure, item fit, convergent validity, internal consistency) of the APS-SF using a sample of university students.

2. METHODS

2.1 STUDY DESIGN

Following the guidelines for the cross-cultural adaptation of questionnaires and rating scales (Beaton et al., 2000; Van De Vijver and Hambleton, 1996), the APS-SF was translated from English to Italian by three expert psychologists, psychometry and teaching and back-translated by an independent English translator to ensure cross-cultural equivalence.

The sample size was determined according to the adaptation to the confirmatory factor analysis (CFA), taking into account the recommendations to use between 5 and 20 subjects per item, with a minimum of 100-150 subjects in the case of short scales. For item response theory (IRT) models, authors suggested that a sample of 200 or fewer subjects could be adequate for differential item functioning (DIF) detection and scale properties evaluation, especially when IRT assumptions are met (Edelen and Reeve, 2007).

2.2 PARTICIPANTS AND PROCEDURE

Data collection was conducted in 2021 using a convenience sample of 340 Italian university students enrolled in a bachelor's degree program in psychology. The sample included 72 males and 268 females ($M_{\text{age}} = 20.75$, $SD_{\text{age}} = 3.50$), primarily from high schools (12% classical high schools, 30% scientific high schools, 42% other high schools, 12% technical or professional schools, 4% other institutions).

The inclusion criteria were the minimum age of 18, the status of university students, and the Italian-language speakers. Before starting the data collection, researchers briefed the participants on the scope of the study, and their agreement was kindly requested. Students who agreed to participate completed an online self-report questionnaire via the MOODLE platform. This study complies in full with the recommendations in the ethical guidelines of the Italian Psychologists Association for research.

2.3 MEASURES

Academic procrastination scale short form (APS-SF; Yockey, 2016). The instrument consists of five items that assess academic procrastination. Participants responded to each item on a five-point Likert scale from 1 (total agreement) to 5 (complete disagreement). The total score is defined as the sum of the items (range 5-25), where higher scores indicate a greater tendency to postpone academic tasks.

Procrastination scale (TPS; Tuckman, 1991; Limone et al., 2020). The instrument consists of 16 items that refer to procrastinating behaviors. Participants responded to each item on a four-point Likert scale from 1 (strongly disagree) to 4 (strongly agree). TPS includes two dimensions of procrastination: a general explanation of procrastination and the likelihood to avoid difficult or unpleasant tasks. The score is defined as the sum of the item scores (range 16-64), where higher scores indicate a higher procrastination tendency. For this study, Cronbach's alpha was .78.

Motivated strategies for learning questionnaire (MSLQ; Pintrich and DeGroot, 1990; Bonanomi et al., 2018). The self-regulation learning strategies (SRLS) scale was administered for this study; it consists of 22 items and comprises two subscales assessing cognitive strategy use (12 items) and self-regulation (10 items). The participants responded to each item on a 5-point Likert scale from 1 (not at all appropriate for me) to 5 (very appropriate for me). The arithmetic mean of the related items is the score of the subscale. For this study, the overall Cronbach's alpha was .75.

2.3 DATA ANALYSIS

The psychometric assessment of the APS-SF (Italian version) included confirmatory factor analysis, item response theory models for differential item

functioning detection and item fit evaluation, internal consistency analysis, and validity. The classical test approach and IRT allow for a more robust psychometric analysis. In particular, the IRT is a model-based approach aiming to estimate the probability of selecting different item response categories according to the individual level of latent trait (academic procrastination in our case) and item characteristics (difficulty and discrimination). This approach is particularly advantageous in educational and psychological assessment because it provides item- and person-level insights, addressing limitations of traditional methods such as classical test theory (CTT), which assumes uniform measurement properties across items (Edelen and Reeve, 2007). Very briefly, the latent trait (usually indicated with θ) can be defined as an unobservable construct measured by means of several manifest variables (e.g., items of a Likert scale); the item difficulty refers to the level of latent trait needed to have a probability $p = .5$ of endorsing the item, and when polytomous items are considered a different difficulty parameter is estimated for each threshold (where the number of thresholds is equal to the number of levels minus 1); the item discrimination expresses the item's ability to distinguish the different levels of the latent trait in the subjects.

The IRT approach provides a collection of techniques for an accurate item- and test-level evaluation, which has proven to be highly useful in scale validation studies. Indeed, it offers robust tools for identifying items that best capture the latent trait and determining the measurement precision across different latent trait levels, which is crucial for both research and practical applications in educational assessment (Edelen and Reeve, 2007). In this theoretical framework, several models have been proposed for both binary and ordinal item responses (see Bartolucci et al., 2015, Tutz, 2020, and Van der Linden, 2018, for an overview). These models differ primarily in the constraints imposed on item parameters (e.g., free or constrained discrimination parameters across items). Additionally, for ordinal items, the relationship between observed indicators and the latent variable can be modeled using either a global (cumulative) logit or a local (adjacent categories) logit, which correspond to the graded response model (GRM) and the generalized partial credit model (GPCM), respectively.

In general, traditional IRT models are grounded on three main assumptions required to be met before the model estimation: one-dimensionality, local independence, and monotonicity. Moreover, also the measurement invariance must be ensured to reach comparable measures.

Confirmatory factor analysis (CFA) was performed to assess the factorial structure and, thus, the one-dimensionality of the latent trait. According to the response scale, the diagonal weighted least square (DWLS) estimator was used as a robust method for non-normality. For the model fit, the study considered the following indexes and test: the Chi-square test statistics, the comparative fit index (CFI), the Tucker-Lewis index (TLI), the root mean square error of approximation (RMSEA), and the standardized root mean square residual (SRMS). As a general rule, a model satisfactorily fits if the CFI and TLI coefficients are .90 or above (Medsker et al., 1994); RMSEA approximately .08 or less (Byrne, 2010), and SRMS .10 or less. The CFA was performed using the R package *lavaan* (Rosseel, 2012, v. 0.6-6).

The residual correlation matrix of the CFA model with one factor allows checking the second IRT assumption of local independence. This assumption implies that the responses to a set of items are independent, conditional to a latent trait (variable). Thus, there should be no significant correlation between item residuals (weak independence condition). About that, correlations below the threshold of .2 are assumed as evidence of independence (Chen and Thissen, 1997). Moreover, a non-parametric IRT model relying on the Mokken scaling was used to assess the monotonicity assumption, consisting of a non-decreasing probability of endorsing higher item response categories as the level of latent trait increases. The scalability coefficient H for each item and the total scale was calculated using the R package *mokken* (Van der Ark, 2007). Values greater than .30 for items and greater than .50 for the entire scale were considered acceptable for monotonicity assurance (Mokken, 1971).

Finally, according to the measurement invariance definition, item parameters are assumed to be invariant across different groups of subjects (e.g., males and females). Violations of this assumption lead to what is known as differential item functioning (DIF; Ackerman, 1992). The DIF occurs when subjects with the same latent trait level but from different groups have a different probability of endorsing the item response categories. The DIF can be consistent along with all the latent traits (uniform DIF) or vary in direction or magnitude according to different latent trait levels (non-uniform DIF). In this study, both uniform and non-uniform DIF for gender were evaluated by exploiting ordinal logistic regression models and considering a change of McFadden's pseudo R^2 equal to .02 as the critical value. The R package *Lordif* (Choi et al., 2011) was used for this aim. In addition to DIF, differential test functioning (DTF) was also

examined to assess whether there is overall bias at the test level with respect to gender. The omnibus measure proposed by Chalmers et al. (2016), referred to as *signed DTF* (sDTF), was employed. This measure identifies whether one group consistently scores higher across a defined range of θ . Values close to 0 indicate negligible or no DTF. The measure is implemented in the R package *mirt* (Chalmers, 2012).

It is worth noting that before testing for measurement invariance, the most appropriate link function (global logit or local logit) was selected by comparing the fit of the two most general models, namely the GRM (global logit) and the GPCM (local logit), through the Bayesian information criterion (BIC; Schwarz, 1978). Afterward, the model selected according to the most appropriate link function was compared with its constrained formulations (see Bartolucci et al. (2015) for a review of IRT models) to test if some item constraints should be set to describe the APS-SF properties better. The IRT models at issue were estimated using the R package *mirt* (Chalmers, 2012), and the BIC was considered the selection criterion.

Once the best model was selected, Orlando and Thissen's $S-X^2$ was used to evaluate polytomous items fitting (Kang and Chen, 2008). An $S-X^2$ p-value $\leq .001$ indicates a poor fit.

Besides, also the following graphical diagnostic tools for IRT model assessment were employed to explore deeper the goodness of the items and the scale: Item information curve (IIC), test information curve (TIC), and item response category characteristic curves (IRCCCs). These curves are a function of the item parameters and allow for the evaluation of item and test measurement precision across the entire range of the latent trait, with a particular focus on the critical area of the construct continuum (in this case, high levels of academic procrastination) (Edelen and Reeve, 2007).

The internal consistency of the factor was measured using Cronbach's alpha coefficient. The correlations between APS-SF and TPS and MLSQ were computed to test the validity of the instrument. Following the recommendations of Clark and Watson (2019), Flake et al. (2017), and testing standards (AERA, APA, and NCME, 2014), the relationships between academic procrastination and other measures, such as procrastination and self-regulated learning strategies, are used to assess the same or similar constructs and provide convergent evidence. In particular, the correlation between the APS-SF and the TPS measures the same psychological dimension. The relationship between self-regulation learning

strategies and the APS highlights the relationship between two correlated psychological dimensions (e.g. Limone et al., 2020).

3. RESULTS

Results include the estimated parameters for the APS-SF and fit indices demonstrating that the model adequately fits the data: $\chi^2(5) = 13.467$, $p = .01$; TLI = .996; CFI = .998; RMSEA = .071 [.026 - .117], SMSR = .034. These values confirmed the adequacy of the one-dimensional structure of the instrument. Item arithmetic means, standard deviations, skewness, kurtosis, and standardized factor loadings are in Table 1.

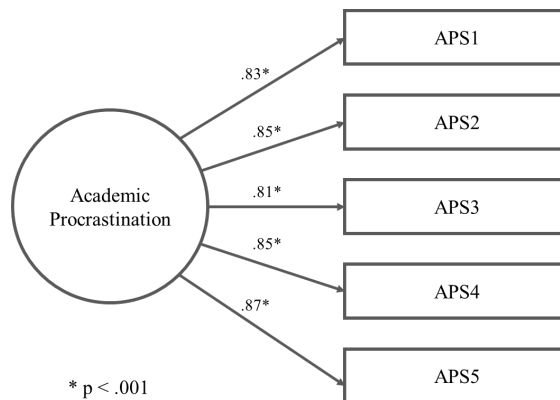
The factor loadings varied from .81 to .87 (see also Figure 1), confirming the soundness of the items used to measure the APS-SF. In the estimated model, no correlations between the residuals of different items were allowed. In a subsequent step, the residual correlation matrix was extracted to verify the IRT local independence assumption. All item residual correlations were below the critical threshold of .2, confirming that the assumption was met. The monotonicity assumption was also confirmed: the scalability coefficient H ranged between .657 and .672 for the single items (see Table 2) and was equal to .670 for the total scale. Regarding measurement invariance for gender, McFadden's pseudo- R^2 changes ranged from .0001 to .0170, detecting neither uniform nor non-uniform DIF. Moreover, a negligible DTF emerged, with sDTF = .48 (95% CI: -.28 to 1.19), which was not significantly different from 0 (p -value = .21). At this step, the GRM (global logit) shows a better fit than the GPCM according to the BIC (3672.09 and 3680.86, respectively).

Table 1: Means, standard deviations of each item, and results of CFA in university students

Item	M	SD	SK	K	(λ)	R ²
I put off projects until the last minute (APS1).	2.33	1.17	0.46	-0.73	.83*	.69
I know I should work on schoolwork, but I just don't do it (APS2).	1.79	1.04	1.10	0.24	.85*	.72
I get distracted by other, more fun, things when I am supposed to work on schoolwork (APS3).	2.47	1.14	0.40	-0.54	.81*	.65
When given an assignment, I usually put it away and forget about it until it is almost due (APS4).	1.56	0.88	1.81	3.29	.85*	.72
I frequently find myself putting important deadlines off (APS5).	1.74	1.08	1.41	1.11	.87*	.76

Note. M = mean, SD = standard deviation, SK= Skewness, K= Kurtosis, λ = factor loading.

* $p < .001$

**Figure 1: CFA results: Standardized estimates of the factor loadings**

Afterward, by comparing GRM with its constrained formulations, namely the rating scale GRM (RS-GRM; Muraki, 1990), the one-parameter GRM (1P-GRM; Van der Ark, 2001), and the one-parameter RS-GRM (1P-RS-GRM; Van der Ark, 2001), the BIC indicated that the RS-GRM was the best model. In the RS-GRM, the item discrimination parameter can freely vary among items, whereas the item difficulty parameter is expressed in an additive way so that items have a different general difficulty level but equal threshold difficulty level. Table 2 reports the estimated parameters for the Italian version of the APS-SF.

Table 2: Scalability coefficient (H), RS-GRM item parameters, and model fit ($S-X^2$ p-value)

Item	H	$S-X^2$ p-value	RS-GRM item parameters					
			a	c	b_1	b_2	b_3	b_4
APS1	.672	.07	2.55	0	1.52	-0.55	-2.93	-5.12
APS2	.664	.02	2.79	-1.81	1.52	-0.55	-2.93	-5.12
APS3	.657	.64	2.21	0.39	1.52	-0.55	-2.93	-5.12
APS4	.666	.05	2.69	-2.67	1.52	-0.55	-2.93	-5.12
APS5	.672	.03	3.24	-2.24	1.52	-0.55	-2.93	-5.12

Note. a = item discrimination; c = item general difficulty; b_1 , b_2 , b_3 , b_4 = threshold difficulty; $c_1 = 0$ for model identifiability.

Even for a low level of procrastination, the easiest item to endorse was Item 4 ("When given an assignment, I usually put it away and forget about it until it is almost due"). In contrast, Item 3 ("I get distracted by other, more fun, things when I am supposed to work on schoolwork") was the most difficult item that requires a high level of procrastination to encounter agreement. On the other hand, Item 5 ("I frequently find myself putting important deadlines off") allows discriminating different levels of procrastination better than the other items do.

Orlando and Thissen's $S-X^2$ index suggested no misfitting items (see Table 2). The item information curves in Figure 2b show that all items are centered on medium-high values of the latent trait and that the most informative one is Item 5, whereas Item 3 is the least informative. Overall, the information provided by the APS-SF scale is satisfactory and maximum for students with high levels of procrastination (the curve in Figure 2a is shifted on the right).

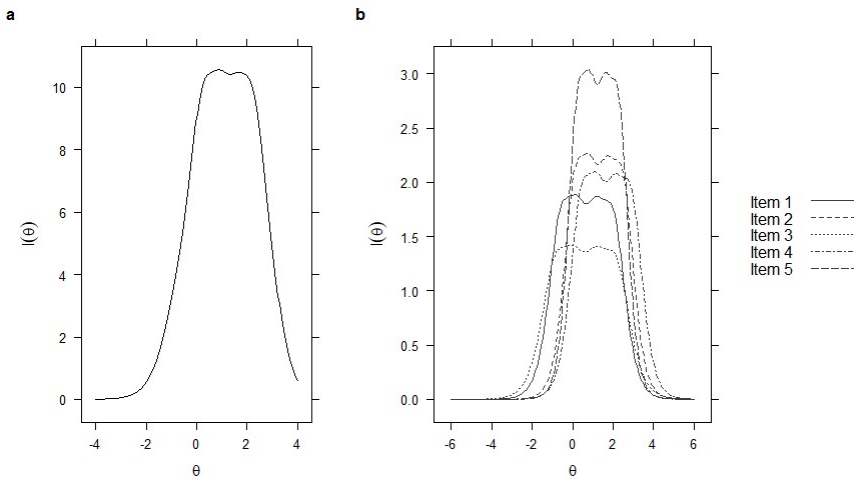


Figure 2: Test (a) and item (b) information curves

Finally, the item response category characteristic curves (IRCCCs) in Figure 3 are well-shaped and the response categories well-separated, indicating that there are no redundant response categories needed to be merged. Moreover, for Item 4 and Item 5, students tend to strongly disagree also for a medium level of latent trait (see the IRCCC for the first category, represented by the solid line in the figure).

Cronbach's alpha was .88, confirming the reliability of the measure. The Italian form of APS-SF also showed adequate convergent validity correlated with APS (see Table 3). Moreover, APS-SF and APS showed the same negative relationship with SRLS.

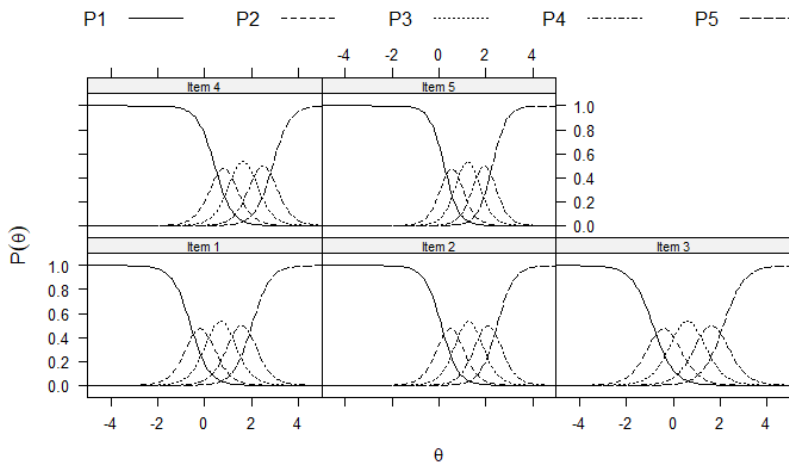


Figure 3: Item response category characteristic curves (IRCCCs): P1 = “Strongly disagree”, P2 = “Disagree”, P3 = “Do not disagree or agree”, P4 = “Agree”, P5 = “Strongly agree”.

Table 3: Means, standard deviation, and bivariate correlations

Measures	<i>M</i>	<i>SD</i>	1	2
1. APS-SF	9.89	4.37	-	
2. TPS	33.97	7.21	.76**	-
3. SRLS	3.62	.40	-.27**	-.16*

Note. *M* = means, *SD* = standard deviations, APS-SF: Academic procrastination scale – short form, APS: Tuckman procrastination scale, SRLS: Self regulation learning strategies.

* $p < .01$, ** $p < .001$

4. DISCUSSION

This study aimed to test the Italian version of the academic procrastination scale – short form psychometric characteristics. The confirmatory factor analyses confirmed the one-dimensional structure of the Italian version of the APS-SF, supporting its use in the Italian context. Cronbach's alpha indicated that the internal consistency of the APS-SF was satisfactory. The association between APS-SF, APS, and MLSQ were further investigated. The results confirmed the convergent validity of the APS-SF, showing the correlation between the APS-SF and APS and the negative relationship between APS-SF and MLSQ.

Item and test proprieties investigation through IRT models showed that the RS-GRM was the best model for describing the Italian version of the APS-SF. Both item and test information were satisfactory and maximum for students with high levels of procrastination. All items reported a more than satisfactory fit to the model; the response categories were well-separated and not redundant. Regarding the measurement invariance for gender, results provided evidence for the absence of uniform and non-uniform DIF, ensuring comparable measures for males and females.

Although the current study contributed to the literature by testing the psychometric properties of the Italian form of the APS-SF, some limitations and future perspectives have to be highlighted. A first limitation regards subjects involved in the study that mainly were female-gendered and homogenously aged, so further studies with wider and more varied samples are awaited. Moreover, the study was carried out only through a single-time assessment of students' procrastination. Therefore, future research will focus on longitudinal studies that would explore students' tendency to procrastinate throughout the whole academic year and the entire enrollment time, hopefully.

One of the limitations of this study was not to consider the relationships between academic procrastination and other variables linked to this construct. Further research should also explore these relationships in-depth, taking into account related psychological variables (such as self-efficacy, self-regulated learning strategies, self-esteem, anxiety, and stress), aiming to assess the procrastination construct in different and more varied conditions.

In conclusion, our findings demonstrate that the APS-SF is a psychometrically valid measure of procrastination in university students. In terms of practical implication, the instrument can be used as a valid alternative

to the long questionnaires composed of different scales, in particular, to reduce the risk of fatigue of the participants. Thanks to its length, this scale can be easily integrated into surveys or batteries of instruments.

Finally, the results of this study demonstrate that researchers can reliably use the APS-SF in future studies to evaluate academic procrastination and explore its associated factors, thereby gaining a deeper understanding of this behavior. Moreover, this measure of procrastination could also be beneficial for learners and teachers in educational contexts. Indeed, as highlighted in Martín-Puga et al. (2022), the APS-SF serves as a useful tool to identify students prone to procrastination, guiding the development of targeted interventions to help them manage this behavior and thus enhance their likelihood of achieving positive academic outcomes.

ACKNOWLEDGMENTS

Rosa Fabbriatore acknowledges the financial support provided by the European Union - NextGenerationEU, in the framework of the GRINS - Growing Resilient, Inclusive and Sustainable project (GRINS PE00000018 - CUP E63C22002140007). The contribution of Giovanbattista Califano was funded by the European Union - NextGenerationEU (Call PRIN 2022 - Italian Ministry of University and Research - RECIPE - cod. 20224ZHNXE). The views and opinions expressed are solely those of the authors and do not necessarily reflect those of the European Union, nor can the European Union be held responsible for them.

REFERENCES

- Ackerman, T.A. (1992). A didactic explanation of item bias, item impact, and item validity from a multidimensional perspective. *Journal of Educational Measurement*. 29, 67-91.
- American Educational Research Association, American Psychological Association, National Council on Measurement in Education, & Joint Committee on Standards for Educational and Psychological Testing (AERA, APA, & NCME) (2014). *Standards for Educational and Psychological Testing*. Washington, DC: American Educational Research Association.

- Bartolucci, F., Bacci, S. and Gnaldi, M. (2015). *Statistical Analysis of Questionnaires: A Unified Approach Based on R and Stata*. Chapman & Hall, CRC Press, Boca Raton, FL, USA.
- Beaton, D. E., Bombardier, C., Guillemin, F. and Ferraz, M. B. (2000). Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine*. 25, 3186-3191.
- Beswick, G., Rothblum, E. D. and Mann, L. (1988). Psychological antecedents of student procrastination. *Australian Psychologist*. 23, 207-217. <https://doi.org/10.1080/00050068808255605>
- Blunt, A. K. and Pychyl, T. A. (2000). Task aversiveness and procrastination: A multidimensional approach to task aversiveness across stages of personal projects. *Personality and Individual Differences*. 28, 153-167. [https://doi.org/10.1016/S0191-8869\(99\)00091-4](https://doi.org/10.1016/S0191-8869(99)00091-4)
- Bonanomi, A., Olivari, M. G., Mascheroni, E., Gatti, E. and Confalonieri, E. (2018). Using a multidimensional Rasch analysis to evaluate the psychometric properties of the motivated strategies for learning questionnaire (MSLQ) among high school students. *Testing, Psychometrics, Methodology in Applied Psychology*. 25, 83-100. <https://doi.org/10.4473/TPM25.1.5>
- Brando-Garrido, C., Montes-Hidalgo, J., Limonero, J. T., Gómez-Romero, M. J. and Tomás-Sábado, J. (2020). Academic procrastination in nursing students. Spanish adaptation of the Academic Procrastination Scale-Short Form (APS-SF). *Enfermería Clínica (English Edition)*. 30, 371-376.
- Burka, J. B. and Yuen, L. M. (1983). *Procrastination*. Reading, Addison-Wesley, Boston.
- Byrne, B. M. (2010). *Structural Equation Modeling with AMOS. Basic Concepts, Applications, and Programming (2nd ed.)*. Routledge Academic, London.
- Caprara, G. V., Fida, R., Vecchione, M., Del Bove, G., Vecchio, G. M., Barbaranelli, C. and Bandura, A. (2008). Longitudinal analysis of the role of perceived self-efficacy for self-regulated learning in academic continuance and achievement. *Journal of Educational Psychology*. 100, 525-534. <http://dx.doi.org/10.1037/0022-0663.100.3.525>.
- Chalmers, R.P. (2012). mirt: A multidimensional item response theory package for the R environment. *Journal of Statistical Software*. 48, 1-29.
- Chalmers, R. P., Counsell, A. and Flora, D. B. (2016). It might not make a big DIF: Improved differential test functioning statistics that account for

- sampling variability. *Educational and Psychological Measurement*. 76, 114-140.
- Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling*. 14, 464-504.
<https://doi.org/10.1080/10705510701301834>
- Chen, W.H. and Thissen, D. (1997). Local dependence indexes for item pairs using item response theory. *Journal of Educational and Behavioral Statistics*. 22, 265-289.
- Choi, S. W., Gibbons, L. E. and Crane, P. K. (2011). Lordif: An R package for detecting differential item functioning using iterative hybrid ordinal logistic regression/item response theory and Monte Carlo simulations. *Journal of Statistical Software*. 39, 1-30.
- Clark, L. A. and Watson, D. (2019). Constructing validity: New developments in creating objective measuring instruments. *Psychological Assessment*. 31, 1412.
- De Paola, M. and Scoppa, V. (2015). Procrastination, academic success and the effectiveness of a remedial program. *Journal of Economic Behavior & Organization*. 115, 217-236.
- Edelen, M.O. and Reeve, B.B. (2007). Applying item response theory (IRT) modeling to questionnaire development, evaluation, and refinement. *Quality of Life Research*. 16, 5-18.
- Ellis, A. and Knaus, W. J. (1977). Overcoming procrastination (New York, Institute for Rational Living). *Scientific Research*. 1, 1-11.
- Ferrari, J. R., Johnson, J. L. and McCown, W. G. (1995). *Procrastination and Task Avoidance: Theory, Research, and Treatment*. Plenum Press, New York.
- Ferrari, J. R., O'Callaghan, J. and Newbegin, I. (2005). Prevalence of procrastination in the United States, United Kingdom, and Australia: Arousal and avoidance delays among adults. *North American Journal of Psychology*. 7, 1-6.
- Flake, J. K., Pek, J. and Hehman, E. (2017). Construct validation in social and personality research: Current practice and recommendations. *Social Psychological and Personality Science*. 8, 370-378.
- Interview study. *Frontline Learning Research*. 3, 47-62.
<https://doi.org/10.14786/flr.v3i2.154>

- Kang, T. and Chen, T. T. (2008). Performance of the generalized S-X² item fit index for polytomous IRT models. *Journal of Educational Measurement*. 45, 391-406.
- Kim, K. R. and Seo, E. H. (2015). The relationship between procrastination and academic performance: A meta-analysis. *Personality and Individual Differences*. 82, 26-33. doi:10.1016/j.paid.2015.02.038.
- Krause, K. and Freund, A. M. (2014). Delay or procrastination – A comparison of self-report and behavioral measures of procrastination and their impact on effective well-being. *Personality and Individual Differences*. 63, 75–80. doi:10.1016/j.paid.2014.01.050
- Licata, F., Citrino, E. A., Maruca, R., Di Gennaro, G. and Bianco, A. (2024). Procrastination and risky health behaviors: A possible way to nurture health promotion among young adults in Italy. *Frontiers in Public Health*. 12, 1432763.
- Limone, P., Sinatra, M., Ceglie, F. and Monacis, L. (2020). Examining procrastination among university students through the lens of the self-regulated learning model. *Behavioral Sciences*. 10, 184. <https://doi.org/10.3390/bs10120184>
- Lindblom-Ylänne, S. A., Saariaho-Räsänen, E. J., Inkinen, M. S., Haarala-Muhonen, A. E. and Hailikari, T. K. (2015). Academic procrastinators, strategic delayers and something betwixt and between: An interview study. *Frontline Learning Research*. 3, 47-62.
- Lonergan, J. M. and Maher, K. J. (2000). The relationship between job characteristics and workplace procrastination as moderated by locus of control. *Journal of Social Behavior and Personality*. 15, 213-224.
- Martín-Puga, M. E., Pelegrina, S., Gómez-Pérez, M. M. and Justicia-Galiano, M. J. (2022). Psychometric properties and measurement invariance of the academic procrastination scale-short form in Spanish children and adolescents. *Journal of Psychoeducational Assessment*. 40, 880-894.
- Medsker, G. J., Williams, L. J. and Holahan, P. J. (1994). A review of current practices for evaluating causal models in organizational behavior and human resources management research. *Journal of Management*. 20, 439-464. <https://doi.org/10.1177/014920639402000207>
- Milgram, N. A., Sroloff, B. and Rosenbaum, M. (1988). The procrastination of everyday life. *Journal of Research in Personality*. 22, 197-212. [https://doi.org/10.1016/0092-6566\(88\)90015-3](https://doi.org/10.1016/0092-6566(88)90015-3)

- Milgram, N., Mey-Tal, G. and Levison, Y. (1998). Procrastination, generalized or specific, in college students and their parents. *Personality and Individual Differences*. 25, 297-316. [https://doi.org/10.1016/S0191-8869\(98\)00044-0](https://doi.org/10.1016/S0191-8869(98)00044-0)
- Mokken, R.J. (1971). *Theory and Procedure of Scale Analysis: With Applications in Political Research*. The Hague, Mouton.
- Muraki, E. (1990). Fitting a polytomous item response model to Likert-type data. *Applied Psychological Measurement*. 14, 59-71.
- Onwuegbuzie, A. J. and Jiao, Q. G. (2000). I'll go to the library later: The relationship between academic procrastination and library anxiety. *College & Research Libraries*. 61, 45-54. <https://doi.org/10.5860/crl.61.1.45>.
- Pintrich, P. R. and De Groot, E. V. (1990). Motivational and self-regulated learning components of classroom academic performance. *Journal of Educational Psychology*. 82, 33-40. <https://doi.org/10.1037/0022-0663.82.1.33>
- Rabin, L. A., Fogel, J. and Nutter-Upham, K. E. (2011). Academic procrastination in college students: The role of self-reported executive function. *Journal of Clinical and Experimental Neuropsychology*. 33, 344–357. <https://doi.org/10.1080/13803395.2010.518597>.
- Rosseel, Y. (2012). Lavaan: An R package for structural equation modeling and more. Version 0.5–12 (BETA). *Journal of Statistical Software*. 48, 1-36.
- Sabini, J., and Silver, M. (1982). *Moralities of Everyday Life*. Oxford University Press, Oxford.
- Schouwenburg, H. C. (1995). Academic procrastination: Theoretical notions, measurement, and research. In J. R. Ferrari, J. L. Johnson and W. G. McCown (Eds.). *Procrastination and Task Avoidance: Theory, Research, and Treatment*. Plenum Press, New York, 71-96.
- Schraw, G., Wadkins, T. and Olafson, L. (2007). Doing the things we do: A grounded theory of academic procrastination. *Journal of Educational Psychology*. 99, 12-25. <https://doi.org/10.1037/0022-0663.99.1.12>.
- Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*. 6, 461-464.
- Simpson, W. K. and Pychyl, T. A. (2009). In search of the arousal procrastinator: Investigating the relation between procrastination, arousal-based personality traits and beliefs about procrastination motivations.

- Personality and Individual Differences*. 47, 906-911. <https://doi.org/10.1016/j.paid.2009.07.013>.
- Solomon, L. J. and Rothblum, E. D. (1984). Academic procrastination: Frequency and cognitive-behavioral correlates. *Journal of Counseling Psychology*. 31, 503-509. <https://doi.org/10.1037/0022-0167.31.4.503>
- Steel, P. (2007). The nature of procrastination: A meta-analytic and theoretical review of quintessential self-regulatory failure. *Psychological Bulletin*. 133, 65-94. <https://doi.org/10.1037/0033-2909.133.1.65>.
- Steel, P., Brothen, T. and Wambach, C. (2001). Procrastination and personality, performance, and mood. *Personality and Individual Differences*. 30, 95-106. [https://doi.org/10.1016/S0191-8869\(00\)00013-1](https://doi.org/10.1016/S0191-8869(00)00013-1)
- Tuckman, B. W. (1991). The development and concurrent validity of the procrastination scale. *Educational and Psychological Measurement*. 51, 473-480. <https://doi.org/10.1177/0013164491512022>
- Tutz, G. (2020). A taxonomy of polytomous item response models. arXiv preprint arXiv:2010.01382.
- Van der Ark, L.A. (2007). Mokken scale analysis in R. *Journal of Statistical Software*. 20, 1-19.
- Van der Ark, L.A. (2001). Relationships and properties of polytomous item response theory models. *Applied Psychological Measurement*. 25, 273-282.
- Van Eerde, W. (2000). Procrastination: Self-regulation in initiating aversive goals. *Applied Psychology: An International Review*. 49, 372-389. <https://doi.org/10.1111/1464-0597.00021>
- Van Eerde, W. (2003). A meta-analytically derived nomological network of procrastination. *Personality and Individual Differences*. 35, 1401-1418. [https://doi.org/10.1016/S0191-8869\(02\)00358-6](https://doi.org/10.1016/S0191-8869(02)00358-6)
- Van der Linden, W. J. (Ed.). (2018). *Handbook of Item Response Theory: Three Volume Set*. CRC Press.
- Van de Vijver, F. and Hambleton, R. K. (1996). Translating tests. *European Psychologist*. 1, 89-99. <https://doi.org/10.1027/1016-9040.1.2.89>
- Yockey, R. D. (2016). Validation of the short form of the academic procrastination scale. *Psychological Reports*. 118, 171-179. <https://doi.org/10.1177/0033294115626825>